



Sigillum Universitatis Ludovici Maximiliani

# Component separation methods for CMB data

Dissertation der Fakultät für Physik

D F P

der Ludwig-Maximilians-Universität München

L M U M

für den Grad des

Doctor rerum naturalium

vorgelegt von Maria-Paola Bottino

aus Genua (Italien)

München, 06.10.2010

---



Sigillum Universitatis Ludovici Maximiliani

**1. Gutachter: Prof. Dr. Simon D. M. White**

**2. Gutachter: Prof. Dr. Jochen Weller**

**Tag der mündlichen Prüfung: 03.12.2010**

*To André*



# Contents

<b>Contents</b>	<b>iii</b>
<b>Zusammenfassung</b>	<b>xv</b>
<b>Abstract</b>	<b>xvii</b>
<b>1 The microwave emissions of the sky</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 The Cosmic Microwave Background radiation . . . . .	1
1.2.1 Spectral distortions . . . . .	3
1.2.2 CMB anisotropies and the inflationary scenario . . . . .	4
1.2.3 CMB polarization . . . . .	9
1.2.4 Power spectrum, cosmological parameters and cosmic variance . . . . .	12
1.3 Non-Gaussianity . . . . .	14
1.3.1 North-south asymmetry . . . . .	15
1.3.2 Amplitude of the quadrupole . . . . .	15
1.3.3 Alignment of the low multipoles . . . . .	16
1.4 Galactic and extra-Galactic components . . . . .	16
1.4.1 Synchrotron emission . . . . .	18
1.4.2 Free-free emission . . . . .	22
1.4.3 Thermal dust emission . . . . .	25
1.4.4 Anomalous Dust emission . . . . .	28
1.4.5 The <i>WMAP</i> Haze . . . . .	31
1.4.6 Extragalactic emissions . . . . .	34
1.5 Foreground models . . . . .	39
1.5.1 The Planck Sky Model . . . . .	39

1.5.2	The foreground model based on MEM solutions . . . . .	40
1.5.3	The foreground model proposed by Ghosh <i>et al.</i> . . . . .	41
1.5.4	Comparison of foreground models . . . . .	41
<b>2</b>	<b>Methods for cleaning CMB data and the algorithm</b>	<b>45</b>
2.1	Introduction . . . . .	45
2.2	Cross-correlations method . . . . .	46
2.3	Internal Linear Combination (ILC) method . . . . .	47
2.4	Component separation methods . . . . .	50
2.4.1	The Independent Component Analysis . . . . .	51
2.4.2	The pre-processing phase . . . . .	52
2.4.3	for CMB data analysis: CMB cleaning and foregrounds analysis . . . . .	54
2.4.4	: CMB prior and loss-less input data compressing . . . . .	56
<b>3</b>	<b>Foreground analysis of the <i>WMAP</i> three-year data with</b>	<b>59</b>
3.1	Introduction . . . . .	59
3.2	and its use for foreground component studies . . . . .	59
3.3	Data used in the analysis . . . . .	61
3.4	Monte Carlo simulations and calibration of method . . . . .	62
3.5	Analysis with the Haslam map as synchrotron template . . . . .	62
3.5.1	Spectral index of foreground emissions. . . . .	65
3.5.2	Single Year Analysis . . . . .	68
3.6	Analysis with the K-Ka map as synchrotron template . . . . .	68
3.6.1	Spectral properties of foreground emissions . . . . .	71
3.7	Evaluation of the Q-, V- and W-band sky maps after foreground cleaning . . . . .	71
3.8	An ‘iterative’ blind component separation study using <i>WMAP</i> data pre-cleaned using templates . . . . .	74
3.8.1	Foreground residuals . . . . .	74
3.8.2	CMB component . . . . .	76
3.9	Discussion . . . . .	78
<b>4</b>	<b>New insights into foreground analysis of the <i>WMAP</i> five-year data using</b>	<b>81</b>

4.1	Introduction . . . . .	81
4.2	Data used in the analysis . . . . .	82
4.3	Monte Carlo simulations . . . . .	83
4.4	Coupling coefficients between data and templates . . . . .	84
4.4.1	Spectral index of foreground emissions . . . . .	88
4.5	Evaluation of the <i>WMAP</i> sky maps after foreground cleaning . . . . .	91
4.6	Iterative application of . . . . . on pre-cleaned data . . . . .	95
4.6.1	Is the foreground residual the ‘ <i>WMAP</i> haze’? . . . . .	97
4.7	Power-spectrum evaluation of iteratively cleaned maps . . . . .	99
4.7.1	Partial-sky analysis . . . . .	100
4.7.2	Full-sky analysis . . . . .	101
4.8	Application of . . . . . on <i>WMAP</i> data and templates simultaneously . . . . .	104
4.8.1	What is the minimal mask? . . . . .	107
4.9	Discussion . . . . .	108
<b>5</b>	<b>A refined study of the free-free frequency spectrum</b>	<b>113</b>
5.1	Introduction . . . . .	113
5.2	Fitting procedure used by Dobler <i>et al.</i> . . . . .	113
5.3	Monte Carlo simulations . . . . .	116
5.3.1	Simulations results . . . . .	118
5.3.2	Fit parameters . . . . .	121
5.3.3	Noise impact simulations . . . . .	129
5.4	Real data analysis . . . . .	130
5.4.1	Dependence on the CMB estimation . . . . .	133
5.5	Discussion . . . . .	136
<b>6</b>	<b>Internal analysis of the <i>WMAP</i> five-year data on patches of the sky.</b>	<b>139</b>
6.1	Introduction . . . . .	139
6.1.1	Partitioning of the sky . . . . .	140
6.2	Method of analysis . . . . .	143
6.3	Monte Carlo simulations . . . . .	143
6.3.1	Statistical study of the simulations results . . . . .	152

6.3.2	Local evaluation of skewness and kurtosis . . . . .	154
6.4	Analysis of the <i>WMAP</i> five-year data . . . . .	161
6.4.1	Power spectrum . . . . .	162
6.4.2	Variance . . . . .	168
6.4.3	Local skewness and kurtosis . . . . .	174
6.4.4	Further non-Gaussianity tests . . . . .	174
6.5	ILC: a deeper study of the effect of partitioning the sky . . . . .	183
6.6	Bias correction of the CMB estimation . . . . .	187
6.6.1	Statistical properties of the bias corrected CMB maps . . . . .	190
6.7	Discussion . . . . .	193
<b>7</b>	<b>Conclusions</b>	<b>199</b>
7.1	Unsolved issues and prospects . . . . .	199
7.1.1	Is the anomalous dust component actually due to spinning dust? . . . . .	199
7.1.2	Do we really understand the free-free emission and its spectrum? . . . . .	201
7.1.3	What is the nature of the <i>WMAP</i> Haze? . . . . .	203
7.1.4	Will observations in polarisation help us to improve the component separation ?	204
7.1.5	What is the optimal use of . . . . . and ILC, and how could they be improved? .	204
<b>A</b>	<b>Monte Carlo simulations</b>	<b>207</b>
	<b>Acknowledgements</b>	<b>215</b>

# List of Figures

List of Figures	vii
1.1 CMB frequency spectrum . . . . .	3
1.2 Polarization from Thomson scattering of quadrupole temperature anisotropies . . . . .	11
1.3 Temperature and polarization CMB angular spectra . . . . .	12
1.4 <i>WMAP</i> cosmological parameters . . . . .	13
1.5 Sensitivity of the acoustic peaks to cosmological parameters . . . . .	14
1.6 Maps of the ILC quadrupole and octopole . . . . .	16
1.7 Theoretical frequency spectrum of the foregrounds . . . . .	17
1.8 Haslam map at 408 MHz . . . . .	18
1.9 $H\alpha$ map created by Finkbeiner (2003) . . . . .	24
1.10 Finkbeiner et al. (1999) thermal dust map at 94 GHz . . . . .	26
1.11 Composite spectrum of the dust in the ISM . . . . .	26
1.12 The spinning dust spectrum . . . . .	30
1.13 Comparison between observed and predicted differential source counts . . . . .	35
1.14 Differential source counts . . . . .	36
1.15 Foreground angular power spectra between 30 and 100 GHz . . . . .	37
1.16 Foreground angular power spectra between 143 and 857 GHz . . . . .	38
1.17 MEM maps of the integrated foreground components . . . . .	42
1.18 Difference between the <i>WMAP</i> data and the PSM foreground maps . . . . .	42
1.19 Difference between the <i>WMAP</i> data and the foreground MEM maps . . . . .	43
1.20 Difference between the <i>WMAP</i> data and the GFM foreground maps . . . . .	43
1.21 Difference between the MEM and GFM maps . . . . .	44
2.1 Color-coded map of the 12 regions of the ILC map . . . . .	49

3.1	Scaling factors obtained from <i>WMAP</i> data with and without the thermal dust emission .	65
3.2	Free-free coupling coefficients in intensity units . . . . .	69
3.3	Scaling factors determined for the yearly <i>WMAP</i> sky maps . . . . .	70
3.4	<i>WMAP</i> data cleaned by subtracting the foreground emissions . . . . .	72
3.5	<i>WMAP</i> data for Q-, V- and W-bands cleaned from foreground emissions . . . . .	73
3.6	Difference between the cleaned maps using the Haslam map and the K-Ka map . . . . .	73
3.7	Binned power spectra of the Q-, V- and W-maps cleaned from the foregrounds . . . . .	75
3.8	Residual component maps obtained by . . . . . with cleaned <i>WMAP</i> data . . . . .	76
3.9	CMB maps obtained by . . . . . with <i>WMAP</i> Q-, V- and W-band cleaned data . . . . .	77
3.10	Power spectra of CMB components from Q-, V- and W-band cleaned data . . . . .	77
4.1	Histogram of the coupling coefficient distributions at K-band from simulations . . . . .	83
4.2	Free-free coupling coefficients in intensity units . . . . .	91
4.3	<i>WMAP</i> Q-, V- and W-band cleaned data . . . . .	93
4.4	<i>WMAP</i> Q-, V- and W-band cleaned data using the K-Ka template . . . . .	94
4.5	. . . . . maps of the residual component from the cleaned K-, Ka-, Q-, V-, and W-band data . . . . .	98
4.6	. . . . . maps of the residuals from the cleaned Q-, V-, and W-band data . . . . .	99
4.7	. . . . . maps of the residuals from the cleaned Q-, V-, and W-band data and K-Ka template	100
4.8	Maps of the residual component recovered by . . . . . with and without Haze . . . . .	100
4.9	. . . . . CMB components recovered from cleaned <i>WMAP</i> data . . . . .	102
4.10	Subset of power spectra of the . . . . . CMB components from the iterative analysis . . . . .	103
4.11	Power Spectra of the full-sky CMB components recovered by . . . . .	104
4.12	CMB maps returned by . . . . . with different combinations of data . . . . .	106
4.13	Full-sky power spectra of the full-sky CMB maps . . . . .	106
4.14	Power spectra of the CMB map derived from the five <i>WMAP</i> maps together with three foregrounds . . . . .	107
4.15	Comparison of CMB power spectra . . . . .	108
4.16	Power spectra of the . . . . . , ILC, HILC and NILC maps . . . . .	109
4.17	Minimal mask . . . . .	109
5.1	Dobler et al. (2009) mask . . . . .	114

5.2	Spinning dust spectra for several environment conditions . . . . .	117
5.3	Simulations without WIM emission: weight distributions . . . . .	119
5.4	Simulations with WIM emission: weight distributions . . . . .	120
5.5	Mean spectral behaviour of the synchrotron emission from simulations . . . . .	122
5.6	Mean spectral behaviour of the dust emission from simulations . . . . .	123
5.7	Mean spectral behaviour of the free-free emission from simulations . . . . .	124
5.8	Free-free coupling coefficients . . . . .	131
5.9	Free-free coupling coefficients from <i>WMAP</i> data without CMB emission . . . . .	132
5.10	Free-free coupling coefficients following Dobler et al. (2009) prescriptions . . . . .	134
5.11	Difference between the ILC maps adopted by Dobler et al. (2009) and the <i>map</i> .	136
6.1	Full-sky maps color-coded to show the different sets of regions that were used to generate the <i>map</i> and ILC CMB maps . . . . .	141
6.2	<i>map</i> of the integrated foreground components . . . . .	142
6.3	Sky partitions proposed by Hinshaw et al. (2007) and Park et al. (2007) . . . . .	143
6.4	Statistical distribution of the CMB weights from PSM simulations: region 1 . . . . .	145
6.5	Statistical distribution of the CMB weights from PSM simulations: region 2 . . . . .	145
6.6	Statistical distribution of the CMB weights from PSM simulations: region 3 . . . . .	146
6.7	Statistical distribution of the CMB weights from MEM simulations: region 1 . . . . .	146
6.8	Statistical distribution of the CMB weights from MEM simulations: region 2 . . . . .	147
6.9	Statistical distribution of the CMB weights from MEM simulations: region 3 . . . . .	147
6.10	Statistical distribution of the CMB weights from GFM simulations: region 1 . . . . .	148
6.11	Statistical distribution of the CMB weights from GFM simulations: region 2 . . . . .	148
6.12	Statistical distribution of the CMB weights from GFM simulations: region 3 . . . . .	149
6.13	Foreground residual maps from simulations using a parallel cut along the Galactic plane	150
6.14	Foreground residual maps from simulations using three regions . . . . .	151
6.15	Comparison of the bias maps . . . . .	152
6.16	Scatter plots of the statistics of the CMB maps from PSM simulations . . . . .	156
6.17	Scatter plots of the statistics of the CMB maps from MEM simulations . . . . .	157
6.18	Scatter plots of the statistics of the CMB maps from GFM simulations . . . . .	158
6.19	Skewness mean maps computed from the CMB maps derived from <i>simulations</i> .	159

6.20 Kurtosis mean maps computed from the CMB maps derived from	simulations	. . . 159
6.21 Skewness mean maps computed from the CMB maps derived from ILC simulations		. . . 160
6.22 Kurtosis mean maps computed from the CMB maps derived from ILC simulations		. . . 160
6.23 The CMB components returned by	with three regions	. . . . . 162
6.24 The CMB components returned by	with six regions	. . . . . 165
6.25 The CMB components returned by	with nine regions	. . . . . 166
6.26 The CMB components returned by	with ten regions	. . . . . 167
6.27 Merged CMB maps from	and ILC regional analysis	. . . . . 168
6.28 Difference between the merged CMB maps from	and ILC regional analysis	. . . 168
6.29 ILC CMB maps derived with 12 and 400 regions		. . . . . 169
6.30 Full-sky power spectrum of the merged CMB maps		. . . . . 169
6.31 Power spectrum of the merged CMB maps applying a minimal mask		. . . . . 170
6.32 Auto- and cross-power spectra of the merged CMB maps		. . . . . 170
6.33 Comparison between the power spectra of the	and ILC maps	. . . . . 171
6.34 Maps of the local skewness and kurtosis estimation with $N_{side} = 32$		. . . . . 175
6.35 Maps of the local skewness and kurtosis estimation with $N_{side} = 8$		. . . . . 176
6.36 Symmetry statistic S-maps		. . . . . 177
6.37 Color-coded phase gradients of the CMB maps derived with ILC		. . . . . 179
6.38 Color-coded phase gradients of the CMB maps derived with ILC		. . . . . 180
6.39 Color-coded phase gradients of the CMB maps derived with		. . . . . 181
6.40 Color-coded phase gradients of the CMB maps derived with		. . . . . 182
6.41	analysis: scatter plots of the maximum values of the T- and D-statistics	. . . . . 184
6.42 ILC analysis: scatter plots of the maximum values of the T- and D-statistics		. . . . . 185
6.43 ILC CMB maps derived using 12, 48, 192, and 768 regions		. . . . . 187
6.44 Maps of the relative variance and cross-variance		. . . . . 188
6.45 Scatter plots of the variance and cross-variance maps		. . . . . 189
6.46 CMB maps derived with	and several partitions of the sky, removing the bias	. . . 190
6.47 CMB maps derived with ILC and several partitions of the sky, removing the bias		. . . . 191
6.48 Local skewness and kurtosis maps of the bias cleaned CMB maps using the PSM		. . . . 193
6.49 Local skewness and kurtosis maps of the bias cleaned CMB maps using the MEM		. . . . 194
6.50 Local skewness and kurtosis maps of the bias cleaned CMB maps using the GFM		. . . . 195

A.1	Coupling coefficients distribution determined from simulations . . . . .	208
A.2	Scatter plots of the coupling coefficients from simulations . . . . .	210
A.3	Scatter plots of the coupling coefficients from simulations with the K-Ka template . . .	211
A.4	Scatter plots of the scaling factors derived with the simple $\chi^2$ method . . . . .	212



# List of Tables

<b>List of Tables</b>	<b>xiii</b>
1.1 Dust emissivity at 100 $\mu\text{m}$ and near 30 GHz . . . . .	31
3.1 Coupling coefficients between the <i>WMAP</i> 3-year data and three foregrounds . . . . .	63
3.2 Fitted spectral index and normalisation factor . . . . .	66
3.3 Free-free spectral index and normalisation factor . . . . .	67
3.4 Coupling coefficients between the <i>WMAP</i> Q,V, and W-band data and three foregrounds	71
4.1 Coupling coefficients between the 5-year <i>WMAP</i> data and three foregrounds . . . . .	85
4.2 Coupling coefficients between the <i>WMAP</i> data and three foregrounds using K-Ka . . .	86
4.3 Spectral index and normalisation factor . . . . .	89
4.4 Fit parameters of the free-free coupling coefficients . . . . .	90
5.1 Mean, mode and errors of the fit parameters obtained from simulations with WIM . . .	125
5.2 Mean, mode and errors of the fit parameters obtained from simulations with WIM . . .	126
5.3 Mean, mode and errors of the fit parameters obtained from simulations without WIM . .	127
5.4 Mean, mode and errors of the fit parameters obtained from simulations without WIM . .	128
5.5 Statistical properties of the coefficients from noise simulations without WIM . . . . .	129
5.6 Statistical properties of the coefficients from noise simulations with WIM . . . . .	130
5.7 Fit parameters using the Dobler et al. (2009) model . . . . .	133
5.8 Fit parameters of the coefficients derived as Dobler et al. (2009) . . . . .	135
6.1 and ILC K-, Ka- and Q- bands weights . . . . .	163
6.2 and ILC V- and W- bands weights . . . . .	164
6.3 Variance of the merged CMB maps . . . . .	172
6.4 Regional values of the variance of CMB maps . . . . .	173

6.5	Variance and cross-variance of CMB maps derived from the regional analysis . . . . .	186
6.6	Variance of the CMB maps corrected for the bias . . . . .	192

# Zusammenfassung

Die Auflösung und Sensitivität der Messungen der kosmischen Hintergrundstrahlung (CMB) werden immer besser und die zurzeit laufende Planck-Mission sowie künftige Versuche stellen noch bessere Daten in Aussicht. Die Kontaminierung durch den galaktischen Vordergrund jedoch verzerrt die Messungen der Hintergrundstrahlung. Ihre Effekte müssen daher beseitigt werden, damit die gemessenen Daten für die kosmologische Forschung verwendet werden können, das ist nachwievor problematisch, weil die physikalischen Eigenschaften des interstellaren Mediums noch unzureichend bekannt sind.

In der beiliegenden Dissertation wird die Frage der Trennung der Komponenten für die CMB betrachtet und vor allem eine spezifische Anwendung einer unabhängigen Komponentenanalyse (ICA), namens *ICA*, vorgestellt. Dieser Algorithmus wurde benutzt bei der Vordergrundanalyse von 3-Jahres- und 5-Jahres- *WMAP* Datensätzen. Anschliessend wurden die Eigenschaften der wichtigsten Quellen von diffusen galaktischen Emissionen (z.B. Synchrotron, Bremsstrahlung und Staub) anhand dieses Algorithmus untersucht. In den *WMAP*-Daten wurde die Vordergrundkontaminierung quantifiziert durch Ankupplungskoeffizienten zwischen den Daten und verschiedenen Schablonen, die Beobachtungen von Emissionen in einem Frequenzbereich darstellen, wo wahrscheinlich nur eine physikalische Komponente dominierend ist. Mit diesen Koeffizienten wurde das Frequenzspektrum der galaktischen Komponenten, insbesondere der Bremsstrahlung, gerechnet. Unsere Ergebnisse weisen auf das Bestehen einer besonderen "Beule" im Spektrum hin. Dies wird interpretiert als Signatur einer rotierenden Staubemission im warmen ionisierten Medium. Solche Emission korreliert räumlich mit der  $H\alpha$ -Strahlung, durch die die Bremsstrahlung verfolgt wird.

Dieselben Ankupplungskoeffizienten wurden zur Bereinigung der *WMAP*- Daten angewendet, die dann mit *ICA* weiter analysiert wurden. Dieser interaktive Schritt in der Analyse erweist sich als ein mächtiges Werkzeug für die Bereinigung von CMB Messungen von Residuen, die mit den benutzten Schablonen nicht aufgespürt werden konnten. Es stellt eine einzigartige Methode dar, neue physikalischen Komponente zu entdecken. So fanden wir eine zusätzliche räumlich lokalisierte Emissionskomponente um das galaktische Zentrum, konsistent mit dem so genannten *WMAP*-Nebel.

Um die Eigenschaften der galaktischen Vordergrundemission zu berücksichtigen, schlagen wir vor, die *WMAP* Daten in verschiedenen Himmelsregionen aufzuteilen die Daten wurden darum sowohl mit *ICA* als auch mit der Methode der internen linearen Kombination (ILC) analysiert. Das Temperaturleistungsspektrum ist nicht besonders empfindlich auf feine Details der Korrekturen im Vordergrund, ausser bei den grösseren Skalen (bei niedrigen  $l$ ). Deshalb werden die beiden Methoden mit nicht-Gaussianischen Tests verglichen, zur Auswertung der Residuen. Während *ICA* nur in bestimmten Fällen und in wenigen Gebieten bessere Ergebnisse liefert, ist die Schätzung mit ILC im Allgemeinen erheblich besser, wenn die Anzahl der Gebiete erhöht wird. Andererseits spielt *ICA* eine Schlüsselrolle bei der Erstellung einer Aufteilung des Himmels, die seine Eigenschaften realistisch auseinander hält. Diese Arbeit zeigt, dass das für eine erfolgreiche Gebietsorientierte Analyse unentbehrlich ist.



# Abstract

Measurements of the Cosmic Microwave Background (CMB) emission with increasingly high resolution and sensitivity are now becoming available, and even higher quality data are expected from the ongoing Planck mission and future experiments. Dealing with the Galactic foreground contamination, however, is still problematic, due to our poor knowledge of the physics of the Interstellar Medium at microwave frequencies. This contamination biases the CMB observations and needs to be removed before using the data for cosmological studies.

In this thesis the problem of component separation for the CMB is considered and a highly focused study of a specific implementation of Independent Component Analysis (ICA), called *ICA-MAP*, is presented. This algorithm has been used to perform a foreground analysis of the *WMAP* three and five-year data and subsequently to investigate the properties of the main sources of diffuse Galactic emission (e.g. synchrotron, dust and free-free emission). The foreground contamination in the *WMAP* data is quantified in terms of coupling coefficients between the data and various templates, which are observations of the sky emission at frequencies where only one physical component is likely to dominate. The coefficients have been used to extract the frequency spectra of the Galactic components, with particular attention paid to the free-free frequency spectrum. Our results favour the existence of a spectral ‘bump’, interpreted as a signature of emission by spinning dust grains in the Warm Ionised Medium, which spatially correlates with the  $H\alpha$  radiation used to trace the free-free emission.

The same coupling coefficients have been used to clean the *WMAP* observations, which have then been further analysed using *ICA-MAP*. This iterative step in the analysis provides a powerful tool for cleaning the CMB data of any residuals not traced by the adopted templates. In practice, it is a unique way to potentially reveal new physical emission components. In this way, we detected a residual spatially concentrated emission component around the Galactic center, consistent with the so-called *WMAP* Haze.

In order to take into account the actual spatial properties of the Galactic foreground emission, we proposed an analysis of the *WMAP* data on patches of the sky, both using *ICA-MAP* and the Internal Linear Combination (ILC). Since the temperature power spectrum is reasonably insensitive to the fine details of the foreground corrections except on the largest scales (low  $l$ ), the two methods are compared by means of non-Gaussianity tests, used to trace the presence of possible residuals. While the performance of *ICA-MAP* improves only for particular cases with a small number of regions, the ILC CMB estimation generally ameliorates significantly if the number of patches is increased. Moreover, *ICA-MAP* plays a key role in establishing a partitioning that realistically traces the features of the sky, a requirement we have shown to be paramount for a successful regional analysis.



# Chapter 1

## The microwave emissions of the sky

### 1.1 Introduction

When observed at microwave frequencies, the sky is a superposition of emissions over a variety of angular scales generated by diverse physical mechanisms in our own Galaxy, contributions on small scales from equivalent processes in external galaxies, together with a background radiation of cosmological origin - the Cosmic Microwave Background (hereafter CMB). The former form a barrier to the study of the CMB, and it requires the application of sophisticated techniques to clean the observed microwave sky of the foreground emissions. A deeper comprehension of these foreground emissions, apart from being an interesting scientific goal on its own, helps in the design of cleaning techniques.

An overview to introduce the reader to the subject of CMB, foregrounds removal and analyses is presented in the following.

### 1.2 The Cosmic Microwave Background radiation

The discovery of the CMB radiation is well known as a case of serendipity in science. In 1964, Penzias and Wilson measured an excess of temperature calibrating a communication antenna at the wavelength of 7 cm. It was a signal of 3.5 K, isotropically distributed on the sky. Only later Dicke et al. (1965) proposed it was a signal of cosmological origin, namely a radiation left in the early Universe during a phase of very high temperature and density ( $\rho \sim 10^{25} \text{ g/cm}^3$  and  $T \sim 10^2 \text{ GeV} \sim 10^{15} \text{ K}$  at  $t \sim 10^{-8} \text{ s}$ ). It was the confirmation of what already predicted by Alpher & Herman (1948) following the assumption that primordial elements abundances were produced during an early hot and dense phase dominated by thermal radiation.

The existence of the CMB is considered one of the three observational pillars supporting the cosmological scenario of the Hot Big Bang, together with the light elements primordial abundances and the expansion of the Universe. The last one was proved by Hubble (1929), observing a *red-shift*<sup>1</sup> in the spectrum of external galaxies, which implied the Universe to be expanding with a velocity

---

<sup>1</sup>The red-shift is defined as  $z = \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}}$ , where  $\lambda_{em}$  is the wavelength emitted by the galaxy and  $\lambda_{obs}$  is the one observed.

$v = H_0 d$  (where the *Hubble constant*  $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1} \sim 71 \text{ km /s/ Mpc}$ , according to the most recent measurements, and  $d$  is the physical distance between two galaxies).

The history of the Universe is strictly connected to its expansion. At primordial high temperature radiation is tightly coupled with matter by means of Thomson scattering, so that it behaves as a nearly perfect fluid. At  $t \sim 3 \times 10^5$  years ( $z \sim 1000$ ) the temperature drops to  $\sim 3000$  K and protons are able to capture electrons to form neutral hydrogen and other light elements ( $^3\text{He}$ ,  $^4\text{He}$ ,  $^7\text{Li}$ ). This is the so called *recombination* process. At about the same time, the CMB photons propagate freely with no further scattering, the opacity of the primordial plasma being strongly reduced. The position at which the CMB photons seen today directly interacted with matter for the last time, is the *Last Scattering Surface* (LSS). These photons are what we look at when we observe the CMB today. In this sense the CMB radiation is a unique window to the past of the Universe. In fact observations of the CMB have been improved a lot in the past few decades, transforming the cosmology into a precision science.

The cosmological origin of the CMB is proved by its two main features: it is an isotropic radiation and has a black body spectrum.

The first aspect follows directly from the isotropy and homogeneity of the Universe as a whole. Modern cosmology is based on the assumption that the Universe is homogeneous and isotropic on sufficiently large scales of few Mpc. This is the so called ‘Cosmological principle’ which takes origin from the Copernican principle. Therefore, the remarkable uniformity of the CMB field shows that at the epoch of last scattering, the Universe was to a high degree of precision ( $10^{-4}$ ) isotropic and homogeneous. Today, we know that small deviations from homogeneity and isotropy in the CMB probably represent the seed that have led to the formation of large-scale structures via gravitational instability.

The second feature (the black body spectrum) is required by the fact that the Universe had time to reach the thermal equilibrium at the epoch of recombination. This corresponds to an epoch of the early Universe when the timescale for thermal processes was shorter than the expansion timescale,  $H^{-1}$  (in terms of red-shift it happens at  $z \sim 2 \times 10^6$ ). Both scattering processes (e.g. Compton and inverse Compton effects) and photon-generating processes, such as the thermal bremsstrahlung (or *free-free*) and radiative Compton effect, were involved. The scattering processes established the kinetic equilibrium characterised by the Bose-Einstein spectrum, with a positive chemical potential  $\mu$ . The radiative Compton and free-free processes set the thermal equilibrium, which requires the creation and destruction of photons as well as energy redistribution. Thermal equilibrium, in turn, generates a black body spectrum in the radiation field. Finally, once generated, the spectrum has been maintained by the adiabatic expansion of the Universe<sup>2</sup>.

The black body spectrum was definitively measured by *COBE-FIRAS*<sup>3</sup>, which derived the Planckian curve shown in Figure 1.1. The spectrum has a temperature  $T = 2.735 \pm 0.001$  K (Fixsen & Mather, 2002), within a frequency range of 60-600 GHz.

Moreover, as the cosmic interpretation of the microwave background has been demonstrated,

<sup>2</sup>Assuming the expansion to be adiabatic ( $PV^\gamma = \text{const}$ , with  $\gamma = 4/3$ ),  $T(z) = T_0/(z+1)$  and  $\lambda \propto (1+z)^{-1}$ , so that  $\lambda T$  is a constant. Therefore, because the Planck spectrum is characterized by  $\exp(hc/k\lambda T)$ , the shape of the curve does not change.

<sup>3</sup>*Far Infrared Absolute Spectrophotometer*

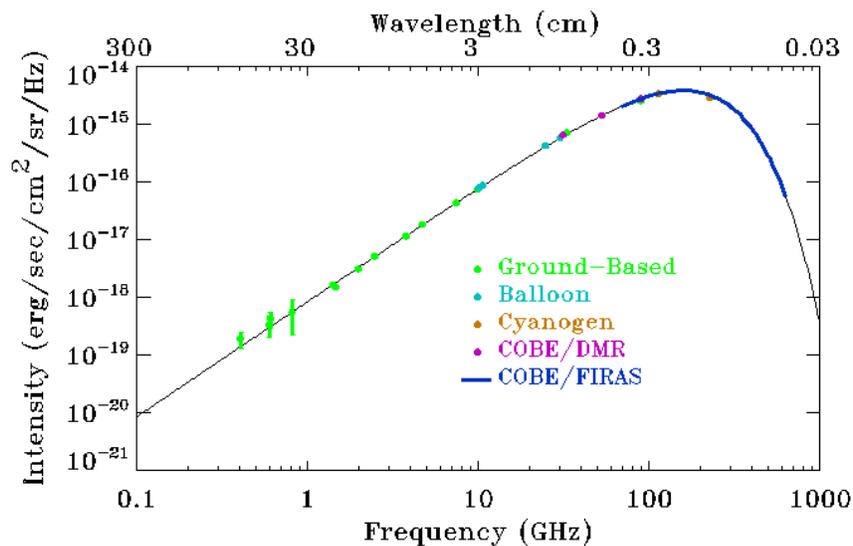


Figure 1.1: The plot shows measurements of the intensity of the cosmic microwave background as a function of observing frequency (or wavelength). The CMB follows the expected black body curve over more than 5 orders of magnitude in intensity. [http://arcade.gsfc.nasa.gov/cmb\\_spectrum.html](http://arcade.gsfc.nasa.gov/cmb_spectrum.html)

increasing interest arose in small departures from perfect isotropy and from a perfect black body spectrum.

### 1.2.1 Spectral distortions

If energy is added to the CMB radiation field after the epoch corresponding to a redshift of  $2 \times 10^6$ , there could have been time to reintroduce kinetic equilibrium, but not the thermal one. In this case, the spectrum would show a deviation from the expected Planckian curve, generally referred to as ‘ $\mu$  distortion’. The interest in detecting these signatures in the spectrum is connected to the fact that the center wavelength and the amplitude of the distortion depends on  $\mu$  and the baryon density.

Also Compton scattering can generate distortions in the spectrum (so called ‘ $y$ -distortions’): the photon number is here preserved, but the energy is boosted by collisions with hot electrons in the ionised material. This scattering process is the so called Sunyaev-Zeldovich effect (see Section 1.4.6 for more details). It affects the entire spectrum with a distortion which can be written in terms of the Compton parameter  $y$  defined as:

$$y = \int \frac{kT_e n_e \sigma_T}{m_e c^2} dl, \quad (1.1)$$

where  $T_e$ ,  $n_e$  and  $m_e$  are the temperature, density and mass of the electrons respectively, while  $\sigma_T$  is the Thompson scattering cross section. So far, no significant deviation from the expected spectrum has been measured. FIRAS measurements have set upper-limits to both the Comptonization parameter ( $-5 \times 10^{-6} < y < 3.5 \times 10^{-6}$ ) and the chemical potential ( $\mu < 7 \times 10^{-5}$ ).

### 1.2.2 CMB anisotropies and the inflationary scenario

Unlike spectral distortions, departures from isotropy have been well detected and become one of the main points of interest of the study of the CMB.

A remarkable anisotropy, at a level of  $\Delta T/T \sim 10^{-3}$  on angular scales of  $180^\circ$  was first detected in 1976. It is not strictly cosmological in origin, but rather induced by the Doppler effect due to the motion of our local frame with respect to the CMB. It is described by:

$$\frac{\Delta T}{T_0} = \frac{v}{c} \cos \theta + \frac{1}{2} \left( \frac{v}{c} \right)^2 \cos 2\theta + O\left( \frac{v}{c} \right)^3, \quad (1.2)$$

where  $\theta$  is the angle between the line of sight and the direction of the motion, and  $v$  is the observer velocity. Note that the second term of the equation above gives the dynamic quadrupole: it is rather small  $\sim 2\mu\text{K}$  (assuming  $v \simeq 370$  km/s) and although it is quite below the intrinsic CMB cosmic quadrupole ( $\sim 15\mu\text{K}$  (Nolta et al., 2009)), sometimes it is removed from the data.

However, apart from the locally induced anisotropy, the CMB field is characterized by intrinsic anisotropies of a smaller scale ( $\Delta T/T \sim 10^{-5}$ ). In fact, before they had been measured by *COBE-DMR*<sup>4</sup>, they have been already theoretically predicted in order to explain the existence of the structures (galaxies and galaxy clusters) that we observe in the Universe. Indeed, this requires to assume the presence of density fluctuations at the last scattering epoch, which in turn, would necessarily induce angular anisotropies in the CMB intensity.

The leading paradigm for what produces these seed fluctuations is *inflation* which seems to well complement the Hot Big Bang theory where it is not self-sufficient. Several problems indeed can not be explained by the Hot Big Bang theory. One of these is the so called ‘horizon problem’, meaning why distant regions of the Universe which could not be causally connected in the past, actually look qualitatively the same. The CMB in particular has the same temperature in every direction to one part in  $10^5$ . Furthermore, the Hot Big Bang theory does not explain why the geometry of the universe is close to flat, if any small deviation from flatness in the past would have grown as expansion proceeded, making the current near-flatness extremely unlikely (‘flatness problem’). Finally, the ‘monopole problem’ refers to the lack of evidence of magnetic monopoles.

According to inflation, in the early phase of its evolution ( $t \sim 10^{-34}$ ), the Universe would have a flat geometry and would be characterised by a Gaussian distribution of initial fluctuations: they would be described by a power spectrum of the form  $P(k) \sim k^{n_s}$ , with spectral index very close to one (Harrison, 1970; Zeldovich, 1972; Bond, 1996; Wright et al., 1996). In these conditions, the Universe would have suffered an exponential expansion, with nearly constant energy density. Consequently, any fluctuation during the expansion would be dilated, becoming density perturbations. Different physical processes are then responsible for the interaction of the radiation with these fluctuations: their efficiency depends on the angular scale and induce spatial variations in the temperature field of the CMB (as described below).

---

<sup>4</sup>Differential Microwave Radiometer

Temperature anisotropies can be described in terms of Spherical Harmonic series:

$$\frac{\Delta T}{T}(\theta, \varphi) = \sum_{lm} a_{lm} Y_{lm}(\theta, \varphi) \quad (1.3)$$

where  $\ell \sim 180^\circ/\theta$  and

$$a_{\ell m} = \int d\theta d\varphi Y_{\ell m}^*(\theta, \varphi) \frac{\Delta T}{T}(\theta, \varphi) \quad (1.4)$$

represent the multipole moments.

According to the simplest inflation models, the CMB temperature fluctuations should be Gaussian at early times (they predict a level of non-Gaussianity  $\leq 10^{-6}$ ) and they should be preserved so by linear evolution of the small fluctuations. Therefore, the  $a_{lm}$  coefficients should be uncorrelated for different  $\ell$  and  $m$ , and independent. In this case, the *angular power spectrum*  $\langle a_{\ell m}^* a_{\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{m,m'} C_\ell$  provides the complete statistical description of the temperature anisotropies. Besides these models, many alternative inflationary scenarios have been proposed (Mollerach, 1990; Moroi & Takahashi, 2001; Enqvist & Sloth, 2002; Lyth & Wands, 2002; Dvali et al., 2004), all of them leading to weak non-Gaussianity. The signatures of non-Gaussianity are often studied by simply parameterising the primordial perturbation,  $\Phi$ , as linear in some Gaussian random field  $\varphi_G$ , but with an additional non-linear term (e.g., a quadratic term), such that  $\Phi = \varphi_G + f_{NL}\varphi_G^2$ . The value of the non-linear coefficient,  $f_{NL}$  gives the level of non-Gaussianity expected.

Note that the power spectrum is related to the two-point correlation function by:

$$C(\theta) = \left\langle \frac{\Delta T}{T}(\hat{\mathbf{n}}_1) \cdot \frac{\Delta T}{T}(\hat{\mathbf{n}}_2) \right\rangle = \sum_{\ell} \frac{(2\ell + 1)}{4\pi} C_\ell P_\ell(\cos \theta), \quad (1.5)$$

where  $\hat{\mathbf{n}}_1$  and  $\hat{\mathbf{n}}_2$  are two unit vectors separated by an angle  $\theta$ , and  $P_\ell$  is the Legendre polynomial of order  $\ell$ . Therefore, if the CMB anisotropy field is Gaussian, all its statistical properties are characterised by the two-point correlation function: all even n-point correlation functions can be expressed in terms of the power spectrum, while the odd ones vanish identically. Thus a non-vanishing three point correlation function, or its Fourier transform, the *bispectrum*, would be an indicator of non-Gaussianity: it is the lowest order statistics able to distinguish non-Gaussian from Gaussian perturbations. Moreover, since primordial non-Gaussianity is model-dependent, detecting or just constraining primordial non-Gaussian signals would allow a better understanding of the physics of inflation.

In conclusion, having a map of the CMB temperature fluctuations and therefore, estimating its power spectrum, gives us a direct image of the LSS, the statistical properties of which are intrinsically connected with the physical mechanisms responsible for the creation of the primordial inhomogeneities. Furthermore, it allows us to gain information about the evolution of the Universe and the structures formation, which impact the photons propagation from the LSS. All this information can be derived from the shape and the features of the power spectrum at different angular scales and, hence, from the values of the cosmological parameters, which can be derived from it.

Anisotropies are generally distinguished in two classes: *primordial* (or *intrinsic*) anisotropies and

*secondary* anisotropies. In fact, the first originated at the time of recombination, while the second arose during the propagation of the photons from the LSS towards us.

## INTRINSIC ANISOTROPIES

Intrinsic anisotropies can be further categorised depending on the scale which they occur at. More specifically, the element of distinction is whether the scale corresponds to causally or not causally connected regions. As mentioned before, the recombination time occurs at  $z \sim 1000$ , which means  $t \sim 380000$  yr: because photons travel with the finite speed of light,  $c$ , at that time they have covered at most a distance equal to  $ct_{rec}$ . This is the so called *causal horizon*, which defines the regions of the Universe where the photons have interacted, modifying the anisotropies. Outside these regions, instead, the perturbations have not suffered modifications, preserving the primordial pattern.

In the following, we will give a brief account on these anisotropies.

### Sachs-Wolfe Effect

At angular scales larger than the horizon at the time of recombination ( $\theta > 2^\circ \sqrt{\Omega_0}$ ), CMB anisotropies are produced by the so called Sachs-Wolfe effect (Sachs & Wolfe, 1967) due to metric perturbations. In the Newtonian context, metric perturbations are related to perturbations in the gravitational potential  $\Phi$  that, in turn, are produced by density perturbations  $\delta\rho$ . This means that, photons climbing out of a potential well will suffer gravitational redshift and time dilation. Likewise, photons from under-dense regions will be blue-shifted. In the hypothesis of adiabatic perturbations and a matter-dominated Universe, the gravitational term is given by  $\frac{\Delta T}{T} = \delta\Phi$ , while the time-dilation term is  $\frac{\Delta T}{T} = -\frac{2}{3}\frac{\delta t}{t} = -\frac{2}{3}\delta\Phi$ . Therefore the final net effect is  $\frac{\Delta T}{T} = \frac{1}{3}\delta\Phi$ . The signature of this effect in the power spectrum is a plateau at large angular scales (low multipoles  $\ell$ ): the *Sachs-Wolfe Plateau*. This is particularly interesting, since this is also the scale range where no causal connection affect the initial perturbations. Therefore they reflect the initial perturbations and consequently, the initial power spectrum of matter density fluctuations, which are responsible for the temperature fluctuations. If the power spectrum is described by the so called Harrison-Zeldovich spectrum, meaning  $P(k) \propto k^{n_s}$  with  $n_s = 1$ , then  $C_\ell \propto 1/(\ell(\ell + 1))$ . Plotting the power spectrum as  $\ell(\ell + 1)C_\ell$  as a function of  $\ell$ , the value of  $n_s$  is directly connected with the plateau of the spectrum at low  $\ell$ .

### Acoustic Oscillations

At angular scales of  $0.1^\circ < \theta < 2^\circ$ , the physical processes responsible for the temperature anisotropies are causally connected and affect the primordial plasma before recombination. As previously explained, at that epoch of the evolution of the Universe, matter and radiation are tightly coupled and behave as a single fluid. Gravitational perturbations, and therefore density perturbations induce acoustic oscillations in the fluid, which are driven by gravity, baryons inertia and photons pressure forces: they can be described in terms of Harmonic oscillations. At the time of recombination, the modes of acoustic oscillations are ‘frozen’ at different phases of oscillations, and therefore at different wavelengths. Modes that are caught at either their maxima or minima of

their oscillation at recombination, correspond to peaks in the power or rms fluctuations. The first peak at scale of  $\sim 1^\circ$  (the so called *Doppler peak*) corresponds to the wave that has been frozen while it was suffering the maximum of compression at the time of last scattering. The secondary peaks at higher multipoles, instead, reflect the higher harmonics of the principal oscillations. Between the peaks, the valleys are partially filled by velocity maxima which are  $90^\circ$  out of phase with respect to the density plasma.

### Silk Damping

The acoustic oscillations decrease exponentially at angular scale  $\theta < 0.1^\circ$ , which corresponds to  $\ell \geq 1000$ . Such decrease, called *Silk damping*, is the dissipation of small-scale perturbations caused by photons random walking out of over-dense regions (Silk, 1968; Smoot, 1997), before recombination. The latter is not an instantaneous phenomenon, but it is long enough to allow the diffusion of photons which washes out the anisotropies on scales smaller than the mean free path of the photons. Therefore, the Silk damping sets the scale for the thickness of the last-scattering surface, and it is responsible for the CMB anisotropies being suppressed on small angular scales after the third peak.

## SECONDARY ANISOTROPY

Secondary anisotropies are generated during the photons propagation from the LSS towards the observer. The physical processes which induce them are also responsible of the structure formation and the evolution of the Universe. Therefore, detailed observations of these effects provides insights on these phenomena after recombination. Gravitational effects as well as re-ionization of the Universe are the main processes involved.

### Variations of the gravitational potential

A first effect responsible for the formation of secondary anisotropies on the CMB field, is time variation of the gravitational potential in the photons trajectory. These can be due to potential decay, gravitational waves connected to tensor perturbations, or non-linear effects associated with structure formation.

Depending on the time at which the variation of the gravitational potential occurs, we have the so called Early *Integrated Sachs-Wolfe* (ISW) effect, Late ISW effect and *Rees-Sciama* effect (Rees & Sciama, 1968). In all these cases, the CMB photons suffer a redshift which induce temperature perturbations:

$$\frac{\Delta T}{T} = 2 \int \dot{\Phi} dl, \quad (1.6)$$

where  $\Phi$  is the gravitational potential and dot represents the derivative with respect to time.

- **Integrated Sachs-Wolfe effect** The gravitational potential decays at a given angular scale, when the evolution of the Universe and its expansion are dominated by a component whose density is homogeneous on that scale. This happens just after the recombination, when the

photons density is not negligible and it causes the  $\Phi$  to decay, producing the early ISW effect: it peaks slightly to the left of the first acoustic peak. Density perturbations stop to grow once the dark energy dominates the expansion. As in the case of the matter-radiation transition, the gravitational potential must then decay. Such a decay, causes an effective heating of the photons in the gravitational well. Like the matter-radiation analogue, it is an intrinsically large effect since the net change due to the decay is 5 times the Sachs-Wolfe effect of  $\Phi/3$ . However, since the opposite effect occurs in the under-dense regions, the contributions are canceled out as photons traverse many crests and troughs of the potential perturbations during the matter-dark energy transition. The effect, called the late integrated Sachs-Wolfe effect, then appears only at large scales or low multipoles.

- **Rees-Sciama effect** When structures begin to form, the evolution of the Universe enters a non-linear regime, such that the approximation of constant  $\Phi$  is not valid any more. Given a single isolated structure, variations are induced along the line of sight due to the evolution of the density profile and, more significantly, to the bulk motion along the line of sight. This effect is relevant for the very small angular scales of the CMB power spectrum.

### Gravitational lensing

Intervening mass of the structures on large scales along the line of sight gravitationally lenses the CMB photons trajectories (Blanchard & Schneider, 1987) and hence distorts the temperature anisotropy field. The photons are deflected according to the angular gradient of the potential projected along the line of sight. Because surface brightness is conserved in lensing, the deflection simply remaps the observed fields as

$$T(\hat{\mathbf{n}}) \rightarrow T(\hat{\mathbf{n}} + \nabla\varphi) \quad (1.7)$$

where  $\varphi$  is the deflection angle. This is of order of a few arcminutes, but the lines of sight are coherently deflected across scales of a few degrees. Since the coherence scale of the acoustic features is larger than the deflection angle, the lensing effect can be calculated by Taylor expanding Equation 1.7. The result is a product of fields so that in harmonic space the modes are coupled to each other across a range  $\Delta\ell \sim 60$  set by the coherence of the deflection. In the temperature power spectrum, this mode coupling results in a smoothing of the acoustic peaks (Seljak, 1996). However, gravitational lensing affects also the polarized signal of the CMB, as explained in Section 1.2.3.

Due to its nature, the lensing effect can be probed not only by variations of the CMB power spectrum, but also by higher order statistics used for detecting non-Gaussianity (Bernardeau, 1998; de Oliveira-Costa et al., 1998) and cross-correlation with the ISW effect (Zaldarriaga & Seljak, 1999).

### Re-ionization

Other processes producing secondary anisotropy are related to local and global re-ionization of the

Universe. The first one produces what is generally called Sunyaev-Zeldovich effect: it is the result of the Compton scattering of the CMB photons by non-relativistic electron gas within clusters of galaxies. This results in an energy change of the CMB photons which can also be interpreted as a net flux coming from the cluster itself. This is why, the SZ effect is generally included among those physical processes which are foreground sources contaminating CMB measurements (see Section 1.4.6).

Also the Universe as a whole is observed to be ionised out of  $z \sim 6$  and it is thought to have undergone re-ionization sometime between  $7 \leq z \leq 30$  (see Barkana & Loeb (2001) for a review). Consequently, a minimum of a few percent of the CMB photons have re-scattered since recombination: the photons we see coming from a given direction may be originated from a completely different one on the LSS. This suppresses uniformly the peaks by a factor of  $e^{-\tau}$ , where  $\tau$  is the optical depth at the re-ionization epoch, for angular scales smaller than the horizon at that epoch. Indeed anisotropies are destroyed by the randomization of the direction in scattering. Since this can be confused with a change in the initial amplitude of fluctuations, it is important to resolve this ambiguity for the study of initial fluctuations and the growth rate of fluctuations. On the other side, re-ionization is interesting since it can provide information about the sources of ionizing radiation from the first astrophysical objects formed in the Universe.

### 1.2.3 CMB polarization

In this section, we outline the basic theoretical framework of the CMB polarization, emphasizing the main aspects that make it an appealing target for cosmological investigation. Excellent reviews on the physics of the CMB polarization are Cabella & Kamionkowski (2004), Kosowsky (1996) and Hu & White (1997).

The polarization of the CMB was predicted soon after the discovery of the CMB. The degree of polarization is expected to be of 5 – 10% on a characteristic scale of tens of arcminutes. This is directly related to the local quadrupole anisotropy incident on the scattering electrons at the LSS. The polarization fraction is a function of the duration of the last scattering. Only those photons that scattered in an optically thin region could have possessed a quadrupole anisotropy: multiple scattering would cause photon trajectories to mix and hence erases anisotropy.

The first detection happened thanks to the experiment DASI<sup>5</sup> (Kovac et al., 2002; Leitch et al., 2002). The DASI results were confirmed and extended (Leitch et al., 2005) almost simultaneously with the release of the CBI<sup>6</sup> (Readhead et al., 2004), CAPMAP<sup>7</sup> (Barkats et al., 2005) and Boomerang<sup>8</sup> (Montroy et al., 2006) results, all of them derived from observations at small angular scales ( $\ell > 100$ ). Moreover DASI, CBI, and Boomerang (Piacentini et al., 2006) teams also reported detections of the temperature-polarization cross correlation.

---

<sup>5</sup>*Degree Angular Scale Interferometer*

<sup>6</sup>*Cosmic Background Imager*

<sup>7</sup>*Cosmic Anisotropy Polarization MAPper*

<sup>8</sup>*Balloon Observations Of Millimetric Extragalactic Radiation and Geophysics*

Polarized radiation is generally described in terms of the Stokes parameters  $Q$ ,  $U$ , and  $V$  (Born & Wolf, 1980). For a quasi-monochromatic electromagnetic wave of frequency  $\omega_0$  propagating in the  $z$ -direction, the components of the wave's electric field vector at a given point in space can be written as

$$E_x = a_x(t) \cos[\omega_0 t - \theta_x(t)], \quad E_y = a_y(t) \cos[\omega_0 t - \theta_y(t)]. \quad (1.8)$$

Polarization implies a correlation between these two components. Note that the time dependence of the amplitudes is a consequence of the waves being quasi-monochromatic (i.e. a superposition of waves with frequencies distributed very closely around  $\omega_0$ ). See e.g. Rybicki & Lightman (1986) for further details. The Stokes parameters are defined as the time averages

$$\begin{aligned} I &\equiv \langle a_x^2 \rangle + \langle a_y^2 \rangle, \\ Q &\equiv \langle a_x^2 \rangle - \langle a_y^2 \rangle, \\ U &\equiv \langle 2a_x a_y \cos(\theta_x - \theta_y) \rangle, \\ V &\equiv \langle 2a_x a_y \sin(\theta_x - \theta_y) \rangle. \end{aligned} \quad (1.9)$$

The parameter  $I$  gives the radiation intensity which is positive definite. The other three parameters can take either sign and describe the polarization state. For unpolarized radiation,  $Q = U = V = 0$ . Moreover, the physical mechanism which generates polarization in the CMB cannot produce any circular polarization (Chandrasekhar, 1960), thus we expect  $V = 0$ .

$Q$  and  $U$  describe orthogonal modes of linear polarization and depend on the axes in relation to which the linear polarization is defined. From Equation (1.9), it is easy to show that when the coordinate system is rotated by an angle  $\alpha$ , the same radiation field is described by the parameters

$$\begin{aligned} Q' &= Q \cos(2\alpha) + U \sin(2\alpha), \\ U' &= -Q \sin(2\alpha) + U \cos(2\alpha). \end{aligned} \quad (1.10)$$

The polarization field can also be described in  $E$  and  $B$  modes (Kamionkowski et al., 1997b; Zaldarriaga & Seljak, 1999)

$$E_{\ell m} \pm iB_{\ell m} = - \int d\mathbf{n}_{\pm 2} Y_{\ell m}^*(\mathbf{n}) [Q(\mathbf{n}) \pm iU(\mathbf{n})], \quad (1.11)$$

in terms of the complete and orthogonal set of spin harmonic functions,  ${}_s Y_{\ell m}$ . This decomposition is analogous to the electromagnetic vector field case where the curl-free electric-field component (E-mode) and the divergence-free magnetic-field one (B-mode) can be distinguished. Represented as  $E(\mathbf{n}) = \sum E_{\ell m} Y_{\ell m}(\mathbf{n})$  and  $B(\mathbf{n}) = \sum B_{\ell m} Y_{\ell m}(\mathbf{n})$ , they describe scalar and pseudo-scalar fields on the sky and hence, are also distinguished by a parity transformation  $\mathbf{n} \rightarrow -\mathbf{n}$ . Parity invariance demands that the cross-correlation between the pseudo-scalar  $B$  and the scalars  $T$  or  $E$  vanishes:  $B$  has the opposite parity of  $T$  and  $E$ .

The physical process responsible for CMB polarization is the Thomson Scattering of the radiation

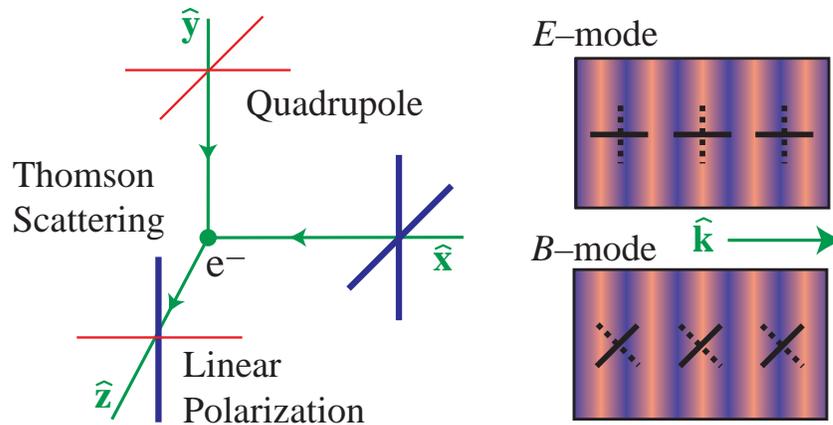


Figure 1.2: Left: Thomson scattering of quadrupole temperature anisotropies (depicted here in the  $\hat{x} - \hat{y}$  plane) generates linear polarization. Right: Polarization in the  $\hat{x} - \hat{y}$  plane along the outgoing  $\hat{z}$  axis. The component of the polarization that is parallel or perpendicular to the wavevector  $\vec{k}$  is called the  $E$ -mode and the one at  $45^\circ$  angles is called the  $B$ -mode. The plot is taken from Hu (2003).

with a quadrupole anisotropy (see e.g. Chandrasekhar (1960)). This can be generated by scalar, vector and tensor perturbations, each of them creating distinct patterns in the CMB polarization (see Figure 1.2). Scalar fluctuations can produce  $E$  modes, while vector and tensor perturbations (e.g. *gravitational waves*), are responsible for the  $B$  modes at large angular scales. Furthermore, the latter can also be created by gravitational lensing by the intervening mass distribution in the Universe: the  $E$  modes are lensed in  $B$  modes at small angular scales.

Figure 1.3, shows all the spectra of the CMB: it is easy to note some differences between the temperature power spectrum and the polarization ones. Specifically, as a direct result of the simple physics of acoustic oscillations, the amplitude of the  $EE$  spectrum is lower than the temperature spectrum by a factor of ten and the oscillatory structure is out of phase: the quadrupole anisotropy at the LSS is closely related to the velocity of the coupled photon-barion fluid, so that maximum of compression and rarefaction (and minimum of velocity) of the fluid corresponds to peak in the temperature spectrum and troughs in the  $EE$  modes. For the same reason the cross-correlation power spectrum between temperature and  $EE$  modes shows peaks corresponding to the interleaved sets of maxima and minima in the two separate components. The final feature of the polarization spectra is the comparative smallness of the  $BB$  signal, which makes it difficult to measure them: even for optimistic values of the tensor-to-scalar ratio  $r$ , the rms signal is only a fraction of  $\mu\text{K}$ , less than 1% of the level of temperature anisotropies at degree scales. Until now, no experiment has been able to actually detect the  $B$  modes, while upper limits have been proposed. In the short term, a tensor-to-scalar ratio of  $r \sim 0.05$  is well within reach of the Planck mission and planned ground-based experiments (Tucci et al., 2005). The next generation of missions will be able to significantly improve the sensitivity to polarized emission by means of arrays of  $10^3 - 10^4$  detectors, and will probably allow us to investigate theories of inflation, down to energy of  $5 \times 10^{15}$  GeV, meaning  $r = 0.008^9$ . A detection of these

<sup>9</sup>The relation between the energy scale inflation  $E_i$  and the tensor-to-scalar ratio at CMB scales is  $E_i = 1.06 \times 10^{16} \text{GeV} \left(\frac{r}{0.01}\right)^{1/4}$

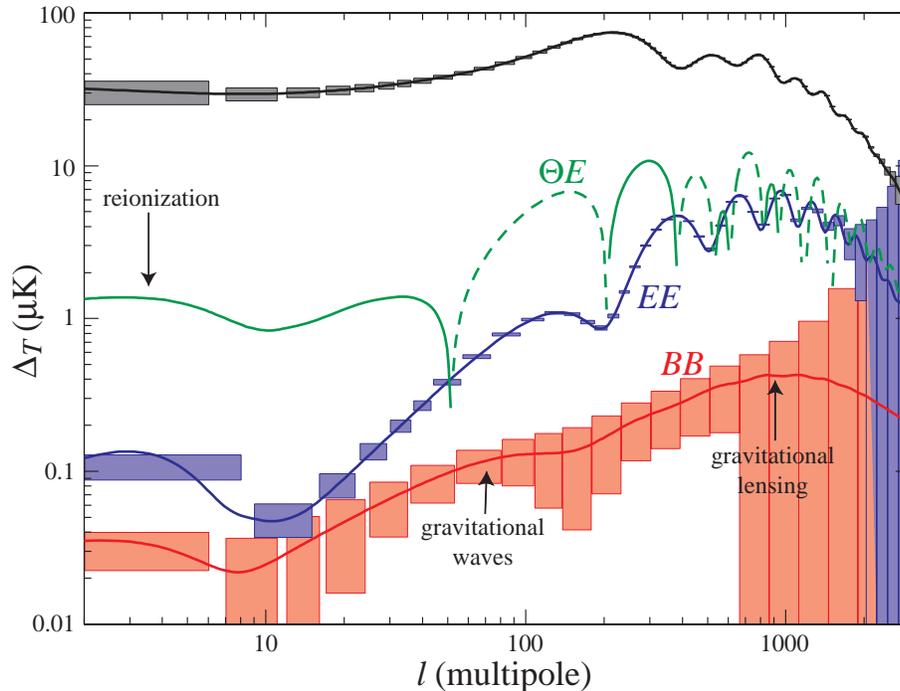


Figure 1.3: Temperature and polarization angular spectra for  $\Omega_{tot} = 1$ ,  $\Omega_{\Lambda} = 2/3$ ,  $\Omega_b h^2 = 0.02$ ,  $\Omega_m h^2 = 0.16$ ,  $n=1$ ,  $z_{ri} = 7$ ,  $E_i = 2.2 \times 10^{16}$  GeV (energy scale of inflation, which corresponds to a tensor-to-scalar ratio  $r \sim 0.01$ ). Dashed lines represent negative cross correlation and boxes represent the statistical errors of the Planck satellite. The plot is taken from Hu (2003).

modes would be of paramount importance since it would confirm inflationary models which invoke gravitational waves (Seljak & Zaldarriaga, 1997; Kamionkowski et al., 1997a) and, at small angular scales, would give information about the matter distribution. On the other side, the study of the CMB polarization in general would give us information about the re-ionization period (Zaldarriaga, 1997).

#### 1.2.4 Power spectrum, cosmological parameters and cosmic variance

An accurate evaluation of the anisotropy pattern of the CMB can determine the value of the parameters which we use to describe the cosmological models of the Universe. Figure 1.4 shows the values of the parameters directly measured or otherwise derived by *WMAP*<sup>10</sup> using the data from the 7-year observations.

Particularly significant is the sensitivity of the  $C_\ell$ 's on the values of some fundamental quantities, such as the total density  $\Omega_0$ , the baryons density  $\Omega_b$ , the Hubble constant, the Cosmological Constant  $\Lambda$ , the spectral index  $n_s$ , the optical depth  $\tau$  and the scalar-to-tensor ratio  $r$ . Figure 1.5 shows as example the variation of the acoustic peaks depending on the values of the total, matter, baryons and dark energy densities.

CMB measurements alone, however are not sufficient to determine the cosmological parameters,

(Baumann et al., 2009).

<sup>10</sup>Wilkinson Microwave Anisotropy Probe

WMAP Cosmological Parameters			
Model: lcdm+sz+lens			
Data: wmap7			
$10^2\Omega_b h^2$	$2.258_{-0.056}^{+0.057}$	$1 - n_s$	$0.037 \pm 0.014$
$1 - n_s$	$0.0079 < 1 - n_s < 0.0642$ (95% CL)	$A_{\text{BAO}}(z = 0.35)$	$0.463_{-0.020}^{+0.021}$
$C_{220}$	$5763_{-40}^{+38}$	$d_A(z_{\text{eq}})$	$14281_{-161}^{+158}$ Mpc
$d_A(z_*)$	$14116_{-163}^{+160}$ Mpc	$\Delta_{\mathcal{R}}^2$	$(2.43 \pm 0.11) \times 10^{-9}$
$h$	$0.710 \pm 0.025$	$H_0$	$71.0 \pm 2.5$ km/s/Mpc
$k_{\text{eq}}$	$0.00974_{-0.00040}^{+0.00041}$	$\ell_{\text{eq}}$	$137.5 \pm 4.3$
$\ell_*$	$302.44 \pm 0.80$	$n_s$	$0.963 \pm 0.014$
$\Omega_b$	$0.0449 \pm 0.0028$	$\Omega_b h^2$	$0.02258_{-0.00056}^{+0.00057}$
$\Omega_c$	$0.222 \pm 0.026$	$\Omega_c h^2$	$0.1109 \pm 0.0056$
$\Omega_\Lambda$	$0.734 \pm 0.029$	$\Omega_m$	$0.266 \pm 0.029$
$\Omega_m h^2$	$0.1334_{-0.0055}^{+0.0056}$	$r_{\text{hor}}(z_{\text{dec}})$	$285.5 \pm 3.0$ Mpc
$r_s(z_d)$	$153.2 \pm 1.7$ Mpc	$r_s(z_d)/D_v(z = 0.2)$	$0.1922_{-0.0073}^{+0.0072}$
$r_s(z_d)/D_v(z = 0.35)$	$0.1153_{-0.0039}^{+0.0038}$	$r_s(z_*)$	$146.6_{-1.6}^{+1.5}$ Mpc
$R$	$1.719 \pm 0.019$	$\sigma_8$	$0.801 \pm 0.030$
$A_{\text{SZ}}$	$0.97_{-0.97}^{+0.68}$	$t_0$	$13.75 \pm 0.13$ Gyr
$\tau$	$0.088 \pm 0.015$	$\theta_*$	$0.010388 \pm 0.000027$
$\theta_*$	$0.5952 \pm 0.0016$ °	$t_*$	$379164_{-5243}^{+5187}$ yr
$z_{\text{dec}}$	$1088.2 \pm 1.2$	$z_d$	$1020.3 \pm 1.4$
$z_{\text{eq}}$	$3196_{-133}^{+134}$	$z_{\text{reion}}$	$10.5 \pm 1.2$
$z_*$	$1090.79_{-0.92}^{+0.94}$		

Figure 1.4: Values of the cosmological parameters derived by *WMAP* with the seven-year observations. [http://lambda.gsfc.nasa.gov/product/map/dr4/params/lcdm\\_sz\\_lens\\_wmap7.cfm](http://lambda.gsfc.nasa.gov/product/map/dr4/params/lcdm_sz_lens_wmap7.cfm)

due to problems of degeneracies: the same power spectrum features may arise from different combinations of their values. Complementary data set are therefore necessary to break these degeneracies (for example Supernova Type Ia measurements).

Furthermore, although there are different limitations on the accuracy of the CMB measurements, the *Cosmic variance* sets the ultimate limit. This is due to the fact that our Universe is a single realization of a stochastic process and seen from one vantage point, i.e. the Earth. Therefore, it does not necessary follow the average over the ensemble of possible realizations. This is equivalent to say that the  $a_{\ell m}$  coefficients are independent identically distributed Gaussian random variables (for a given  $\ell$ ) and therefore the  $C_\ell$  are a  $\chi^2$  distribution with  $2\ell + 1$  degrees of freedom. The variance in  $C_\ell$ 's is then:

$$\frac{\delta C_\ell}{C_\ell} = \sqrt{\frac{2}{2\ell + 1}}, \quad (1.12)$$

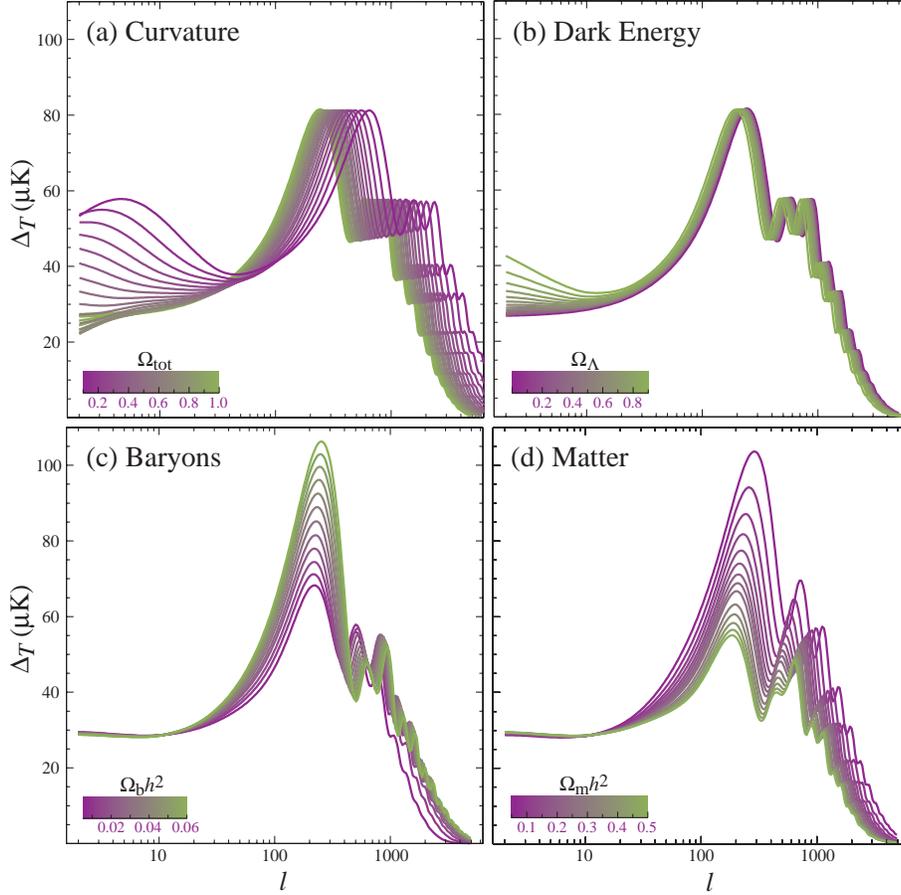


Figure 1.5: Sensitivity of the acoustic temperature spectrum to four fundamental cosmological parameters (a) the curvature as quantified by  $\Omega_{tot}$  (b) the dark energy as quantified by the cosmological constant  $\Omega_\Lambda$  (c) the physical baryon density  $\Omega_b h^2$  (d) the physical matter density  $\Omega_m h^2$ , all varied around a fiducial model of  $\Omega_{tot} = 1$ ,  $\Omega_\Lambda = 0.65$ ,  $\Omega_b h^2 = 0.02$ ,  $\Omega_m h^2 = 0.147$ ,  $n_s = 1$ ,  $z_{ri} = 0$ ,  $E_i = 0$ .

which becomes quite important at low multipoles, since a small number of modes is available.

### 1.3 Non-Gaussianity

As stated before, according to the simplest inflation models, the CMB temperature fluctuations should be nearly Gaussian. The *WMAP* science team (Komatsu et al., 2003; Spergel et al., 2007; Bennett et al., 2010) found the *WMAP* data consistent with the hypothesis of Gaussianity, using Minkowski functionals (which describe the properties of the hot and cold regions at different threshold levels), and the bispectrum (the Fourier analogue of the three-point correlation function). They concluded that there is no compelling evidence for deviations from the  $\Lambda$ -dominated cold dark matter ( $\Lambda$ CDM) standard model, which is generally an acceptable statistical fit to *WMAP* and other cosmological data.

However, other works have tested the Gaussianity of the *WMAP* data in many different ways, some of them finding significant deviations from it. Examples are the analyses based on phase correlations (Chiang et al., 2003), the genus (Park, 2004), isotropic wavelets (Cruz et al., 2006; Vielva et al., 2004),

1-pdf (Monteserín et al., 2008), isotropy analyses based on local  $n$ -point correlations (Eriksen et al., 2004), local curvature (Hansen et al., 2004).

Among the ‘anomalies’ that have been detected, the most significant are the North-South asymmetry (Eriksen et al., 2004; Hansen et al., 2006), the alignment of the low multipoles (de Oliveira-Costa et al., 2004; Land & Magueijo, 2005), the cold spot (Cruz et al., 2006, 2007; Vielva et al., 2004), the non-Gaussian features detected with directional wavelets (McEwen et al., 2005, 2006, 2008) and the low variance (Monteserín et al., 2008).

### 1.3.1 North-south asymmetry

Several accurate analyses have revealed a statistically significant north-south ecliptic asymmetry in the large-angle correlations strength of the *WMAP* data (see for example Hansen et al. (2004) and Eriksen et al. (2004)). This phenomenon could be a consequence of a systematic error in the data processing or in the instrument characterization, residual foregrounds, and large-angle correlations induced by the incomplete sky CMB data.

Eriksen et al. (2004) found the results to be stable with respect to the choice of the Galactic cut and the frequency band, which suggests it is an effect of foreground residuals. Moreover, since a similar asymmetry is found in the *COBE*-DMR map (the axis of maximum asymmetry is close to the one found in the *WMAP* data), the asymmetry is possibly not connected to *WMAP* systematics.

Bernui (2008) confirmed such large-angle anisotropy in three- and five-year *WMAP* CMB maps containing only the multipole components  $4 \leq \ell \leq 10$ . They also suggested that this hemispherical asymmetry is unlikely to be due to systematics.

### 1.3.2 Amplitude of the quadrupole

The *WMAP* data have also shown an important suppression of power at large scales. More precisely, the power of the quadrupole is substantially reduced with respect to the value of the best-fit  $\Lambda$ CDM model. This peculiarity emerged since the first data of *COBE*-DMR appeared (Smoot et al., 1992; Bennett et al., 1996). Later, this behavior was confirmed by the *WMAP* data. In fact, the *WMAP* five year data give a quadrupole power of  $213 \mu K^2$ , while the expected value in the  $\Lambda$ -dominated cold dark matter model is about  $1207 \mu K^2$  (Nolta et al., 2009).

If this discrepancy turns out to have a cosmological origin, then it could have consequences for our understanding of the Universe and, in particular, for the standard inflationary picture.

de Oliveira-Costa et al. (2004) gave a probability of such a low quadrupole as 1 in 20, although it has been estimated with an a posteriori statistic: it applies to the one data point chosen for being furthest from the expectation. The *WMAP* science team, instead, considered the quadrupole still consistent with the best-fit  $\Lambda$ CDM model. Other analyses (see for instance Land & Magueijo (2005)) confirmed that the amplitude of the quadrupole itself is not statistically anomalous. Instead, Copi et al. (2006) examine the significance of the power deficit at large scales using the angular correlation function outside the Galactic plane region and find that it is anomalous at the 99.97% level.

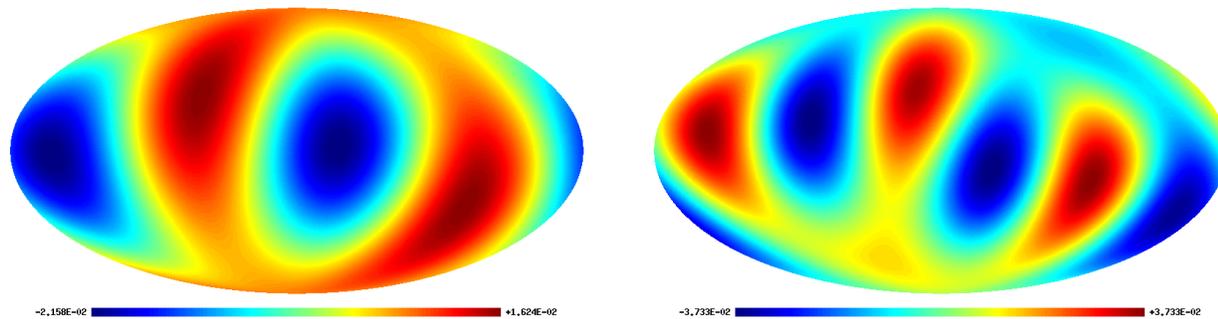


Figure 1.6: Maps of the quadrupole (left) and octopole (right) obtained from the *WMAP* five-year ILC map.

### 1.3.3 Alignment of the low multipoles

The lowest multipoles, especially the quadrupole ( $\ell = 2$ ) and the octopole ( $\ell = 3$ ), of the *WMAP* data have been found to be anomalously planar and aligned (Bielewicz et al., 2005; Copi et al., 2004, 2006; de Oliveira-Costa et al., 2004). Figure 1.6 shows the quadrupole and octopole of the ILC derived from the *WMAP* five-year data. Both multipoles present maxima and minima following a planar shape, whose perpendicular axis points towards similar direction called *axis of evil*. The axes of the two multipoles are separated by  $\sim 10^\circ$ . The probability that the two directions are separated by that angle or less by chance is  $\sim 1.5\%$ . Further alignments have also been claimed for higher multipoles  $\ell \leq 5$  (Land & Magueijo, 2005) and  $\ell = 6$  and  $\ell = 7$  (Freeman et al., 2006). The northern end of the alignment points towards  $(\theta, \varphi) = (30^\circ, 260^\circ)$ , a direction close to the CMB dipole one whose northern end is at  $(\theta, \varphi) = (42^\circ, 264^\circ)$ .

The estimation of the low multipole components is made difficult if a mask is applied to the data. Varying the mask produces significant changes in their amplitude estimates, especially for the quadrupole, implying consequently uncertainties in the determination of their axes. In fact, detailed analyses of this effect tend to weaken the significance of the detection (de Oliveira-Costa & Tegmark, 2006; Land & Magueijo, 2007).

## 1.4 Galactic and extra-Galactic components

One of the biggest obstacles to observe the anisotropies in the CMB are *foregrounds*, other sources of radiation which also emit at microwave frequencies. They are diffuse emissions from our Galaxy, and also the radiation produced by point sources such as radio galaxies, and by clusters of galaxies which scatter CMB photons through the Sunyaev-Zeldovich effect. Among them, at large angular scales, the strongest contamination comes from the Galactic foregrounds: they include certainly synchrotron, free-free emission (or *thermal bremsstrahlung*) along with thermal (vibrational) emission from dust. The first two components dominate at frequencies below  $\sim 100$  GHz, while the third one is the strongest emission at higher frequencies. Recently another component has been accepted to be an independent source of contamination produced by our Galaxy. This is the so called ‘anomalous emission’ from dust: its nature is still not well identified, although we know that it is a dust correlated emission at microwave

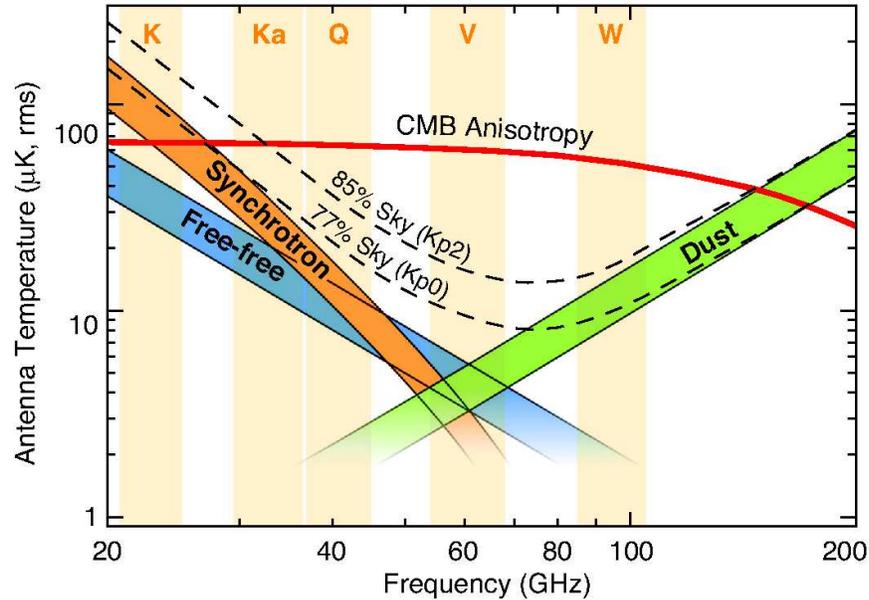


Figure 1.7: The figure shows the rms anisotropy as a function of frequency from the CMB (red line) and three known sources of foreground emission: synchrotron, free-free, and thermal dust emission. The composite galactic emission for two sky cuts, retaining 77% and 85% of the sky respectively, are shown as dashed lines. The five *WMAP* radiometer bands are indicated in the background.

wavelengths, but inconsistent with the thermal mechanism. A subject of ongoing debate is the so called ‘*WMAP* haze’ which Dobler & Finkbeiner (2008a) attribute to *hard* synchrotron emission distributed around the Galactic center, whose origin is again uncertain.

A simple way of avoiding foreground contamination is to select the regions where to take the observations from. This is the main idea behind ground based or balloon CMB experiments, whose measurements are generally concentrated in specific patches of the sky, where the foregrounds are known to be weak. Besides that, masks of the Galactic plane are generally applied to the data in order to exclude those regions where the contamination’s reach their strongest contribution. Both these approaches, however, have the drawback of missing information enclosed in the cut sky.

To avoid these issues it is generally more convenient to use multi-frequency measurements as well as to select the frequency of observation. Specifically in the case of Galactic components, as shown in Figure 1.7, each emission has a different spectral behavior, which can be approximately described by a power law. Their contribution reaches a minimum around 60 GHz, which defines the optimal window for CMB measurements: the cosmological emission, indeed, dominates over the foregrounds. Multi-frequency observations, chosen where the cosmological emission dominates over its contaminants, are helpful in order to discriminate the different contributions in the sky, based on the specific spectral and spatial properties of each of them.

In the following sections, we will give a brief description of the physical processes responsible for such contamination, highlighting their main properties. A better understanding of those is the key for cleaning the CMB data, to be consequently used for cosmological studies. Of course, the better we

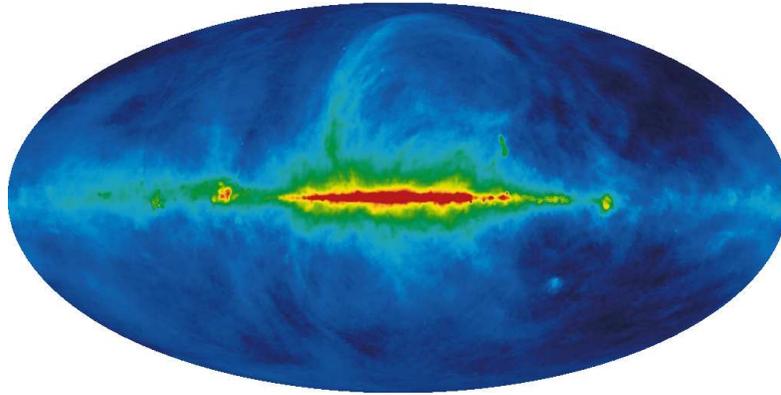


Figure 1.8: We show the Haslam map at 408 MHz, derived from 4 separate surveys. Scan striping and contributions from strong point sources have been mitigated (compared to the original map) through the use of filtering in the Fourier domain. It has an angular resolution of  $0.85^\circ$  and a sensitivity of 2 K.

know the Galactic components, the better we can subtract them from CMB observations.

### 1.4.1 Synchrotron emission

The synchrotron radiation arises from the interaction of relativistic electrons with the Galactic magnetic field. They are the electron component of the cosmic rays which pervade our Galaxy and are continually produced in supernovae explosions that occur in the plane of the Galaxy and then diffuse into the halo.

Figure 1.8 shows the all-sky 408 MHz map (Haslam et al., 1982), which is generally used to trace the synchrotron emission from our Galaxy, although it contains partially emission from free-free: this is particularly true along the Galactic plane and in the Gould Belt system which extends up to  $30^\circ$  above the plane in the Galactic center region and below the anticenter. Reich & Reich (1988) estimate the relative fraction at  $b = 0^\circ$  to be of order 10–20%. However, the contribution at intermediate and high latitudes is effectively negligible and hence the 408 MHz sky map is generally assumed as the default template for the synchrotron foreground component. The main feature of this map is the strong emission at high latitudes in the Northern Hemisphere, the so called North Polar Spur, which is most likely the synchrotron emission produced in a supernova remnant (SNR). Other similar features can be seen extending from the plane: they are all probably the low surface brightness remains of old SNRs.

Indeed, supernovae play an important role in supplying the cosmic ray electrons which escape from the SNRs and diffuse out into the interstellar medium. Those with the highest energy will eventually escape from the Galaxy altogether. When the energy will not be large enough, instead, the electrons will remain in the Galaxy. The synchrotron emission which we observe as a contaminant of the CMB is actually produced by those electrons: they are held in the Galaxy and, spiraling in the magnetic field, produce the diffuse radio halo which we observe. The latter, as a consequence of the mechanism which produces it, represents the best evidence of the structure of the magnetic field of the Galaxy.

Let us consider an electron of mass  $m$  and charge  $e$ , with a relativistic energies  $E = \gamma E_0$  where  $E_0 = mc^2$  (0.51 MeV) which moves in a magnetic field  $\mathbf{B}$ . The equations which describes the motion

of such a particle are:

$$\frac{d}{dt}(\gamma m \mathbf{v}) = \frac{e}{c} \mathbf{v} \times \mathbf{B} \quad (1.13)$$

$$\frac{d}{dt}(\gamma m c^2) = e \mathbf{v} \cdot \mathbf{E} = 0 \quad (1.14)$$

where  $\gamma$  is the Lorentz factor and  $\mathbf{v}$  is the velocity of the electron. By solving these equations, we obtain that the particle follows an helical motion, as the result of the combination of a circular uniform motion around the magnetic field lines, and a uniform motion along the lines. Specifically, they spiral with a Larmor frequency,  $\nu_L$ , given by

$$\nu_L = \frac{eB}{2\pi mc} = 2.8 \left( \frac{B}{1 \text{ Gauss}} \right) \text{ MHz}. \quad (1.15)$$

We can then accordingly define its gyrofrequency as

$$\nu_g = \frac{eB}{2\pi \gamma mc} = \frac{\nu_L}{\gamma}. \quad (1.16)$$

The emission from such a relativistic electron is in an instantaneous cone angle along its trajectory of  $\sim \gamma^{-1}$  steradians. Therefore, the spectrum of a single electron consists of the harmonics of the apparent pulse duration ( $\sim \nu_a^{-1}$ , with  $\nu_a \sim \gamma^2 \nu_L$ ) as seen by an external observer, and has a maximum emission at the critical frequency of  $\nu_c \sim 3/2 \nu_a$ . In a magnetic field of intensity  $B$ , we have

$$\nu_c = \frac{3eB\gamma^2}{4\pi mc} = 4.2 \left( \frac{B}{\text{Gauss}} \right) \gamma^2 \text{ MHz} = 1.6 \times 10^7 \left( \frac{B}{\text{Gauss}} \right) \left( \frac{E}{\text{GeV}} \right)^2 \text{ MHz}. \quad (1.17)$$

Given a distribution of relativistic electrons, they produce an energy spectrum which will depend on the energy distribution  $N(E)$  of the particles. The intensity of the radiation will be expressed by

$$I(\nu) = \int \int P(\nu, \mathbf{B}, E) N(E, \mathbf{l}) dE d\mathbf{l}, \quad (1.18)$$

where  $P(\nu, \mathbf{B}, E)$  is the power emitted at a frequency  $\nu$  by the single electron of energy  $E$  in the magnetic field of amplitude  $B$ .

Assuming the electron energy distribution to be well described by a power law in the form

$$N(E) dE = N_0 E^{-p} dE, \quad (1.19)$$

the emission intensity will be given by

$$I(\nu) \propto N_0 B^{(p+1)/2} \nu^{-(p-1)/2}. \quad (1.20)$$

That is, the emission spectrum itself is described by a power law, whose spectral index  $\beta_{\text{sync}}$  is related

to the one of the energy distribution as following:

$$\alpha_{sync} = -(p - 1)/2. \quad (1.21)$$

Taking into account the conversion between the intensity and the temperature ( $T(\nu) = I(\nu)c^2/2k\nu^2$ ), we have

$$T(\nu) \sim \nu^{-(p+3)/2} = \nu^{-\beta_{sync}} \quad (1.22)$$

The possibility of constraining the value of the spectral index is currently one of the greatest challenges of the studies of this foreground emission. Indeed, the value of  $\beta_{sync}$  suffers of the variations in the intensity of the magnetic field, as well as of the variations of the electrons density and energy. The last one specifically, is affected by cosmic ray propagation, energy loss, and degree of confinement.

More than 90% of the observed synchrotron emission arises from a diffuse component with a direction-dependent spectral index that generally lies in the range  $-0.5 > \alpha > -1.1$  ( $2.5 < \beta_{sync} < 3.1$  and  $2.0 < p < 3.2$ ). The cosmic ray electrons in the Galaxy substantially outlive their original supernova remnants and slowly lose energy while traveling large distances across the Galaxy. The low-frequency cut-off of synchrotron emission arises both from the single-electron synchrotron spectrum ( $F(x) = x \int_x^\infty K_{5/3}(\xi)d\xi$ , where  $K_{5/3}$  is the modified Bessel function) low-frequency cut-off, and from self-absorption and free-free absorption, both of which become increasingly efficient at lower frequencies. Besides, we expect the electrons spectrum to be steep at high energies ( $\sim$  GeV), since the most energetic electrons escape from the Galaxy: indeed recent studies have demonstrated that  $p$  varies from 2.7 to 3.3.

The complexity in measuring precisely the value of the spectral index of the synchrotron emission, is due to the limitations in the available observations: they are generally affected by systematic errors connected to the specific properties of the instruments and to the scanning strategy itself. The first full-sky map at the frequencies where the synchrotron emission is dominant, is the map at 408 MHz, already shown in Figure 1.8. It is the result of the combination of observations taken at Jodrell Bank and Bonn in the north and Parkes in the southern Hemisphere. It has an angular resolution of  $0.85^\circ$  and a sensitivity of 2 K.

There is also a map at 1420 MHz, which initially covered only the north Hemisphere and afterwards has been extended to the south hemisphere (Reich & Reich, 1986; Reich et al., 2001). Finally there are incomplete maps at 2326 MHz (with  $\delta < 30^\circ$ ) (Jonas et al., 1998) and at 38 MHz and 820 MHz (Lawson et al., 1987).

Studies indicate a range of spectral indices  $\beta_s$  from 2.3 to 3.0 between 408 MHz and 1420 MHz (Reich & Reich, 1988) and the presence of spurious baseline effects in the survey 408 MHz (Davies et al., 1996) that affect the spectral index determination for weaker features. The combination of these effects, plus the large frequency gap between 408 MHz and microwave wavelengths, does suggest that the sky map scaled by a single effective index may not be representative of the synchrotron emission at *WMAP* frequencies. Nevertheless, its use for foreground studies is well established in the CMB literature. However, Hinshaw et al. (2007) suggested that a better tracer for the synchrotron emission

at *WMAP* frequencies is the difference between the K and Ka *WMAP* data. This at least in part compensates for possible errors introduced by using the Haslam template, presumably accounting for the change in morphology of the emission at microwave frequencies. Moreover, using this template with a single fitted scaling per frequency is likely to be sufficiently accurate even given modest departures from a single spectral index. It is also likely to be a good choice because the intrinsic systematic measurement errors are smaller than for the Haslam map.

Different studies have been carried out taking advantage of all these observations. Lawson et al. (1987) have evaluated the spectral index to have a value between 2.55 and 2.8 for frequencies within 100 and 800 MHz respectively. They also found a typical variation of  $\Delta\beta_{sync} = \pm 0.15$  considering couples of frequencies with  $\delta\nu = 2 - 3$  (Reich & Reich, 1986). Developing a destripping technique, applied to the maps at 408 and 1420 MHz, Davies et al. (1996) obtained a value for the spectral index between 2.8 and 3.2. Improving this technique applied to the same maps together with the one at 2326 MHz, Platania et al. (2003) have defined a distribution of values of the spectral index with a mean value of  $\beta_{sync} = 2.66 \pm 0.12$ . The *WMAP* science team has proposed similar results. They associated the variations of the spectrum with the propagation of the cosmic ray electrons in the Galaxy. Specifically, Bennett et al. (2003) assumed the radiation to be produced by the electrons which outlive the SNRs and they distinguished between the electrons which are able to escape from the halo and those who are trapped into the disk. In the first case they evaluated  $\beta \sim 2.5$  which increases to  $\sim 3.0$  in the second one. Platania et al. (1998), studying the spectrum of the local cosmic rays, found a variation of the index equal to 0.5 and therefore a value of  $\beta_{sync} = 3.1$  for frequencies between 1 and 10 GHz.

At frequencies higher than 10 GHz, multi-frequency correlation analysis have been carried out by many authors. For example, comparing maps at 10 and 15 GHz from the experiment of Tenerife, with the Haslam map and the map at 1420 MHz, de Oliveira-Costa et al. (1997) have measured a value of  $\beta_{sync} \sim 3.0$ . Moreover, Davies et al. (2006), using the first year of observations of *WMAP*, have constraint the spectral index to be 2.85. The same result has been confirmed by Banday et al. (2003): they found a value of 2.5 for the DMR frequencies. In the next chapters, we will describe the analysis performed with the code `synchrofit` using the three- and five-year observations of *WMAP*, which also allowed us to estimate the synchrotron spectrum.

Synchrotron emission can be highly polarized: since the acceleration suffered by the electrons is perpendicular to the magnetic field and therefore, the emission is linearly polarized in the same plane. The degree of linear polarization is  $\sim 75\%$ , which depends on the spectral index  $p$  of the electrons:

$$\Pi = \frac{p + 1}{p + 7/3} \quad (1.23)$$

Although theoretically quite strong, this level of polarization is almost never observed due to non-uniform magnetic field directions along a line of sight and the rotation of the polarization direction by Faraday rotation ( $\propto \lambda^2$ ), which generally reduces the degree of observed polarization to  $< 20\%$ .

### 1.4.2 Free-free emission

The free-free emission (or *bremssstrahlung*) is the radiation which arises from collisions between thermal electrons and protons or other ions in the warm ionized interstellar medium, near the star forming regions of our Galaxy. The ultraviolet radiation from newly formed hot stars ionized the surrounding neutral gas.

The electrons velocity distribution is well described by a Maxwellian at a temperature  $T$  ( $T \geq 10^4$  K):

$$N_e(v)dv = 4\pi N_e \left( \frac{m_e}{2\pi kT} \right)^{3/2} v^2 \exp\left(-\frac{m_e v^2}{2kT}\right) dv \quad (1.24)$$

Taking into account the total distribution of the electrons and the ions, it is possible to compute the emissivity per volume unit along the line of sight:

$$\alpha(\nu) = \frac{4e^2}{3m_e hc} \left( \frac{2\pi}{3km_e} \right)^2 N_e N_i Z^2 T_e^{-1/2} \nu^{-3} (1 - e^{h\nu/kT_e}) \langle g_{ff} \rangle \sim 0.018 Z^2 N_e N_i T_e^{-3/2} \nu^{-2} \langle g_{ff} \rangle \quad (1.25)$$

where the approximation comes from the fact that  $h\nu \ll kT_e$  and  $\langle g_{ff} \rangle$  is the Gaunt factor averaged on the velocities.  $N_e$  and  $N_i$  are the number of electron and ions per unit of volume,  $h$  is the Planck constant and  $k$  is the Boltzmann constant.  $m_e$  is the electron mass, of charge  $e$  and temperature  $T_e$ , while  $\nu$  is the frequency of observation. The optical depth is then derived computing the integration along the line of sight:

$$\tau(\nu) = 0.018 T_e^{-3/2} \nu^{-2} \int N_e N_i dl \langle g_{ff} \rangle \quad (1.26)$$

We can assume to have a neutral plasma, with an electron temperature which is constant along the line of sight and larger than 20 K (the warm ionised medium is at about  $10^3 - 10^4$  K): if the frequency of observation is smaller than 100 GHz and the Gaunt factor can be factorized as a product of a power law of the frequency and the electron temperature, we obtain the following approximation for the optical depth:

$$\tau(\nu) \sim 0.08235 T_e^{-1.35} \left( \frac{\nu}{1\text{GHz}} \right)^{-2.1} \left( \frac{EM}{\text{cm}^{-6}\text{pc}} \right) \quad (1.27)$$

where the quantity  $EM = \int N_e N_i dl$  is the Emission Measurements. The brightness spectrum has a spectral index of  $2.10 \pm 0.03$  and depends slightly on the temperature, the plasma density and the frequency of observation, dependencies which are all included in the Gaunt factor. From the radiative transfer equation, we finally derive the definition of the brightness temperature for the free-free emission:

$$T_b^{ff} = T_e(1 - e^{-\tau}) \sim \tau_\nu T_e \sim 26 \mu\text{K} \left( \frac{10\text{GHz}}{\nu} \right)^{2.1} \left( \frac{10^4\text{K}}{T_e} \right)^{0.35} \left( \frac{EM}{\text{cm}^{-6}\text{pc}} \right) \text{ for } \tau \ll 1 \quad (1.28)$$

A discrete component is generally distinguished from a diffuse one. The first one is connected to the regions of ionized Hydrogen (HII) mainly concentrated along the Galactic plane ( $|b| < 5^\circ$ ), although present also at higher latitudes, like the Orion Nebula.

By collecting the information from previously published lists and catalogs, Paladini et al. (2002)

produced a comprehensive catalog of 1442 Galactic HII regions, which they use to produce a Synthetic Catalog at 2.7 GHz. These regions are particularly useful since they can be used to calibrate polarization CMB instruments, with an angular resolution smaller than  $1^\circ$ : indeed, the free-free emission is unlikely to be polarized. The maximum of polarization is  $\sim 5-10\%$  (Rybicki & Lightman, 1986), but since most of the compact HII regions are symmetric, the integrated linear polarization is canceled out. Indeed, Dickinson et al. (2007) found values lower than 0.6%.

A typical example of this kind of objects is the Rosette Nebula (NGC 2237-38-39) located in the anticenter region of the Galaxy: it is an evolved HII region, with an expanding shell of gas at a distance of 1.1 kpc and moving at a few tens of  $\text{km s}^{-1}$ . The electron gas reaches a temperature of  $\sim 8000$  K, which is the equilibrium temperature between energy loss by line radiation and energy gain from the ultraviolet radiation of the recently formed central star cluster (NGC 2244). The radio spectrum of the nebula fits the index given above, over a wide frequency range, as expected for an optically thin ionized region ( $\tau_\nu \ll 1$ ). Two contrasting spectra are instead seen in adjacent HII regions (G291.3-0.7 and G291.6-0.5) on the Southern Galactic plane observed by Dickinson et al. (2007): they are associated with NGC 3576 and NGC 3603, and they have a typical optically thin spectrum, except than at low frequencies where the spectrum becomes that of a black body ( $\alpha = 2.0$ ).

The diffuse component, instead, is the most unknown emission among the Galactic foregrounds. At radio frequencies, indeed, it is easily identifiable near the Galactic plane, while it is not the case at higher latitudes, due to the mixing with the synchrotron emission. Theoretically, it should be easy to separate the two components since they have different frequency spectra. However at frequencies lower than 10 GHz, the synchrotron emission is dominant at high and medium latitudes. At higher frequencies instead, where the free-free emission should overcome the synchrotron one, the signal is too faint, making its detection very difficult. This is why we do not have a full-sky map of this component. Currently, most of the knowledge of the thermal bremsstrahlung comes from the spectral line emission of the ISM. At optical wavelengths, the Balmer lines of Hydrogen are used to measure the surface brightness of the ionized emission over the entire Galaxy. Specifically, this is the case for the  $H\alpha$  emission. Where this emission is not available due to the dust absorption, the gap can be significantly covered by the contribution of the corresponding radio recombination lines (RRLs), although very weak.

The intensity of the  $H\alpha$  is given by

$$I_{H\alpha} = 0.36R \left( \frac{EM}{\text{cm}^{-6}\text{pc}} \right) \left( \frac{T_e}{10^4\text{K}} \right)^{-\gamma} \quad (1.29)$$

where  $\gamma = 0.9$  for  $T_e \leq 2.6 \times 10^4$  K and  $\gamma = 1.2$  for  $T_e > 2.6 \times 10^4$ . Finally,  $1R = 1 \text{ Rayleigh} = \frac{10^6}{4\pi} \text{ photons}/(\text{cm}^2 \text{ s sr}) = 2.41 \times 10^{-7} \text{ erg}/(\text{cm}^2 \text{ s sr})$  at the wavelength of  $\lambda_{H\alpha} = 6563\text{\AA}$ . Combining Equation 1.29 with Equation 1.28, we have the following relation between the  $H\alpha$  intensity and the free-free brightness temperature:

$$T_b^{ff} = 1.68\mu\text{K} < g_{ff} > \left( \frac{T_e}{10^4\text{K}} \right)^{0.4 \text{ to } 0.7} \left( \frac{\lambda}{\text{cm}} \right)^2 \left( \frac{I_{H\alpha}}{R} \right) \sim 7\mu\text{K} \left( \frac{T_e}{10^4\text{K}} \right)^{0.55 \text{ to } 0.85} \left( \frac{\lambda}{\text{cm}} \right)^{2.1} \left( \frac{I_{H\alpha}}{R} \right) \quad (1.30)$$

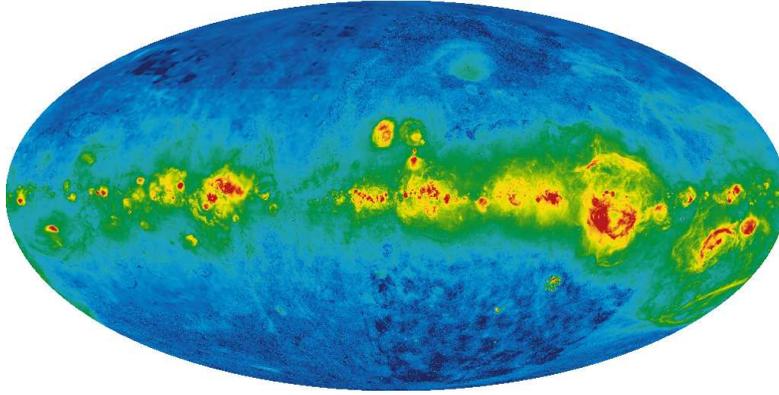


Figure 1.9: Data from WHAM, VTSS and SHASSA have been combined by Finkbeiner (2003) to form an all-sky composite  $H\alpha$  map. No correction has been made for extinction effects. The  $H\alpha$  intensity from 0.03 to 160 R (in logarithmic scale) is provided in units of Rayleigh.

Clearly, it is necessary to assume a value of  $T_e$  appropriate to the area of the sky considered. RRLs data for HII regions along the Galactic plane, show a significant increase of  $T_e$  with the Galactocentric distance,  $R$ , whose dependence has been estimated by Paladini et al. (2004) to be:

$$T_e = (4170 \pm 120) + (314 \pm 20)R. \quad (1.31)$$

However, there are also indication of the fact that  $T_e$  of diffuse HII regions is actually different from the regions of the plane. From observation and theoretical studies, we expect the diffuse gas to have a higher  $T_e$  than the HII regions with ionizing stars. Nevertheless a value of  $T_e = 8000$  K seems to be reasonable.

A full-sky map of the  $H\alpha$  emission can now be made by combining the Wisconsin H-Alpha mapper (WHAM) northern Fabry Perot survey of Haffner et al. (2003) with the Southern H-Alpha Sky Survey Atlas (SHASSA) filter survey of Gaustad et al. (2001). The WHAM survey has an angular resolution of  $1^\circ$ , while the SHASSA resolution is 6 arcmin, although it does not cover the region of the sky at scales  $\geq 10^\circ$ . By smoothing the latter survey to a resolution of  $1^\circ$  and applying a baseline correction of cosecant form at declination  $< -30^\circ$ , an all-sky map can be constructed (Dickinson et al., 2003). An other map has been create by Finkbeiner (2003) at a resolution of 6 arcmin, assembling the surveys previously described with the Virginia Tech Spectral-Line Survey (VTSS)Dennison et al. (1998). In this case, the WHAM data have been used to calibrate the zero level. The resultant map is shown in Figure 1.9: although the emission is mainly concentrated on the plane, there is also significant  $H\alpha$  emission at intermediate Galactic latitudes. Particularly, a large emission is seen in the Gould Belt system, which includes the Orion region in the southern Galaxy ( $l = 180^\circ \pm 30^\circ$ ) and Sagittarius-Scorpio region in the north ( $l = 0^\circ \pm 30^\circ$ ).

In general, although the two signal are well correlated, it is not easy to extrapolate the free-free emission from the  $H\alpha$  emission. First of all, the data are contaminated by the  $H\alpha$  emission of the earth atmosphere. The two components can be distinguished only taking advantage of the Doppler effect, due to the motion of the Earth with respect to the Solar system: it induces a separation in the lines. A

further complication, however, introduced by the fact that this effect is reduced by the OH emission at a wavelength of  $6568.68 \text{ \AA}$ , very closed to  $\lambda_{H\alpha}$ . Anyway, the largest source of uncertainty is the dust absorption of the  $H\alpha$  emission. This can be estimated using the Galactic absorption calibrated against the  $100 \mu\text{m}$  dust emission (Schlegel et al., 1998). Assuming the well known reddening curve in the optical range of  $A(H\alpha) = 0.81A(V)$ , the absorption at the  $H\alpha$  wavelength is given by

$$A(H\alpha) = 2.51 E(B - V) = 0.0462 D^T f_d \text{ magnitudes}, \quad (1.32)$$

where  $D^T$  is the  $100 \mu\text{m}$  surface brightness in units of  $\text{MJy sr}^{-1}$  and  $f_d$  is the dust fraction responsible of the absorption along the line of sight. Specifically, a correction of half this value is sufficient, assuming that dust and gas are uniformly mixed along the line of sight. However, Dickinson et al. (2003) derived a value for  $f_d$  which is not consistent with this hypothesis: they instead adopted a value of  $f_d = 0.33$ . Besides, Davies et al. (2006) and Banday et al. (2003) favour  $f_d = 0$  at high latitudes ( $|b| > 20^\circ$ ).

Using Equation 1.32, Dickinson et al. (2003) converted the  $D^T$  dust map to an absorption map which can be applied to the  $H\alpha$  intensity  $I_{H\alpha}$ , in order to obtain the proper correction:

$$I_{H\alpha}^{corr} = I_{H\alpha} \times 10^{D^T \times 0.0185 \times f_d} \quad (1.33)$$

### 1.4.3 Thermal dust emission

Dust is diffusely present in the entire Galaxy, with a concentration along the Galactic plane and near the stellar formation regions. Dust grains, heated up by the interstellar radiation field, absorb UV and optical photons and re-emit them in the far IR range of frequencies. This is the dominant emission at frequencies larger than  $100 \text{ GHz}$ .

The *COBE*-DIRBE<sup>11</sup> full-sky maps at  $100$ ,  $140$  and  $240 \mu\text{m}$ , at an angular resolution of  $0.7^\circ$ , were used to trace this emission (Kogut et al., 1996b). The Infrared Astronomical Satellite (IRAS) has then produced a more sensitive full-sky map of dust emission at  $100 \mu\text{m}$  with a resolution of  $6 \text{ arcmin}$ . These data sets have been re-calibrated using the *COBE*-DIRBE data. Furthermore, artifacts due to the zodiacal light have been reduced and point sources have been removed (Schlegel et al., 1998). The result is shown in Figure 1.10: the strongest emission lies along the Galactic plane, while the Gould Belt system is visible at intermediate latitudes as well as in the  $H\alpha$  emission, although the specific structures are different.

The intensity of the radiation depends on the shape, composition and size distribution of the dust grains. Figure 1.11 shows the spectrum of a typical sample of interstellar dust (Desert et al., 1990), which includes three components:

- Grains composed by polycyclic aromatic hydrocarbons (PAHs) containing only 30 to 100 atoms, with a diameter of  $\sim 10 \text{ \AA}$ . They have a rich spectrum in the  $10 \mu\text{m}$  wavelength range.
- Very small grains: they have generally a radius of  $0.001$  to  $0.01 \mu\text{m}$  and emit at  $25 - 60 \mu\text{m}$ . They are most likely carbon-dominated grains with possibly a mix of other metals.

<sup>11</sup>Diffuse Infrared Background Experiment

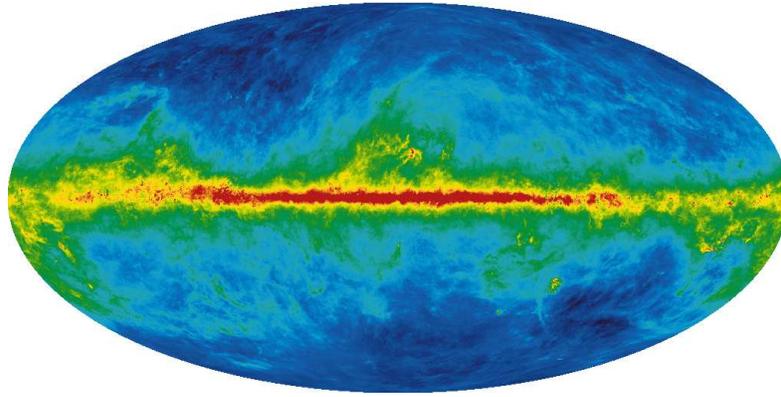


Figure 1.10: Using a combination of *COBE* and *IRAS* data, Finkbeiner et al. (1999) derived models for thermal dust emission at microwave and sub-millimeter frequencies. This map is based on their best-fit two-component Model 8, at a frequency of 94 GHz. It has a resolution of 6.1 arcmin.

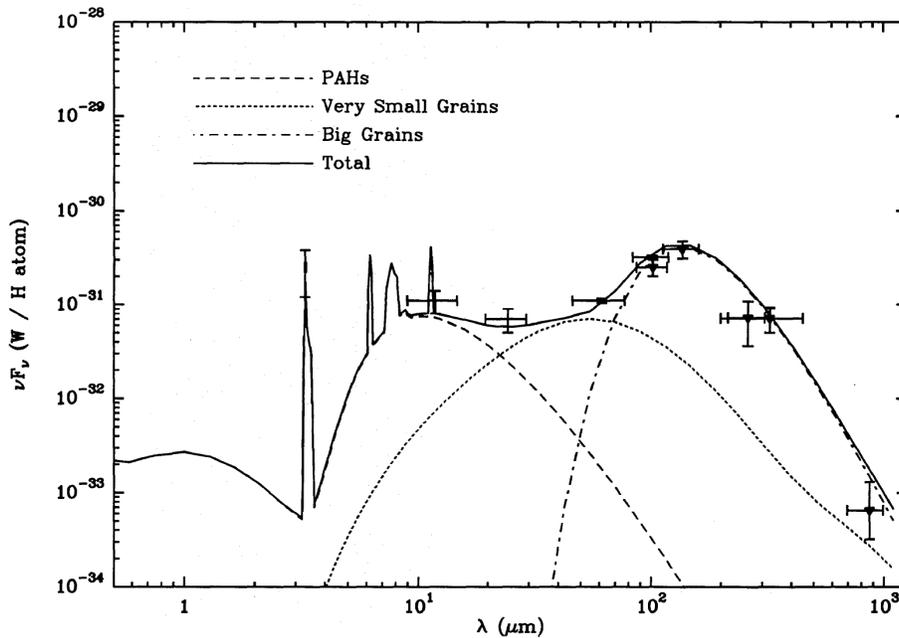


Figure 1.11: Composite spectrum of the dust in the ISM. Three components are identified. Taken from Desert et al. (1990).

- Large grains up to  $1 \mu\text{m}$  in diameter: their emission is well described by a black body peaking at  $100 - 200 \mu\text{m}$ . Silicates are probably the main component of these grains, although in order to explain the optical emission, carbon has to be included as well.

There is also a strong spatial correlation between the dust emission and the gas of the interstellar medium: indeed, the dust radiation is associated with the line emission of the gas, specifically with the neutral hydrogen (HI). Therefore, the dust emission depends also on the chemical composition of the interstellar gas and the quantity of dust grains with respect to the gas. Consequently, the emission presents strong spatial variations.

Let us assume a dust grain population with a size distribution  $f(a)$ , where  $a_{\min} \leq a \leq a_{\max}$ . We can

define  $Z_d$  the ratio between the mass of the gas and the dust. The dust emissivity is then given by (Dwek et al., 1997):

$$I(\nu) = \frac{\mu m_H Z_d}{\langle m_d \rangle} N_H \int_{a_{min}}^{a_{max}} \int P(a, T) B_\nu(T) dT \quad (1.34)$$

where  $\mu$  is the atomic weight of the gas,  $N_H$  the total column density,  $\langle m_d \rangle$  is the mean mass of the grains population and  $\sigma(a, \nu)$  is the cross section of the dust particles with radius  $a$  and at a frequency  $\nu$ .  $B_\nu$  is the surface brightness given by the Planckian law at the temperature  $T$  and  $P(a, T)$  is the probability for a dust particle of having a temperature within  $T$  and  $T + dT$ . This expression, can then be approximated by a modified black body law:

$$I(\nu) \propto \nu^\alpha B_\nu(T_d) \quad (1.35)$$

where  $\nu$  is the frequency of observation,  $\alpha$  the emissivity index and  $T_d$  is the dust temperature. There are strong variations of the temperature throughout the interstellar medium, depending on the strength of the local radiation field: it varies from 10 K to 25 K, with most of the dust being at 15-20 K. However, in the HII regions, the temperature can be larger than 40 K, due to the strong radiation produced by the central ionizing stars.

It is interesting to define the value of the spectral index  $\alpha$ . For this purpose, several studies have been undertaken and different models have been proposed. Dwek et al. (1997) have determined the dust emissivity spectrum for wavelengths of 3.5 to 1000  $\mu\text{m}$ , using the *COBE*-DIRBE and FIRAS data, together with the IRAS ones. From this spectrum, they derived a composition model of the Galactic dust which includes graphite and silicates, but also PAHs. The first ones dominate at wavelengths larger than 140  $\mu\text{m}$  with a temperature of 17-20 K for graphite, and 15-18 K for silicates. Kogut et al. (1993) and Kogut et al. (1996a) using the *COBE*-DMR and DIRBE data, instead, have found the dust spectrum at high Galactic latitudes, to be well described by a modified black body with a spectral index equal to 2 and a temperature of 18 K. The DIRBE data at 100 and 240  $\mu\text{m}$  have been used also by Schlegel et al. (1998), who derived a map of the dust temperature distribution on the entire sky: this varies within 17 and 21 K. Therefore, the same shape of the spectrum ( $\sim \nu^2$ ) has been determined.

However, a single power law is actually not able to fit properly the FIRAS data at frequencies larger than 800 GHz. This result has introduced the idea of multiple components models, partially developed by Finkbeiner et al. (1999). They elaborated a series of empirical models (FDS), guided by physics rather than physical like the Dwek model, based on the *COBE*-DIRBE 100 and 240  $\mu\text{m}$  maps tied to the *COBE*-FIRAS spectral data in the range 0.14 to 3.0 mm. The study was supported by a laboratory analysis of Pollack et al. (1994) which demonstrated that different kinds of grains contribute in a different way to the total emissivity and dominate at different frequencies. Specifically, main contributors are carbons and amorphous silicates.

Among the eight models proposed by Finkbeiner et al. (1999), the number 7 and 8 have shown the better agreement with the FIRAS data. They are both two components models: in the first case, the two components have a temperature of 9.6 and 16.4 K, with a spectral index of 1.5 and 2.6 respectively. For the second one, instead, the temperature is equal to 9.4 and 16.2 K and the spectral index is 1.67 and

2.7 respectively. The map shown in Figure 1.10 is actually the one derived using the model 8. It is at 94 GHz and has a claimed nominal resolution of 6.1 arcmin.

Over the *WMAP* range of frequencies, the thermal dust emission in antenna temperature is sometimes considered to scale with a power-law dependence  $T_A \sim \nu^{\beta_d}$ , with values of 1.7 – 2.2 for  $\beta_d$  assumed by different authors. In the analysis done for this work, we adopted the FDS8 predicted emission at 94 GHz, as the reference template for dust emission, and computed a single global scaling between this template and the *WMAP* frequencies.

#### 1.4.4 Anomalous Dust emission

Analysing the *COBE*-DMR data at 31.5, 53 and 90 GHz, Kogut (1999) found a statistically significant correlation with the DIRBE data at 100, 140 and 240  $\mu\text{m}$ . He basically discovered the existence of a dust correlated emission for scales larger than  $7^\circ$ , not explained by a simple thermal dust model. A two-component fit of the correlated *COBE* data to a model with dust (with spectral index  $1.5 < \beta_{dust} < 2$ ) and radio emission, returned a spectral index  $\beta_{radio} \sim 2.1_{-0.8}^{+0.6}$  for the unknown radio component. This result induced them to associate the radiation to free-free emission, whose theoretical spectral index is  $\beta = -2.15$ . Later, the same result has been confirmed and extended in frequencies (down to 14 GHz) by Leitch et al. (1997) and de Oliveira-Costa et al. (1997) with the OVRO<sup>12</sup> and Saskatoon data respectively. However, the free-free hypothesis was discarded soon: given the correlation between the free-free emission and the  $\text{H}\alpha$  emission, the proposed explanation would require an analogous correlation between the anomalous radio emission and the  $\text{H}\alpha$  emission. Instead, the first one was measured to be 5-10 times larger than the second one.

A deep scan around the north celestial pole at 14.5 and 31.7 GHz by Leitch et al. (1997) showed a strong correlation with 100  $\mu\text{m}$  dust and free-free like spectrum. However, they also confirmed a lack of associated  $\text{H}\alpha$  emission at the expected level. Therefore the emission was proposed to be associated with ionized gas at  $T_e = 10^6$  K. However, again, this hypothesis has been ruled out: such a hot gas would produce X-rays, but no significant detection of them has been measured by ROSAT<sup>13</sup> at 1/4 KeV. Furthermore, temperatures of  $10^6$  K require energy higher than  $10^2$  times the one released by a Supernova explosion.

All these results substantiated the model of electric dipole emission from small spinning dust grains proposed by Draine & Lazarian (1998a,b). The same PAHs grains which absorb stellar radiation and cool down by radiating in IR bands, undergo collisions with gas-phase species and can become charged. Therefore, they carry a significant permanent dipole moment (of a few tens of a Debye). As a result of the frequent collisions, the spinning PAHs will produce electric dipole emission from rotational levels. The spinning dust paradigm has been increasingly accepted as the explanation of the ‘anomalous component’: this would explain naturally both the correlation with dust, since the same grains are responsible also for the thermal emission, and the absence of correlated  $\text{H}\alpha$  radiation.

Draine & Lazarian (1998b) have computed the emissivity  $j_\nu$  of the grains, for different size  $a$  and

<sup>12</sup>Owens Valley Radio Observatory

<sup>13</sup>Röntgensatellit

electric dipole moment  $\beta$ , in the 10-100 GHz range. Assuming the angular velocity of the spinning grains to have a Maxwellian distribution:

$$j_\nu = \left(\frac{8}{3\pi}\right)^{1/2} \frac{1}{c^3} \int da f(a) \frac{\mu^2 \omega^6}{\langle \omega^2 \rangle^{3/2}} \exp\left(\frac{-3\omega^2}{2 \langle \omega^2 \rangle}\right) \quad (1.36)$$

where  $\mu$  is the intrinsic electric dipole moment of the grain without charge ( $\mu \sim \beta N^{1/2}$ , where  $N$  is the total number of atoms of the grain),  $\omega$  is the angular velocity and  $f(a)$  is the size distribution of the grains.

The emissivity was found to be sensitive to the properties of the gas such as the density and the temperature. Specifically, they distinguished the contribution produced by five typical environments in the ISM with various canonical densities and temperatures in order to compute corresponding spinning dust model spectra:

- CNM cold neutral medium
- WNM warm neutral medium
- WIM warm ionized medium
- MC molecular clouds
- DC dark clouds

Depending of the component, the emissivity and the peak frequency change, although the latter lies always between 20 and 30 GHz. Figure 1.12 shows the spectrum expected for these components. Recently Ali-Haïmoud et al. (2009) presented a comprehensive treatment of the spectrum of electric dipole emission from spinning dust grains, updating the commonly used model of Draine & Lazarian.

Following the theoretical expectations, the distinguishing characteristic of spinning dust emission would be the turn-over in the spectrum at lower frequencies, around 10-15 GHz depending on the dimension of the grains. The first detection of this feature came out combining the Tenerife data at 10 and 15 GHz with the *COBE*-DMR data (de Oliveira-Costa et al., 1997): a peak was measured at  $\sim 20$  GHz in the radio emission, per unit of 100  $\mu\text{m}$  dust emission. This result was confirmed by combining the same data set with the *WMAP* data, thanks to the larger frequency coverage. Specifically, it was demonstrated that the low-frequency part of the spectrum was inconsistent with the synchrotron emission for intermediate latitudes. A similar analysis was carried out by Banday et al. (2003) with the DMR and the 19 GHz data.

The existence of this component has been confirmed more recently by the analysis of the *WMAP* data performed by Bonaldi et al. (2007), Davies et al. (2006), and Miville-Deschênes et al. (2008). Moreover, Ysard et al. (2010) have carried out a correlation analysis of the IRAS data with a full-sky spinning dust template at 23 GHz produced by Miville-Deschênes et al. (2008). They found a strong correlation between the template and the 100  $\mu\text{m}$  IRAS band and, more important, with the 12  $\mu\text{m}$  band. This has been interpreted as the confirmation of the spinning dust model, since small dust grains (PAHs) carry the 12  $\mu\text{m}$  flux.

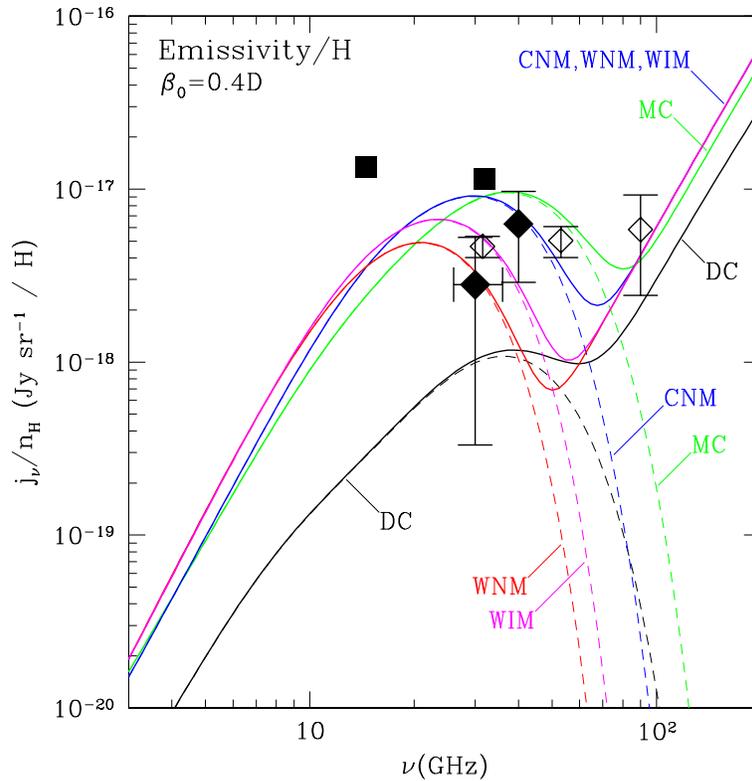


Figure 1.12: The spectrum of the spinning dust. The radio emissivity is given in terms of Jy/sr per unit hydrogen column density. The dotted curves show the spectra for the different components. The solid lines are the spectra when the corresponding thermal emission is added. Observational data from *COBE*-DMR (Kogut et al., 1996a), Saskatoon (de Oliveira-Costa et al., 1997) and OVRO (Leitch et al., 1997) are also shown. The plot is taken from Draine & Lazarian (1998b).

Beside the spinning dust grains, other scenarios have been proposed as possible explanations for the anomalous component. Fullerenes of 60-200 carbon atoms have been suggested to be responsible for the anomalous component (Iglesias-Groth, 2006). Large grains of submicron radius may have a significant contribution in the 10 to 100 GHz range. Draine & Lazarian also proposed that magnetic dipole emission may be emitted by large grains over a broad range between 10 and 100 GHz. Spontaneous magnetization is required in order to have a peak at the observed frequency: this could be possible thanks to the presence of magnetite and metallic iron. Currently, magnetic dipole emission can not be ruled out as a possible explanation for the anomalous component of dust at microwave frequencies.

Some hints about the processes that occur in the anomalous emission of dust can be provided by observation of individual compact dust clouds. Watson et al. (2005) discovered very strong anomalous emission in the Perseus molecular cloud (G 159.6-18.5) with the COSMOSOMAS<sup>14</sup> experiment at 11-17 GHz: the derived spectrum is consistent with the spinning dust model. Furthermore, a strong correlation has been found with the IRAS 25  $\mu\text{m}$  band suggesting that the anomalous emission is produced by *very small spinning dust grains*. More recently, a morphological study of the Lynds Dark

<sup>14</sup>COSMOlogical Structures On Medium Angular Scales

Source	Dust emissivity $\mu\text{K (MJy/sr)}^{-1}$	Reference
<b>HII regions</b>		
6 HII regions (mean)	$3.3 \pm 1.7$	Dickinson et al. (2007)
LPH96	$5.8 \pm 2.3$	Dickinson et al. (2006)
<b>Cool dust clouds</b>		
15 regions <i>WMAP</i>	$11.2 \pm 1.5$	Davies et al. (2006)
All-sky <i>WMAP</i>	$10.9 \pm 1.1$	Davies et al. (2006)
LDN1622	$24.1 \pm 0.7$	Casassus et al. (2006)
G159.618.5	$17.8 \pm 0.3$	Watson et al. (2005)

Table 1.1: Comparison of 100  $\mu\text{m}$  dust emissivities for HII regions and cooler dust clouds, from data at or near 30 GHz (Dickinson et al., 2007). Emissivities, in units  $\mu\text{K (MJy/sr)}^{-1}$ , have been normalised to 31 GHz.

Nebula LDN1622 was carried out by Casassus et al. (2006) at 31 GHz with CBI. Combining these observations with the *WMAP* data, a well defined spectrum consistent with the spinning dust emission is produced, with a peak around 30 GHz. They also showed a strong correlation with all the IRAS data, in particular with the 25  $\mu\text{m}$  band, which again supports the hypothesis of spinning dust emission from very small grains.

In order to confirm this idea, it is important to study the anomalous emission from compact HII regions: in fact very small dust grains are not expected to exist in those regions. Therefore, any detection of anomalous emission would discard the hypothesis of very small grains and would leave space to other interpretations. Dickinson et al. (2006) observed the LPH96 (G201.66+1.64) with the CBI at 31 GHz and found that the spectrum is basically consistent with simple free-free. A more extensive study of the anomalous emission in HII regions in the southern hemisphere, has been undertaken by Dickinson et al. (2007) (see table 1.1). Again the derived spectrum was found to be consistent with free-free emission, although some excess of emission has been measured in six regions. This excess could be associated with spinning dust emission or other emission mechanisms.

We would clearly need an accurate template for this anomalous emission to model the foreground amplitudes at *WMAP* frequencies, but this remains elusive. Finkbeiner (2004a) and Davies et al. (2006) have proposed that the FDS template modulated by some power-law of the dust temperature provides a better fit than the unperturbed sky map. However, Bonaldi et al. (2007) suggest that although the anomalous emission is tightly correlated with thermal dust, the correlation is not perfect. Nevertheless, we expect that the FDS model provides a representative template and we adopted it for our analysis.

#### 1.4.5 The *WMAP* Haze

The first detection of the so called ‘*WMAP* Haze’ is due to Finkbeiner (2004a). Analysing the *WMAP* data, he found an approximately circular excess of diffuse emission centered around the Galactic center of radial extent  $\sim 20^\circ$  and luminosity  $L_{\text{haze}} \sim 10^{37} \text{ erg s}^{-1}$  between 20 and 60 GHz. Later, Dobler &

Finkbeiner (2008a) defined its profile as well described by:

$$\mathbf{h} \propto \begin{cases} \frac{1}{r} - \frac{1}{r_0} & \text{for } r < r_0; \\ 0 & \text{for } r > r_0, \end{cases} \quad (1.37)$$

where  $r$  is the angular distance to the Galactic center and  $r_0$  is arbitrarily set equal to 45 degrees.

Initially attributed to free-free emission produced by  $\gtrsim 10^5$  K gas (Finkbeiner, 2004a), this hypothesis was immediately discarded since a counterpart emission in X-ray was not detected (Finkbeiner, 2004b). Finkbeiner (2004b) showed instead that the energetics and profile of the haze emission could be explained by synchrotron emission from the by-products of dark matter annihilation for a generic weakly interacting ( $\sim 100$  GeV) particle with a standard dark matter halo profile. This idea has been pursued further in several subsequent works (Hooper et al., 2007, 2008; Zhang et al., 2009; Cholis et al., 2009; Cumberbatch et al., 2009).

Other studies carried out different interpretations. The haze has been associated with synchrotron emission produced by pulsars (Zhang et al., 2009; Kaplinghat et al., 2009) or, more recently, by cosmic rays from astrophysical sources (McQuinn & Zaldarriaga, 2010). However, McQuinn & Zaldarriaga (2010) also pointed out that both these interpretations are not able to justify the following observations:

1. Dobler & Finkbeiner (2008a) found the haze spectrum in the microwave to be hard, measuring a power-law electron spectrum with index  $p \sim 2$ . Such a spectrum is much harder than the galactic synchrotron towards other directions (where  $p \sim 3$ ).
2. The distribution of known galactic cosmic ray sources appears not to be sufficiently concentrated toward the Galactic center in order to create the haze (Finkbeiner, 2004b; Dobler & Finkbeiner, 2008a; Zhang et al., 2009). The distribution of supernovae and pulsars is estimated to peak at  $\sim 4$  kpc from the Galactic center, or  $\sim 30^\circ$  (Lorimer et al., 2006). In contrast, the haze intensity is strongly increasing with decreasing angle at  $< 20^\circ$  from the Galactic center (Dobler & Finkbeiner, 2008a).
3. A similar excess towards the Galactic center is not present at 408 MHz. If the haze is from synchrotron, this suggests that the source of the haze does not contribute significantly to the population of cosmic ray electrons at  $\sim 5$  GeV. In addition, the 408 MHz emission appears very disk-like without a significant enhancement in the haze region.

On the other hand, these criteria can be satisfied by dark matter annihilation (Finkbeiner, 2004b), although other authors argued that the annihilation cross-section would need to be significantly boosted (Cumberbatch et al., 2009).

A preliminary analysis (Dobler et al., 2009) of the data produced by the Fermi Space Telescope has shown the existence of a  $\gamma$ -ray ‘haze’ with a similar spatial morphology of the *WMAP* haze. It has been interpreted as due to the same hard population of electrons that generate the haze in the microwave (Dobler et al., 2009) which interact via inverse Compton effect with the starlight. As the *WMAP* Haze, also the  $\gamma$ -ray haze could be explained in terms of dark matters annihilation (Cholis et al., 2009).

However, the Fermi collaboration has not claimed any excess in the galactic center region over the standard diffuse  $\gamma$ -ray background (Abdo et al., 2010; Casandjian et al., 2009). Moreover, Linden & Profumo (2010) pointed out that what was measured by Dobler et al. (2009) could be just an artifact of the analysis performed, i.e. an incorrect foreground removal. In fact, they used templates which seem to be inappropriate and to underestimate both the  $\pi^0$  decay and inverse Compton scattering (ICS) contributions to the  $\gamma$ -ray emission, specifically in the Galactic center region.

The reliability of the models used to trace the different components of the sky is a crucial point also for the detection of the haze in the microwaves. Finkbeiner (2004a) identified the haze as the result of an analysis where the *WMAP* maps at different frequencies were fitted with a model of the foreground components of the sky, as described by specific templates. In particular, the Haslam map at 408 MHz has been used to trace the synchrotron emission, although the *WMAP* observations lie at much larger frequencies. Relying on such a crude extrapolation of the morphology of synchrotron emission can thus potentially introduce unphysical residuals. Specifically, Mertsch & Sarkar (2010) have shown by means of simulations, that this leads to residuals of the same order as the claimed haze, along the Galactic disk. Therefore, they support the idea that the *WMAP* haze is actually an artifact of inappropriate template subtraction, rather than evidence of dark matter annihilation.

Indeed, other authors (Cumberbatch et al., 2009), have noted that the significance of the *WMAP* haze is substantially reduced by allowing for spatial variation in the frequency-dependence of synchrotron emission in the inner and outer parts of the Galaxy.

Doubts about the actual existence of the haze come also from a recent study by Gold et al. (2010), where they attempted to detect the haze in polarization in the *WMAP* sky maps, although unsuccessfully. Moreover, the total polarized synchrotron emission within  $\sim 10, 20$  and  $30$  degrees from the Galactic center is consistent with a spectral index of  $-1.2$ , which is much softer than the value proposed by Dobler & Finkbeiner (2008a): they found that the haze contributes roughly half of the diffuse intensity in synchrotron within  $10^\circ$  (and a decreasing fraction at larger angles). Therefore, Gold et al. (2010) seems to suggest that if the haze exists, it must contribute a much smaller fraction to the polarized emission.

However, as pointed out by McQuinn & Zaldarriaga (2010), the Gold et al. (2010) results can be still consistent with the hypothesis of the *WMAP* haze as a new independent synchrotron emission. First of all, the possibility that the haze emission is not highly polarized is actually plausible, since in the inner regions of the Galaxy the polarized fraction of the synchrotron emission is lower than outer regions, probably due to the strong entanglement of the Galactic magnetic field along the Galactic plane. Furthermore, local structures in the haze region could be dominant in polarization with respect to the haze. More importantly, the structures we know be responsible for the polarized synchrotron emission were not excluded in the analysis performed by Gold et al. (2010), but in the one of Dobler & Finkbeiner (2008a). Finally there is the possibility that the estimation of the spectral index for the electron was wrong and that it is not so hard as it has been claimed. Based on the analysis of Gold et al. (2010), in order to reproduce the amplitude of the haze, a spatial variation in the synchrotron spectral index of only  $0.25$  between  $408$  MHz and  $23$  GHz is required.

A more recent study of the Fermi data proposed by Su et al. (2010) seems to give the stronger

confirmation in favour of the *WMAP* Haze. Indeed, they has demonstrated the existence of two large gamma-ray bubbles, extending 50 degrees above and below the Galactic center, with a width of about 40 degrees in longitude. The hard spectrum emission associated with these bubbles seems to be spatially correlated with the *WMAP* Haze, and more importantly, the edges of the bubbles line up with features in the ROSAT X-ray maps at 1.5-2 keV. This result would actually confirm the Haze as a new synchrotron emission and would associate it to a large episode of energy injection in the Galactic center, such as past accretion events onto the central massive black hole, or a nuclear starburst in the last  $\sim 10$  Myr. Besides, the hypothesis of dark matter annihilation or decay would be consequently ruled out.

In the analysis proposed in Chapters 3 and 4, we tried to address the issue of the existence of the *WMAP* Haze, attempting to confirm the results proposed by Dobler & Finkbeiner (2008a). We refer the reader to these chapters for a more detailed analysis of our results.

### 1.4.6 Extragalactic emissions

Together with diffuse components, CMB data are also contaminated by extragalactic foregrounds, namely radio and far-infrared sources, and the Sunyaev-Zeldovich effect due to clusters of galaxies. In the next sections, we will give a brief description of these effects to give a more complete picture of the foregrounds which contaminate the cosmological measurements. Note however, that this is not object of study of the analysis presented in this thesis. Since we focused our attention on the study of diffuse Galactic components, we neglected the extragalactic contribution just cutting out these sources from the data. This is possible applying a mask, which is generally produced according to existing catalogs. Specifically, we adopted the masks provided by the *WMAP* science team.

#### Radio and far-infrared sources

Extragalactic sources affect CMB observations both because of single objects which can be studied and individually subtracted, and because of an effective noise generated by all those sources whose emission is too low to be identified. In each resolution element, there will be  $N$  sources in a given flux density range, with a Poisson uncertainty  $N^{1/2}$ . This uncertainty in  $N$  translates to a temperature fluctuation and makes a contribution to the power spectrum at the corresponding angular scale of the CMB foreground (Toffolatti et al., 1998).

Resolvable sources are mainly extragalactic radio synchrotron emitters with a range of spectral indexes which depends on their activity. The most luminous ones can be derived from the known source counts over a wide frequency range and therefore, can be individually studied and subtracted. Furthermore, complete samples of faint sources are available to evaluate the foreground at frequencies  $\nu \leq 10$  GHz.

Figure 1.13 shows the number counts for 1.4, 5 and 8.4 GHz (Toffolatti et al., 1998) extending from 10 Jy down to flux densities as low as  $10 - 100 \mu\text{Jy}$ . The star-burst galaxies are the main contributors at the lower flux densities. Active Galactic Nuclei have generally  $S < 1$  mJy and they are found at substantial redshift. Finally, ellipticals with a steep spectrum and quasars dominate the lower frequencies.

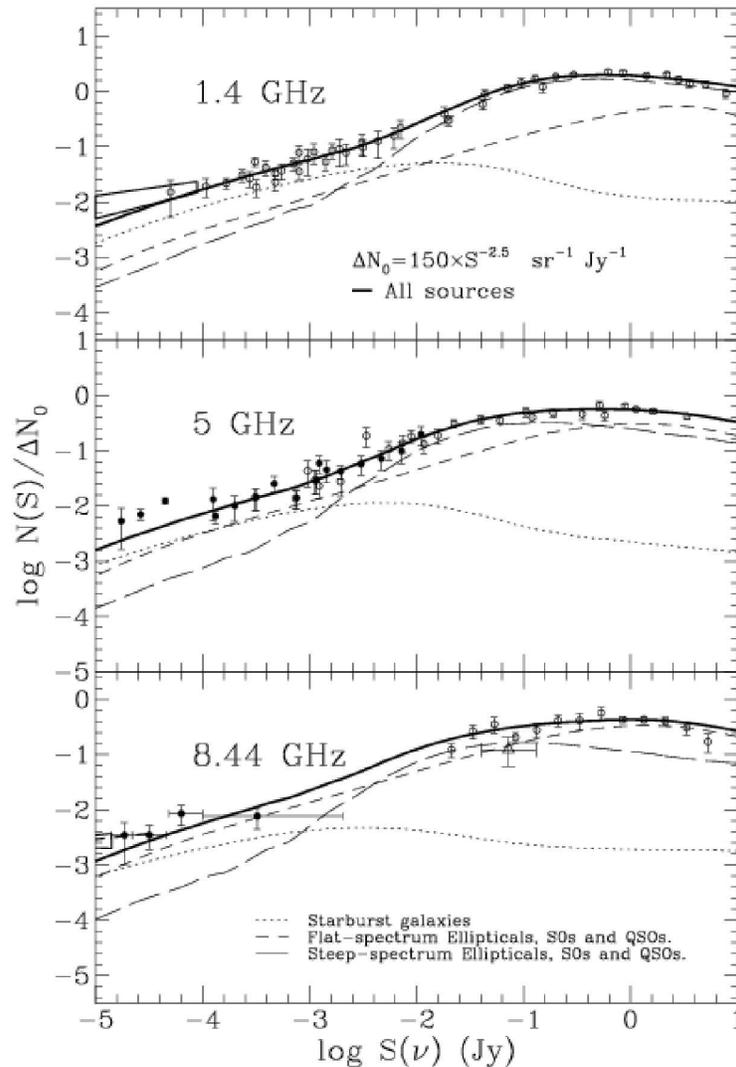


Figure 1.13: Comparison between observed and predicted differential source counts at 1.4, 5.0 and 8.44 GHz. The predictions for different classes of radio source are shown: starburst galaxies (dotted line), flat-spectrum ellipticals SOs and QSOs (dashed line) steep spectrum ellipticals SOs and QSOs (long dashed line). Taken from Toffolatti et al. (1998).

Two classes of objects are thought to be part of the unresolved point sources:

- Radio sources, meaning radio galaxies with a flat spectrum such as Cygnus A, BL Lacs objects, quasars (e.g. 3C48) and blazars. Their contribution dominates at frequencies between 30 and 100 GHz.
- Far-infrared galaxies whose emission is dominant at higher frequencies, because rich of dust. These are mainly the spirals but also SOs and ellipticals. Because of the dust, they have a spectrum increasing with the frequency:  $S(\nu) \propto \nu^{3.5}$ , where  $S(\nu)$  is the flux emitted at the frequency  $\nu$ . In this case, counts are made even more difficult by the fact that they depend on their evolution, which is again uncertain.

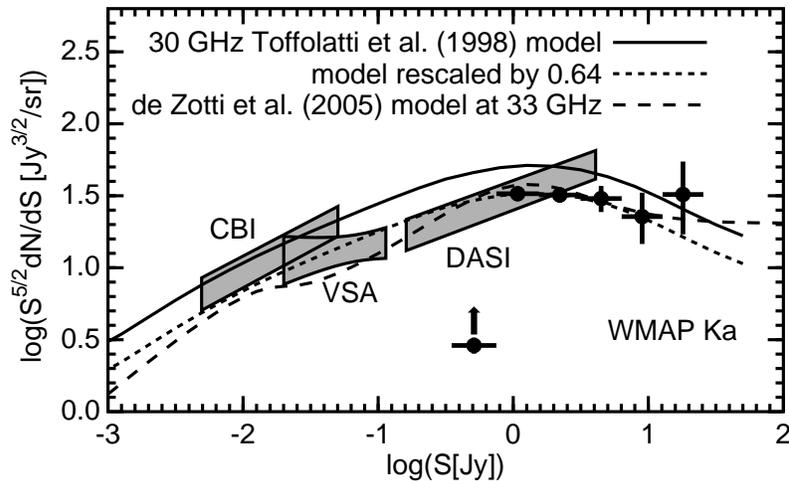


Figure 1.14: Differential source counts from the *WMAP* five-year catalog compared to the Toffolatti et al. (1998) model, and to CBI counts at 31 GHz (Mason et al., 2003), 33 GHz VSA<sup>19</sup> counts (Cleary et al., 2005), and DASI 31 GHz counts (Kovac et al., 2002). Models from Toffolatti et al. (1998) and de Zotti et al. (2005) are shown as well. Error bars for *WMAP* are statistical only. The *WMAP* catalog in the 0.35 to 0.75 Jy bin is quite incomplete, leading to the low data point with the upward arrow on the plot.

Big efforts have been done to enlarge our knowledge of unresolved point sources at the frequencies of interest of CMB measurements, and to draw a realistic model. Indeed, an incorrect determination of this contamination can lead to a significant bias on the estimation of the cosmological parameters as demonstrated by Huffenberger et al. (2006, 2008).

de Zotti et al. (2005) updated and improved the evolutionary models of Toffolatti et al. (1998), taking into account new high-frequency data from the 15-GHz 9C (Waldram et al., 2003), 18-GHz ATCA<sup>15</sup> catalogs (Ricci et al., 2004) and *WMAP* (first year; (Bennett et al., 2003)). Based on those models, Toffolatti et al. (2005) estimated the contribution of the various populations of extragalactic sources to the counts at 30 GHz (the frequency of BIMA<sup>16</sup> (Dawson et al., 2002) and CBI (Mason et al., 2003; Readhead et al., 2004) experiments) and at 150 GHz (ACBAR<sup>17</sup> experiment, Kuo et al. (2004)).

The *WMAP* data themselves have been used to identify point sources, both internally - without relying on information provided by other point sources surveys - (Wright et al., 2009), as well as in combination with other observations (Colombo & Pierpaoli, 2010).

Unresolved point sources produce a simple white noise power spectrum with the same power in all multipoles, so that the point sources fluctuations become increasingly important at small angular scales (high  $\ell$ ). It becomes comparable to the CMB power at  $\ell = 500$  at 30 GHz and  $\ell = 2000$  at 100 GHz. The point sources exceed the diffuse foregrounds at intermediate Galactic latitudes at  $\ell = 100$  and  $\ell = 700$  for 30 GHz and 100 GHz respectively (see Figure 1.15). Throughout  $\ell = 10 - 2000$ , a similar linear rise with  $\ell$  is shown in Figure 1.16 for the contribution of the dusty galaxies at frequencies  $\nu \geq 100$  GHz. Specifically, at frequencies of  $\sim 300$  GHz, it is greater than the CMB at  $\ell = 1000$ , while

<sup>15</sup>Australia Telescope Compact Array

<sup>16</sup>Berkeley Illinois Maryland Array

<sup>17</sup>Arcminute Cosmology Bolometer Array Receiver

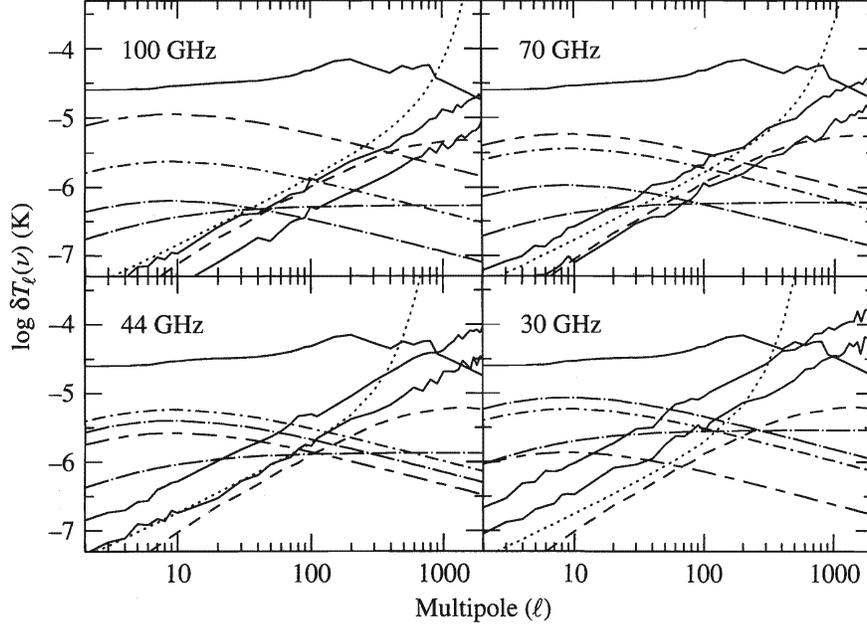


Figure 1.15: The angular power spectra of the components contributing to the foreground at frequencies in the range 30 to 100 GHz of the Planck Surveyor. The plots at each frequency are in terms of a temperature fluctuations  $\delta T_\ell = [\ell(2\ell + 1)C_\ell/4\pi]^{1/2}$  as a function of angular multipole  $\ell$ . The diagonal solid lines show the contribution of extragalactic sources. The upper of the two lines is where sources brighter than 1 Jy are identified and removed; the lower is for 0.1 Jy. The CMB power spectrum in the heavy line. The heavy dashed lines are for the anisotropies due to the SZ effect. The diffuse Galactic components estimated for Galactic latitudes  $|b| > 30^\circ$  are synchrotron (dots - long dashes line), free-free (dots - short dashes line) and thermal dust (long - short dashes line). Taken from Rubiño-Martín et al. (2010).

at 857 GHz they exceed the CMB above  $\ell = 30$ . Moreover, at the same frequency, the surface density of dusty galaxies becomes so high that they form a continuous background of emission, exceeding both the CMB and the diffuse components (Rubiño-Martín et al., 2010).

### Sunyaev-Zeldovich Effect

Another point-like source contribution to the CMB foregrounds is the Sunyaev-Zeldovich (SZ) effect (Sunyaev & Zeldovich, 1972) produced by galaxy clusters containing a hot electron gas ( $T_e \sim 10^8$  K). The inverse Compton scattering of CMB photons with this hot and diffuse electron gas trapped in the potential well of the cluster produces a systematic shift of the photons from the Rayleigh-Jeans side to the Wien side of the CMB spectrum. Therefore, the signature of this effect in the CMB frequency spectrum is a distortion with a negative temperature at low (Rayleigh-Jeans) frequencies and a positive temperature at high (Wien) frequencies. The change from negative to positive is at a frequency of 217 GHz. Qualitatively, the intensity variation is given by:

$$\Delta I = \frac{2(kT)^3}{(hc)^2} y g(x) \quad (1.38)$$

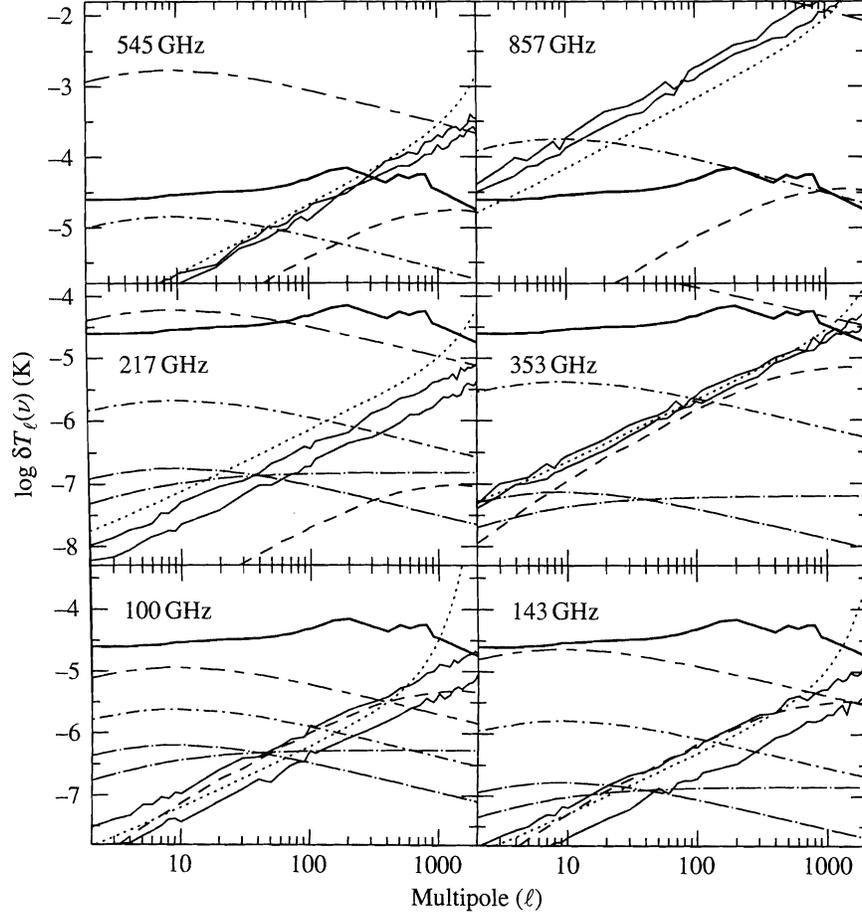


Figure 1.16: The angular power spectra of the components contributing to the foreground at frequencies in the range 143 to 857 GHz of the Planck Surveyor. The plot format for each of the components is as given in Figure 1.15. The diffuse dust component (long dash-short dash line) is the dominant diffuse foreground over this frequency range. The dusty galaxy Poisson component (liner rising plots) exceeds the CMB for all  $\ell$  values at 857 GHz. Taken from Rubiño-Martín et al. (2010).

where  $T$  in the CMB temperature and  $\nu$  is the frequency of observation. The dimensionless parameter  $y$  (*Comptonization parameter*) depends the temperature and density of the electron gas along the line of sight:

$$y = \int \frac{kT_e}{m_e c^2} n_e \sigma_T dl \quad (1.39)$$

where  $T_e$  is the temperature of the electron of mass  $m_e$  and density  $n_e$ ,  $c$  is the speed of light and  $\sigma_T$  is the Thomson cross section. Finally,  $g(x)$  is given by

$$g(x) = \frac{x^4 e^x}{(e^x - 1)^2} \left[ x \coth\left(\frac{x}{2}\right) - 4 \right] \quad (1.40)$$

with  $x = h\nu/kT$ . Since the function  $g(x)$  gets null at  $\nu = 217$  GHz, the SZ effect can be detected via multi-frequency measurements.

Beside the thermal SZ effect, there is also a *kinetic* effect, due to the Doppler shift induced by

Thomson scattering of the CMB photons off an electron gas in motion with respect to the CMB. The effect is a red shift of the photons of the order of  $v_{bulk}/c$  (with  $v_{bulk} < 1000$  Km/s), together with spectral distortions with the same frequency scaling as the CMB fluctuations. Because of that, the two of them are in principle indistinguishable. However, since the effect arises from galaxy clusters with typical size of 1 arcminute, it is possible to disentangle them by applying a spatial filtering. Moreover, the kinetic effect is a factor 10 lower than the thermal one, therefore having a limited impact on the CMB power spectrum.

Given the  $z$  distribution of the galaxy clusters, it is possible to derive the power spectrum of the fluctuations in the CMB produced by SZ effect. For this purpose, it is convenient to use simulations where various properties of cluster have to be taken into account (for example the dispersion of cluster sizes and electron temperatures). On a cosmological scale, the evolution of the cluster as a function of the red-shift  $z$  must be considered as well. The power spectrum will depend on the cosmological model and it will have a peak at  $\ell \sim 2000 - 5000$ . Indeed, Holder et al. (2007) showed that the temperature of the decrement is independent of the red-shift, while the angular size decreases with it. Since there is a mean redshift of cluster formation, there is also a minimum angular size (maximum  $\ell$ ) for the SZ fluctuations.

The SZ power spectrum is compared with the CMB one and that of the other foreground components in Figure 1.15 and Figure 1.16: it is possible to notice that the SZ contribution is important relative to the CMB only for  $\ell \geq 1000$ , where the Silk damping damps down the CMB fluctuations. Moreover, it is also possible to see the minimum of this effect at 217 GHz: for lower frequencies the radio point sources are dominant, while at higher frequencies the dusty galaxies give a stronger signal than the SZ effect. This suggests that in order to have a good detection of the SZ effect, it is important to have observation at frequencies in the range of 40 and 150 GHz and in the  $\ell$  range of 1000 to 5000.

## 1.5 Foreground models

Many efforts have been employed to define realistic models of the foreground components. They can thus be used to simulate observations of typical CMB experiments and to test the techniques developed to analyse the real data.

Here, we briefly describe the models generally used as foreground templates at the *WMAP* frequencies, and that we also adopted in the analysis presented in Chapter 6. We refer to Chapter 2 for more details about the techniques used to derive them.

### 1.5.1 The Planck Sky Model

Developed for the Planck mission, the Planck Sky Model (PSM) provides simulations of the sky emissions in the frequency range of CMB experiments (Planck User Manual, 2008).

Specifically, the synchrotron emission is modeled as an extrapolation in frequency of the 408 MHz

map of Haslam et al. (1982). The derived map  $S_\nu(r)$  at frequency  $\nu$  is then given by:

$$S_\nu(r) = S_{0.408}(r) \left( \frac{\nu}{0.408} \right)^{\beta(r)} \quad (1.41)$$

where  $\beta(r)$  is the spectral index map. The different models implemented in the PSM differ by the  $\beta(r)$  map used: in the model proposed by Miville-Deschênes et al. (2008) and adopted for our analysis,  $\beta(r)$  has been deduced from the 23 GHz polarization data from *WMAP*.

Besides, the free-free template is a mix of two maps. The  $H\alpha$  map of Dickinson et al. (2003) is used in diffuse regions of the sky while the MEM map (Bennett et al., 2003) is used in regions where dust extinction is important. In all cases, the extrapolation in frequency is done following Dickinson et al. (2003) assuming an electron temperature  $T_e = 7000$  K. Furthermore, the free-free spectrum is assumed constant on the sky.

The model 7 of Finkbeiner et al. (1999) is adopted to trace the thermal dust emission. This uses a spectral extrapolation of the IRAS 100  $\mu\text{m}$  map with the spatial variation of the dust grain equilibrium temperature given by the 240/140  $\mu\text{m}$  DIRBE maps. The PSM allows to use a version of the IRAS 100  $\mu\text{m}$  from which HII regions were removed. These sources have a specific temperature which is not well estimated by the low resolution 240/140  $\mu\text{m}$  map. The HII regions are therefore treated separately.

Finally, the anomalous emission is included as spinning dust contribution and it is modeled by the spinning dust template proposed by Miville-Deschênes et al. (2008). The scaling in frequency is done using a constant spectrum on the sky given by the Warm Neutral Medium (WNM) emission model proposed by Draine & Lazarian (1998b). It is worth to note, however, that there are uncertainties in this model as well as in the models deduced for other environmental conditions. In fact, the latest models from Hoang et al. (2010) find shift factors of between 2 and 4 in peak frequency if additional physics is taken into account. Moreover, the WNM model does not have the power that seems to be needed at V-band (see fits in Chapters 3 and 4). Therefore, a power-law may well be the best model since the anomalous component along a line of sight is likely to be a sum over many models with different peak frequencies.

The polarization part of the signal is also included: the dust polarization is computed according to equation 15 of Page et al. (2007). To this, a bi-symmetrical spiral (BSS) Galactic magnetic field model is used, also allowing to consider the line-of-sight depolarization due to the turbulent part of the magnetic field (see Miville-Deschênes et al. (2008)). On the other side, synchrotron polarization is based on the 23 GHz polarization data of *WMAP*, extrapolated in frequency with the same spectral index map as the synchrotron intensity.

### 1.5.2 The foreground model based on MEM solutions

Despite their complicated noise properties, the *WMAP* science team has used the MEM solutions as foreground models at the frequency of observation of the satellite. They are derived as described in Bennett et al. (2003) and Gold et al. (2009).

The data are previously cleaned of the CMB contribution using the ILC map, despite issues over

whether residual foregrounds in the ILC bias the MEM foreground estimates (see Chapter 2 for a more detailed introduction to the ILC method and its limitations). They are subsequently fitted with templates of each of the foregrounds, to each pixel independently. For the foreground model, the *WMAP* team assumed the prior spatial distribution and spectral behavior of foreground components by using the Haslam map at 408 MHz (Haslam et al., 1982) for the synchrotron emission, the Model 8 of Finkbeiner et al. (1999) for the dust emission and the extinction corrected  $H\alpha$  map from Finkbeiner (2003) for the free-free, already described in previous sections. The synchrotron spectrum is allowed to vary for each pixel, while the free-free and dust spectral indices are assumed to be constant over the sky, with a value of  $\beta_{free-free} = -2.14$  and  $\beta_{dust} = 2$  (Gold et al., 2009).

### 1.5.3 The foreground model proposed by Ghosh *et al.*

Most recently, Ghosh et al. (2010) derived a set of maps of the foreground components<sup>20</sup> at the frequency of observation of *WMAP* and with the same angular resolution.

They achieved this result by subtracting an estimate of the CMB from the maps, and then post-processing the remaining set of CMB-cleaned *WMAP* data to obtain maps of the foreground emissions. Specifically, they extracted a CMB map using the ILC method implemented in the needlet space (Delabrouille et al., 2009): this kind of approach allows for taking into account the variability of the statistical properties of foregrounds and noise both in pixel and harmonic space. In particular, it permits better reconstruction of the CMB in the vicinity of the galactic plane, which is the region of major interest for foreground science.

After CMB subtraction, the estimated foreground map at a given frequency still contains significant noise, which dominates on small scales. The variance of the foreground map in a single channel (and at the resolution of the channel), is minimal if the Wiener-filtered version of the needlet ILC map is subtracted. Since Galactic foregrounds are strongly concentrated in the galactic plane, the filtering should depend on the location on the sky. Specifically, as a reasonable compromise between efficiency and simplicity, they approximated the foreground emission as following a plane parallel slab model. They then applied a filter for each zone of galactic latitude.

### 1.5.4 Comparison of foreground models

The MEM maps are shown in Figure 1.17, as example of how the models look like. They have an effective resolution of 1 degree which does not allow for detecting significant differences: the main features are very similar to all the three models.

A systematic comparison of these models with respect to the real *WMAP* data, gives an idea of the actual differences of the models and of the level of consistency of the former with respect to the observations.

It is particularly interesting to compare the PSM with the MEM and GFM maps, and to see the effect of the different processes used to generate them. While the MEM and the GFM maps have been

<sup>20</sup>[http://www.apc.univ-paris7.fr/APC/Recherche/Adamis/fg\\_wmap-en.php](http://www.apc.univ-paris7.fr/APC/Recherche/Adamis/fg_wmap-en.php)

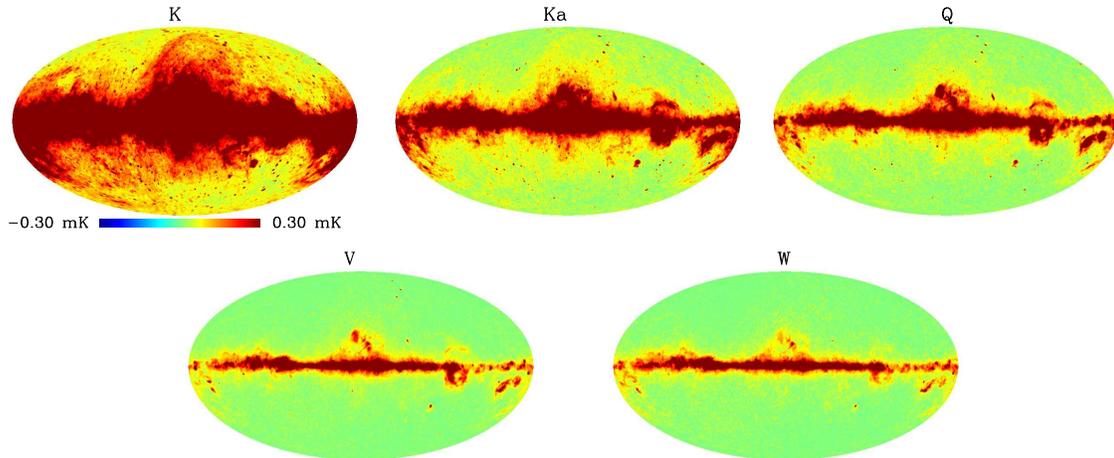


Figure 1.17: Plot of the maps of the integrated foreground components as derived using the MEM, at the five frequency of observation of *WMAP*. All the maps have an effective angular resolution of 1 degree.

derived directly from the *WMAP* data pre-cleaned of the CMB emission via ILC estimations, the PSM is the result of the combination of the best templates of the single components that are assumed to contribute to the foreground contamination. These templates are derived from an ancillary data set and are properly scaled at the frequency of interest, according to specific models. Therefore, although very similar to the data, both the MEM solutions and the GFM maps are affected by the bias due to the subtraction of the ILC map, which contains foreground residuals (see Chapters 2 and 6). Furthermore, the MEM maps have complex noise properties. All these problems should be overcome by the PSM, whose limitation however, is the fact of relying on other data than the *WMAP* ones.

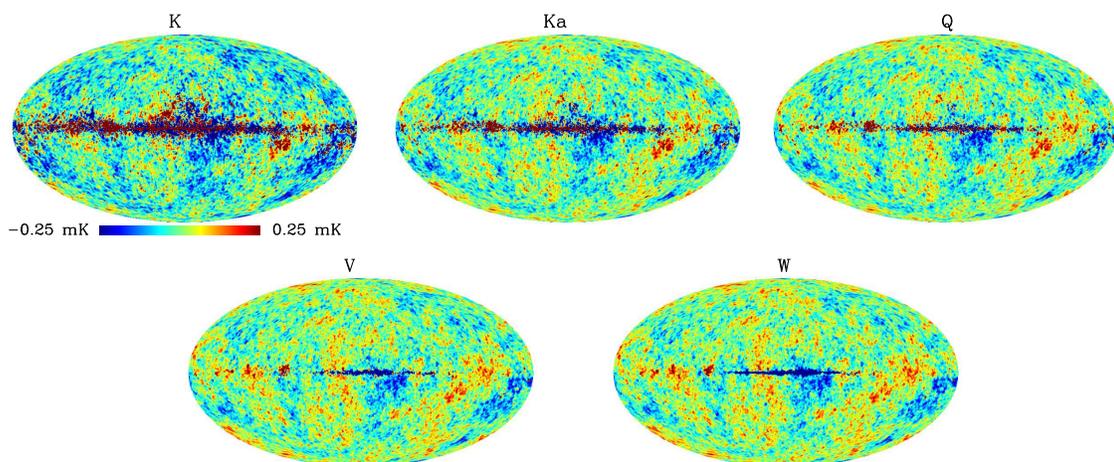


Figure 1.18: Difference between the *WMAP* data and the integrated foreground maps derived using the PSM.

Figures 1.18, 1.19 and 1.20 show the difference between the integrated foreground maps derived

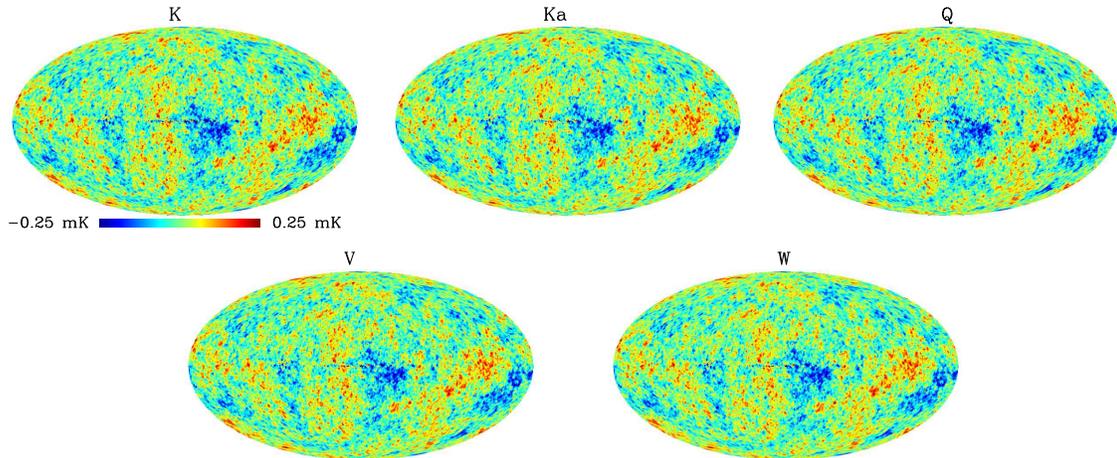


Figure 1.19: Difference between the *WMAP* data and the integrated foreground maps derived using the MEM maps.

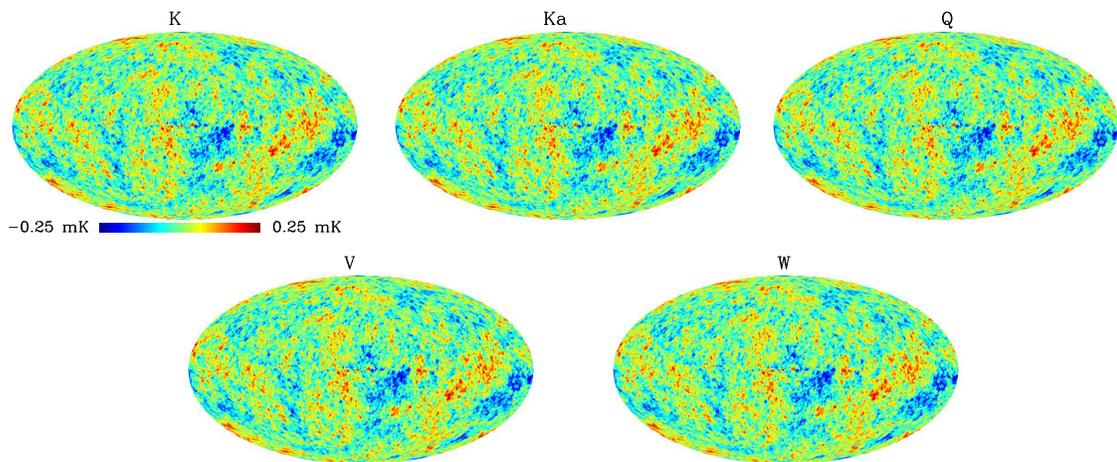


Figure 1.20: Difference between the *WMAP* data and the integrated foreground maps derived using the GFM maps.

using the three models and the *WMAP* data. In the ideal case, these maps should show a pure CMB emission. This is the case for the MEM and GFM models. However, as clearly visible in the plot, when the PSM is used, a quite strong contamination remains along the Galactic plane. Its amplitude varies depending on the band considered: therefore, this gives an estimation of the foreground emissions not traced by the model. As expected, the strongest signal comes out at the lowest frequencies, since this is the range where the uncertainty about the Galactic emissions is largest, and the cross-talk among the components is stronger. Therefore, although in principle the PSM provides a good estimation of the Galactic emissions, the other two models should be probably preferred as tracers of the foreground components for the *WMAP* data.

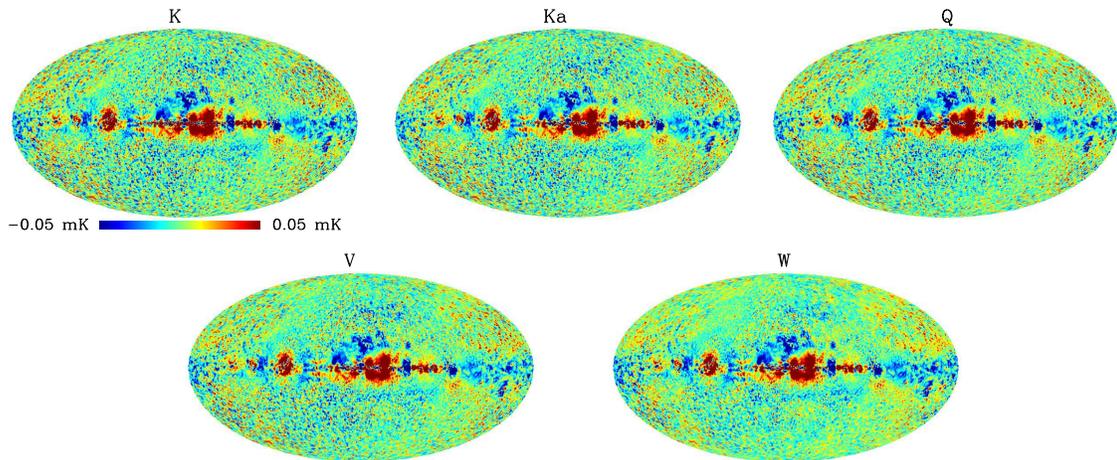


Figure 1.21: Difference between the integrated foreground maps derived using the MEM maps and the GFM.

The MEM and GFM maps are quite similar to each other, although some discrepancies come out looking at the difference maps shown in Figure 1.21. The main features appear along the Galactic plane, and it is interesting to note that the amplitude changes only very little with the frequency. Moreover, the latter is very small, making the difference practically non influential in the analysis.

## Chapter 2

# Methods for cleaning CMB data and the algorithm

### 2.1 Introduction

The foreground emissions strongly contaminate CMB observations, making it difficult to use the latter for cosmological studies. Several methods have been developed in order to clean CMB data, all of them taking advantage of multi-frequency observations and therefore, of the different properties of the emissions in the sky (see Chapter 1).

Two classes can be distinguished, depending on the target of the method itself. The first one consists of the so called foreground cleaning methods which simply attempt to estimate the CMB signal as clean as possible from any contamination, with no concern for the detailed separation of the foreground emission into physical components. The knowledge of foregrounds is not only interesting science on its own, but it is also the key for better cleaning CMB data. Therefore, a different approach to the problem of cleaning CMB data requires to quantify the actual contamination due to foreground components and to subtract it from the data. The focus is now constituted by the analysis and separation of the different components: any new information about the foregrounds can be used to fine-tune the technique used to disentangle the components, consequently improving the cleaning of the data. This idea has driven the development of the so called component separation methods, both for diffuse and discrete sources. Component separation techniques decompose multi-frequency input data into separate physical emission components, one of which, in this context, is the CMB.

Note, however, that, as shown in Chapter 3, component separation methods can have limited success, providing an adequate reconstruction of the CMB, but not of the foreground emission. Based on the idea that the observed signal is the results of the superposition of independent sources, the efficiency of the method deteriorates in the regions where foregrounds are strongly correlated and mixed. In these regions, the complexity of the spectral properties and morphology of the foreground components make it difficult to disentangle the different sources. In these cases, nevertheless, component separation methods effectively constitute foreground cleaning methods (see Chapters 3 and 4).

## 2.2 Cross-correlations method

The simplest way to produce a cleaned CMB map useful for cosmological analysis (by maintaining well specified noise properties) is to simultaneously fit a set of externally derived template maps to the CMB observations. The cross-correlation analysis is realised by means of a least-squares fit of a CMB observation to one or more templates. A cleaned CMB map is therefore derived by cleaning the data of the foreground contamination according to the evaluated correlation. Due to the use of predefined models of the foreground components, this is a non-blind approach for cleaning CMB data. When several components are correlated simultaneously, the problem is described by a matrix equation and the error can be determined from the diagonals of the matrices. It is also optimal to include information about the CMB emission (flat spectrum) and the other components (by fixing the spectrum behavior). This is particularly useful in order to limit the errors introduced by the cross-correlation between the CMB and the foreground templates: such correlation indeed, biases the evaluation of the spectrum of the foreground components and, consequently, introduces residuals in the yielded CMB emission. Conversely, fixing the spectra of the foreground components has the disadvantage of losing information about the actual spectral behaviour, which is the main target of the analysis itself.

This method has been used also by the *WMAP* science team, and then slightly modified with the arrival of the new data release. Bennett et al. (2003) proposed a simple correlation analysis of all the five frequencies maps with three templates, namely the Haslam map, the FDS model 8 at 94 GHz (developed by Finkbeiner et al. (1999)) and the  $H\alpha$  map for the synchrotron, dust and free-free emission respectively. However, they also experimented with a correlation analysis which involved only the Q, V and W bands, and where the synchrotron and free-free spectrum were fixed to follow a power law with index equal to -2.7 and -2.15 respectively. Furthermore, since the three years data were published, the *WMAP* analysis has been modified by using a different synchrotron template. Indeed they tried to improve the model which describes the low-frequency emissions considering the difference map (in thermodynamic temperature) between the K and Ka band maps: the CMB contribution is canceled out, leaving only foregrounds and noise. The resultant map is certainly a better description of the low-frequency contaminant as observed by *WMAP*, taking into account spatial spectral variations. However the interpretation of the measured correlation becomes complicated due to the mixing between all the components present at these frequencies. Although the synchrotron emission is the dominant component, a certain amount of free-free emission is also present, as well as the anomalous component. As before, they introduced some constraints relatively to the spectral behaviour of the foreground emissions: specifically, they imposed that all the coefficient were positive-defined, and that the dust and free-free coefficients followed a power law spectrum with spectral index equal to 2 and -2.14 respectively.

A cross-correlation analysis has been used by many other authors. Of particular interest for this work (see Chapter 5) is the analysis proposed by Finkbeiner (2004a) and Dobler & Finkbeiner (2008b). The main significant modification is the fact that they first subtracted a CMB estimation from the data, according to different ILC solutions of the *WMAP* data. They claimed it was necessary in order to avoid the cross-correlation between the CMB and foreground emissions, although this introduces a

major problem due to presence of foreground residuals in the ILC map: practically, subtracting this map to the data implies to add some spurious contamination (see next section for more details about ILC).

Finally, another interesting work in that direction is the one proposed by Davies et al. (2006) who performed a cross-correlation analysis of the *WMAP* first-years data restricted on patches of the sky, which have been selected so that only one component could be dominant over the others.

### 2.3 Internal Linear Combination (ILC) method

The ILC method falls in the category of those methods aimed at obtaining a map of the CMB free from foreground contamination, without studying the nature and properties of the single foreground components. Proposed by Tegmark & Efstathiou (1996), it has been used in a simplified way by the *WMAP* science team as a quick *blind* method to derive a clean map of CMB emission.

The approach to the problem is indeed simple. Let us assume to have  $k$  observed CMB maps at different frequencies, but with identical beams. Each of them can be written in thermodynamic temperature as:

$$T(\nu_k) = T_{\text{CMB}} + T_{\text{residual}}(\nu_k), \quad (2.1)$$

where  $T_{\text{CMB}}$  and  $T_{\text{residual}}(\nu_k)$  are statistically independent. It is possible to use them to generate a CMB map as cleaned as possible from any contribution but the cosmological emission, by simply linearly combining them as follow:

$$T = \sum_{i=1}^k w_i T(\nu_i). \quad (2.2)$$

Therefore, it is not necessary to have any information about the foreground components, since anything we need to know is already contained in the observations.

The frequency independence of the CMB spectrum requires that:

$$\sum_{i=1}^k w_i = 1, \quad (2.3)$$

Therefore, the resulting map may be written as

$$T = T_{\text{CMB}} + \sum_{i=1}^k w_i T_{\text{residual}}(\nu_i). \quad (2.4)$$

and the  $k - 1$  free weights can be chosen to minimize the contribution from the residuals. This means minimizing the variance of  $T$ , assuming that the CMB component is statistically independent of the foregrounds and the noise: indeed, in these conditions, the variance of the error is minimum when the

variance of the ILC map itself is minimum:

$$\text{Var}(T) = \text{Var}(T_{\text{CMB}}) + \text{Var}\left(\sum_{i=1}^k w_i T_{\text{residual}}(v_i)\right). \quad (2.5)$$

Following the prescription given by Eriksen et al. (2004), the variance minimization can be achieved by means of Lagrange multipliers. Indeed, it turns out that the solution is obtained by solving the following linear system of equations:

$$\begin{bmatrix} 2\mathbf{C} & -\mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}, \quad (2.6)$$

where  $\lambda$  is an arbitrary constant, and  $\mathbf{w} = (w_1, \dots, w_k)^T$  are the ILC weights. Their definition is given by a straightforward linear algebra:

$$w_i = \frac{\sum_{j=1}^k C_{ij}^{-1}}{\sum_{jk} C_{jk}^{-1}}. \quad (2.7)$$

Finally,

$$C_{ij} \equiv \langle \Delta T_i \Delta T_j \rangle = \frac{1}{N_{\text{pix}}} \sum_{p=1}^{N_{\text{pix}}} (T^i(p) - \bar{T}^i)(T^j(p) - \bar{T}^j) \quad (2.8)$$

is the map-to-map covariance matrix. This is the procedure which has been adopted by the *WMAP* science team since the data of three years of observations have been released.

Away from the galactic plane and on small scales, the best linear combination for cleaning the CMB of foregrounds and noise may be very different from what it is close to the galactic plane and on large scales. In fact, the foreground properties vary strongly over the sky as a result of spatially dependent spectral indexes. A very natural idea to improve the ILC is to decompose the sky maps in several regions. They can be independently used to derive ILC maps which can consequently be added together to obtain a full-sky final map.

It is important to note that the ILC method does not work properly (biasing the result) if some correlation actually affects the components of interest. In particular, small data sets are always empirically correlated at some level, although they are realisations of uncorrelated random processes. This is why it is inadvisable to apply the ILC method on too small subsets of the original data (very small regions). Besides that, however, there is a level of correlation which can not be overcome. It has been generally referred to as *Cosmic Covariance* (Chiang et al., 2009) and it is due to the fact that the foregrounds and the CMB are correlated amongst themselves. Over an ensemble of universes the true CMB and foreground are indeed expected to be uncorrelated, but any particular sky pattern (such as the one we happen to observe) will generate nonzero measured correlations simply by chance.

Different partitions of the sky have been proposed in the literature as well as different implementations of the ILC method itself. Some of them have been considered as a reference for the results of the work described later on in Chapter 4.

The *WMAP* science team (Bennett et al., 2003) derived an ILC map dividing the sky into 12 regions:

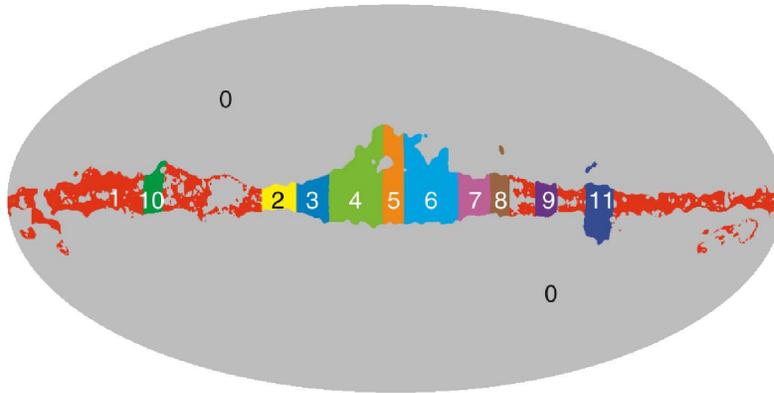


Figure 2.1: Full-sky map color-coded to show the 12 regions that were used to generate the five-year ILC map.

as shown in Figure 2.1, 11 regions lie in the inner Galactic plane, within the Kp2 cut, while the rest of the sky is treated as a single area. Thus, a full sky map is obtained by co-adding the maps derived with the individual regions: in order to suppress boundary effects, they also smoothed the edges of the regions with an effective Gaussian beam of amplitude  $1.5^\circ$ . Although the ILC method has been adopted by the *WMAP* science team as a possible solution for cleaning the data, they have also recognized the limitation of the results in terms of a possible usage for cosmological purposes. It is essentially due to the not perfect cleaning of foregrounds, which leaves a significant residual (a *bias*), mostly concentrated along the Galactic plane, but also quite significant at higher latitudes. The *WMAP* science team itself suggested to use this maps only up to a multipole  $\ell = 32$ , unless a mask is applied to exclude the region of the Plane. Moreover, the level and properties of the noise, as well as the limited angular resolution are other significant limits which have to be taken into account. They have estimated the bias due to the CMB-foreground correlation by Monte-Carlo simulations, and subtracted it from the composite map. Even though a bias correction has been done, the resultant map has still some drawbacks: it is assumed that the foreground emissions have constant properties in the regions, which is probably not realistic. Moreover, it is difficult to access the impact on the final map of the arbitrariness of the choice of the regions driven by *a priori* knowledge of foreground emissions. Finally, it is assumed that the foreground model used to determine the bias is accurate, but there might be a mismatch with the real data which is difficult to quantify.

The work proposed by Park et al. (2007) goes in the direction of improving all these aspects: in fact, a new set of regions is defined as the result of a more detailed study of the spatial variations of the spectral properties of the signals. 400 regions are defined from  $20 \times 20$  spectral index bins: 20 of them describe the spectral properties of the sky between the K and V band, and 20 the spectral index between V and W. Although the work is certainly a good attempt of improving the ILC performance, the method proposed is weakened by the fact that the spectral variations are defined on the basis of the MEM maps derived by *WMAP* (Bennett et al., 2003), which is not the optimal estimation of the foreground contaminants of the CMB emission. Moreover, the MEM solution uses the result of the *WMAP* ILC as a prior, which is subtracted from the data before using the MEM method to separate galactic foregrounds, hence automatically including all the drawbacks of the ILC map.

Another attempt in terms of taking into account spectral and spatial variations of the foregrounds, are the methods suggested by Kim et al. (2008) and Delabrouille & Cardoso (2007). In the first case the ILC method is implemented deriving pixel dependent weights, which are computed as maps via harmonic decomposition. Moreover, the method has been improved with an estimation of the bias: an iterative foreground reduction method has been developed, where perturbative corrections are made for the cross term which appears not to be negligible. In the second case, Delabrouille & Cardoso (2007) combined the ILC method with a Wiener filtering, however everything is now implemented on the frame of spherical wavelets called needlets. Needlets were introduced by Narkowich et al. (2006) as a particular construction of a wavelet frame on the sphere. They have been studied in a statistical context (e.g. Baldi et al. (2009, 2007)). The most distinctive property of the needlets is the fact that they can be simultaneously localised in the spherical harmonic domain and in the spatial domain (as is generally the case with wavelets). In fact, they are quasi-exponentially (i.e., faster than any polynomial) concentrated in pixel space and exactly localized on a finite number of multipoles.

## 2.4 Component separation methods

In this section, we will give a description of the ICA approach to the component separation problem. It is particularly interesting, since the algorithm used for large part of this work is actually one of the several implementations of this kind of approach. Nonetheless, it is important to note that several methods have been recently developed and used by the CMB community, which will not be presented here. A useful review is presented by Delabrouille & Cardoso (2007) and Leach et al. (2008).

A general formalism can be introduced. We assume that the sky radiation at a given frequency  $\nu$  can be described as a linear combination of  $N$  different sources, produced by different physical processes, whose frequency and spatial dependencies can be factorized into two terms:

$$\tilde{x}(\mathbf{r}, \nu) = \sum_{j=1}^N \bar{s}_j(\mathbf{r}) f_j(\nu). \quad (2.9)$$

Generally the radiation is observed by a combination of an optical system and a  $M$ -channel measuring instrument. Therefore, a further assumption is that the observations can then be described as the convolution of the sky signal with a frequency dependent azimuthally symmetric beam response and a bandpass  $t_\nu(\nu')$  (the frequency response of the channel) which is effectively a delta function.

If we introduce the instrumental noise  $\epsilon$ , and we assume that the beam function is frequency independent  $B(\mathbf{r}, \nu) = B(\mathbf{r})$ , the observed signal for each position  $r$  is expressed by

$$\mathbf{x}(\mathbf{r}) = \mathbf{A} \bar{\mathbf{s}}(\mathbf{r}) * B(\mathbf{r}) + \epsilon(\mathbf{r}) = \mathbf{A} \mathbf{s}(\mathbf{r}) + \epsilon(\mathbf{r}), \quad (2.10)$$

where  $\mathbf{x}$  and  $\epsilon$  are vectors with  $M$  rows, and the star represents the convolution of the beam point spread function (PSF) with the sky signals  $\bar{\mathbf{s}}$ , indicated simply as  $\mathbf{s}$  afterward. Finally,  $\mathbf{A}$  is the so called mixing matrix, whose elements define the scaling coefficients  $a_{\nu,j} = f_j(\nu)$ . It is a  $M \times N$  matrix, which

does not depend on the pixel for a single linear mixture.

The problem of component separation consists of inverting the system of equations 2.10. However, this is not trivial, since it is not a simple linear inversion, due to the instrumental noise term. Therefore, statistical techniques are necessary. Some of them, make use of *a priori* knowledge of the signals, so that the inversion becomes linear, others, instead, are *blind* approaches.

### 2.4.1 The Independent Component Analysis

The *Independent Component Analysis* (ICA) is a blind approach to the component separation problem, where nothing is assumed about the spectral properties of the components to be separated. The starting point of ICA is the very simple assumption that the component  $s_i$  are statistically *independent* and that all of them, but at most one, have *non-Gaussian* distribution. The last assumption, specifically, fits well the astrophysical problem of separating the CMB emission from the foreground components, since the latter have non-Gaussian distributions, unlike the CMB emission.

The component separation is achieved using the Central Limit Theorem. It states that the distribution of a sum of independent random variables tends toward a Gaussian distribution. Therefore, assuming that the components of the vector  $\mathbf{s}$  in Equation 2.10 are independent as required by the ICA approach, the solution to the component separation problem is obtained finding a linear transformation  $\mathbf{W}$  of the vector  $\mathbf{x}$  such that the transformed vector  $\mathbf{y} = \mathbf{W} \mathbf{x}$  has independent components. Indeed, considering a single component of  $\mathbf{y}$ , we have:

$$\mathbf{y} = \mathbf{w}^T \mathbf{x} = \mathbf{w}^T \mathbf{A} \mathbf{s} \quad (2.11)$$

where  $\mathbf{w}^T$  is a row of the matrix  $\mathbf{W}$ , we have neglected the noise contribution and  $\mathbf{y}$  turns out to be defined as a linear combination of the independent sources  $\mathbf{s}$ . Following the central limit theorem, it means that the distribution is closer to a Gaussian one than the single components. Conversely, it will become least Gaussian when it in fact equals one of the components  $\mathbf{s}$ . Therefore, the main point is to find  $\mathbf{W}$  so that the non-Gaussianity of the variables  $\mathbf{y}$  is maximised. Of course, it has to take into account the instrumental noise.

A fundamental result of information theory is that a Gaussian variable has the largest entropy among all random variables of equal variance. The concept of entropy (Shannon & Weaver, 1949) defines the amount of information we have in the probability distribution of a variable. For a discrete random variable  $X$  on a finite set of possible values  $x_i$  with probability distribution function  $p(x_i) = p(X = x_i)$ , the entropy is defined as :

$$H(X) = - \sum_{i=1}^N p(x_i) \log(p(x_i)). \quad (2.12)$$

It easily applies to images which are constituted by  $N$  pixels, such as observed maps. The entropy is then interpreted as the numbers of independent ways of obtaining the image, given the single pixels. In the continuous case where the variable  $X$  can assume any real value  $x$  with probability density  $p(x)$ , the

*differential entropy* is introduced, which is defined as follows:

$$H(X) = - \int_{-\infty}^{\infty} p(x) \log p(x) dx. \quad (2.13)$$

Maximising the entropy of a variable means to chose a probability distribution such that the amount of information we have, relative to the variable, is the least informative possible. Therefore, to obtain a measure of non-Gaussianity, a slightly modified version of the definition of the entropy (Equation 2.13) is used, called *neg-entropy*:

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y}), \quad (2.14)$$

where  $\mathbf{y}_{gauss}$  is a Gaussian random variable of the same covariance matrix as  $\mathbf{y}$ . It is always non-negative and it is zero if and only if  $\mathbf{y}$  has a Gaussian distribution.

The advantage of using the neg-entropy for evaluating the non-Gaussianity of a variable is the fact that it is well justified by statistical theory. However, it is computationally very difficult since it would require an estimate of the probability density function (pdf). Thus, simpler approximations of the neg-entropy are very useful. One of those has been proposed by Hyvarinen (1999) and Hyvarinen & Oja (2000): if the noise is assumed to be additive, signal-independent, white, Gaussian, and stationary, and its covariance matrix is known, the Gaussian moments of  $\mathbf{y}$  are shown to be robust estimates of the desired function. Therefore, a general approximation is:

$$J(\mathbf{y}) \propto [EG(\mathbf{y}) - EG(\mathbf{y}_{gauss})]^2, \quad (2.15)$$

where  $E$  is the expectation value,  $\mathbf{y}_{gauss}$  is a Gaussian variable of zero mean and unit variance, and  $G$  is a non-quadratic function. Three good choices for  $G$  are:

$$p(u) = u^3, \quad g(u) = u \exp(-u^2), \quad t(u) = \tanh(u). \quad (2.16)$$

In general, other estimators of non-Gaussianity can be used. In fact, depending on the involved statistic, different implementations of the ICA approach have been proposed, such as JADE (Cardoso & Souloumiac, 1993) and SMICA (Spectral Matching Independent Component Analysis, introduced by Delabrouille et al. (2003)).

## 2.4.2 The pre-processing phase

A specific implementation of the ICA approach described in the preceding section, is given by the algorithm. It is an algorithm based on a fixed-point iteration scheme where the non-Gaussianity of  $\mathbf{w}^T \mathbf{x}$ , expressed by the approximate form shown in Equation 2.15, is maximised.

As any other ICA algorithm, it requires some useful pre-processing of the data, in order to reduce the complexity of the computations.

### Centering

The most basic pre-processing is to *center* the data, which means to subtract the mean  $\mathbf{m} = E\mathbf{x}$  to the data  $\mathbf{x}$  in order to have zero mean variables. This implies that  $\mathbf{s}$  has the same property as well. The offset of each independent component can be recovered in the end of the separation process.

### Whitening

Another useful pre-processing strategy is to whiten the observed variables. This correspond to linearly transforming the data  $\mathbf{x}$  to a new vector  $\hat{\mathbf{x}}$ , which is white, meaning that its components are uncorrelated and their variance is equal to unity. In other words, it means that the covariance matrix of the new vector is the identity matrix:

$$E\{\hat{\mathbf{x}}\hat{\mathbf{x}}^T\} = \mathbf{I} \quad (2.17)$$

The whitening of a variable is always possible, by means of the eigenvalue decomposition of the covariance matrix  $E\{\mathbf{x}\mathbf{x}^T\}$ .

If the covariance matrix  $\Sigma$  of the system noise is known and the covariance matrix  $\mathbf{C}$  of the zero mean data is calculated by evaluating  $\mathbf{C} = E\{\mathbf{x}\mathbf{x}^T\}$ , the whitening procedure consists in evaluating a modified noise covariance matrix as follows:

$$\hat{\Sigma} = (\mathbf{C} - \Sigma)^{-1/2} \Sigma (\mathbf{C} - \Sigma)^{-1/2} . \quad (2.18)$$

Therefore, the whitened data  $\hat{\mathbf{x}}$  are defined as:

$$\hat{\mathbf{x}} = (\mathbf{C} - \Sigma)^{-1/2} \mathbf{x} . \quad (2.19)$$

The utility of whitening resides in the fact that the new mixing matrix  $\hat{\mathbf{A}} = (\mathbf{C} - \Sigma)^{-1/2}$  is orthogonal. Therefore, the number of parameters to be estimated is reduced: instead of estimating the number of rows of the original matrix  $\mathbf{A}$  ( $n^2$  parameters), only the new matrix  $\hat{\mathbf{A}}$  has to be computed, meaning  $n(n-1)/2$  parameters.

### The algorithm

After the pre-processing phase, the separation algorithm estimates the matrix  $\mathbf{W}$  row by row. Thus the single components  $\mathbf{y} = \mathbf{w}^T \hat{\mathbf{x}}$  ( $\mathbf{w}^T$  is a row of matrix  $\mathbf{W}$ ) are recovered.

As already stated is a noise-robust iterative algorithm where  $\mathbf{w}$  is found reaching a convergence criterion following a fixed-point scheme:

1. Choose an initial vector  $\mathbf{w}$ ;
2. update it through

$$\mathbf{w}_{new} = E\{\hat{\mathbf{x}}g(\mathbf{w}^T \hat{\mathbf{x}})\} - (I + \hat{\Sigma})\mathbf{w}E\{g'(\mathbf{w}^T \hat{\mathbf{x}})\}$$

where again  $E$  denotes expectation over all the available samples,  $g$  is one of the regular non-quadratic functions defined in Equation 2.16, and  $g'$  is its first derivative;

3. let

$$\mathbf{w}_{new} = \frac{\mathbf{w}_{new}}{\|\mathbf{w}_{new}\|} .$$

4. Compare  $\mathbf{w}_{new}$  with the old one; if it did not converge (i.e., the difference is larger than a fixed value  $\epsilon$ ), go back to (ii), if it converged, begin another process.

This procedure maximizes the non-Gaussianity of the component  $\mathbf{w}^T \hat{\mathbf{x}}$ . However, since the components have to be independent, therefore at least uncorrelated, an orthogonalization step (e.g. Gram-Schmidt) must be inserted between steps (ii) and (iii). It implies that if a certain number,  $k$ , of rows of  $\mathbf{W}$  have been evaluated, to estimate the  $(k + 1)$ -th row, it has to be searched in the subspace orthogonal to the first  $k$  rows.

Given the separation matrix  $\mathbf{W}$  and the estimation for the independent components in the data, it is possible to derive the following definition of the data (by simply inverting equation  $\mathbf{y} = \mathbf{W}\mathbf{x}$ ):

$$x_{\nu}^{(j)} \text{ (physical units at frequency } \nu) = (W^{-1})_{\nu j} y_j \quad (2.20)$$

which specifies the contribution of the independent sources for each observation at a given frequency  $\nu$ . Then, the frequency scaling for the component  $j$  between frequencies  $\nu$  and  $\nu'$  is given by

$$\frac{x_{\nu j}}{x_{\nu' j}} = \frac{(W^{-1})_{\nu j}}{(W^{-1})_{\nu' j}} . \quad (2.21)$$

Furthermore, it is also possible to have an estimation of the noise relative to the CMB reconstructed map. Since the system noise covariance matrix  $\Sigma$  is assumed to be known, it is possible to build noise constrained realizations  $\mathbf{n}_x$  for each frequency channel. Once the matrix  $\mathbf{W}$  has been recovered, the corresponding noise realizations in the outputs  $\mathbf{n}_y$  are

$$\mathbf{n}_y = \mathbf{W}\mathbf{n}_x . \quad (2.22)$$

### 2.4.3 for CMB data analysis: CMB cleaning and foregrounds analysis

The ICA approach for the component separation problem is particularly interesting since it allows a *blind* reconstruction of the signals. Given this feature, it is sensible to use the algorithm in the astrophysical context where the CMB emission needs to be separated from the foreground radiations, whose properties, however, are poorly known. It is even more advantageous, if the CMB data are measured in polarization, since very little is known about the polarized foreground components. Moreover, the algorithm properly treats the noise and allows a good estimate of it, unlike other methods: although the weights are derived not linearly, the maps are simply linear combinations of the input data.

The code has been used for different kinds of analysis of CMB observations. Indeed, it is suitable for both the approaches mentioned in the introduction to this chapter. The work of this thesis is actually

the result of several applications of `ica_fm`, as a tool for analysing CMB data, aiming both to clean the cosmological signal of any contamination and to enlarge our knowledge of the foregrounds.

An internal analysis of multi-frequency CMB data is generally used to extract a clean map of the cosmological signal, regardless of the nature and properties of the contaminants. This approach makes `ica_fm` very similar to the ILC method. Given  $N$  input maps, derived from the same set of observations at different frequencies, the code returns  $N$  signals, defined as a linear combination of the input data: they are the best estimation of the components to be separated. Therefore, among all of them, one is the best estimation of the CMB itself.

`ica_fm` has been utilised for this kind of application for both the *COBE*-DMR and *WMAP* multi-frequency observations of the microwave sky in Maino et al. (2002) and Maino et al. (2007) respectively.

A great limitation of this approach, however, is the fact that the weights which define the best estimation of the sources to be separated, are unique for the entire sky. In other words, the solution does not take into account spatial variation of the foreground emissions. This is why, in this thesis, we have also investigated a regional analysis of CMB observations with the code. In a way similar to the one proposed for ILC, we experimented with an internal analysis of the *WMAP* data, in different patches of the sky separately. This is the work presented in Chapter 6, where we perform a comparison between the results of the ICA approach and those derived with ILC. The analysis proposed is also interesting in order to define the partition of the sky which realistically reflects the properties of the sky and, therefore, improves the estimation of the CMB.

In general, it has been found that when used for internal analysis, the foreground components returned by the method are not optimal – a fully reliable component separation is not achieved. Nevertheless, one of the output maps of the analysis can, in general, be associated explicitly and robustly with the CMB signal. In an alternative approach, Maino et al. (2002) proposed the analysis of a single microwave sky map in combination with observations at wavelengths where only single foreground physical emission processes dominate. These various foreground templates are considered to be representative models for the morphology of the foreground emission at that wavelength, except for a scale-factor to be determined. Application of the `ica_fm` algorithm then provides information on foreground emission, not directly from the spectral properties of the reconstructed and ambiguous Galactic components, but from the recovered CMB component. Referring to Equation 2.11, the reconstructed CMB component is defined as a weighted linear combination of the input data, namely the microwave sky map and putative templates of Galactic emission. Those coefficients associated with the foreground templates define the effective contamination of the CMB anisotropy signal by each foreground. After normalising the weights so that the factor associated with the microwave sky map is unity, we can then interpret the modified template weights as correlation coefficients between the foregrounds and a given channel of the microwave data. Therefore with `ica_fm`-derived CMB sky maps at different frequencies, we can derive the frequency dependence of the coupling coefficients associated with a given template or physical foreground component. This is the approach proposed also in Chapters 3 and 4, for the analysis of the *WMAP* three- and five-year data respectively.

In addition, in Chapter 4, we propose an alternative application of `ica_fm` where the *WMAP*

observations are analysed simultaneously with the foreground templates (see Section 4.8). Finally, we experiment with a combination of the foregrounds and internal analysis of the CMB observations: the first one is used to clean the data of the foreground contamination, which are then internally analysed to derive a CMB map deprived of any residuals left in the first step of the analysis (see Section 4.6.)

#### 2.4.4 : CMB prior and loss-less input data compressing

Several improvements to the code are possible. Two of them have been already implemented in a new version of the algorithm, called .

The first innovation introduced is the implementation of the component separation where the CMB is sought under the constraint of following a Planckian spectrum. A detailed description of the way this constraint can be introduced in the computation, is given by Vio & Andreani (2008). One of the main results of this work is the fact that the constrained CMB derivation does not differ from the one implemented by the ILC code.

A second modification is a generalization of the pre-whitening procedure, which properly treats the noisy and redundant data sets, preserving all the available information. The standard data compression and Principle Component Analysis techniques are combined in order to find a linear combination of the input maps which preserves all the information concerning the sky signal, but which consists of the independent data.

As before, for each sky pixel,  $p$ , we model the data as,

$$\mathbf{x}_p = \mathbf{s}_p + \mathbf{n}_p \quad (2.23)$$

where the vectors  $\mathbf{x}_p$ ,  $\mathbf{s}_p$  and  $\mathbf{n}_p$  are the measured maps, sky signals, and noise, respectively. Note however, that now each sky signal map,  $\mathbf{s}_p$ , denotes a total, combined sky signal, which is presumed to be a linear combination of the physical components,  $\mathbf{c}_p$ . Therefore, for each sky pixel, we can write,

$$\mathbf{s}_p = \mathbf{A}\mathbf{c}_p, \quad (2.24)$$

where the matrix  $\mathbf{A}$  is the mixing matrix already defined, and  $\mathbf{c}_p$  denotes the amplitudes of the  $n_c$  sky components we want to estimate (previously called  $\mathbf{s}$ ).

The eigenvalue decomposition is used to define the null space of the signal correlation matrix,  $\mathbf{C} = \langle \mathbf{s}\mathbf{s}' \rangle$ :

$$\mathbf{C} = \mathbf{U} \mathit{diag}(\epsilon_i) \mathbf{U}^t \quad (2.25)$$

with  $\mathbf{u}_i$  being the eigenvectors of  $\mathbf{C}$ , and  $\mathit{diag}(\epsilon_i)$  the diagonal matrix of the corresponding eigenvalues. The null space is spanned by a subset of the  $\mathbf{u}_i$  vectors, which we will call  $\mathbf{V}$ . Denoting with  $\mathbf{B}$  the remaining vectors, we have:

$$\mathbf{U}^t = [\mathbf{B}'\mathbf{V}']. \quad (2.26)$$

Any signal maps  $\mathbf{s}$ , can be found only up to an unknown arbitrary vector from the null space. Therefore, to make the solution unique, it is convenient to impose a set of additional constraints, such

as that the solutions have no component in the null space, i.e.,

$$\mathbf{V}\mathbf{s} = 0. \quad (2.27)$$

It implies that:

$$\mathbf{s} = \tilde{\mathbf{B}}^t \mathbf{y} \quad (2.28)$$

where  $\mathbf{y}$  is a vector of the length,  $n_c$ , which defines the number of the independent components.  $\tilde{\mathbf{B}} = \text{diag}(\epsilon_i^{-1/2})\mathbf{B}$ , where only the non-zero  $\epsilon$  values are included. Plugging that into Equation 2.24 we obtain,

$$\mathbf{x} = \tilde{\mathbf{B}}^t \mathbf{y} + \mathbf{n}. \quad (2.29)$$

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

$$\tilde{\mathbf{y}} = \left( \tilde{\mathbf{B}}\mathbf{N}^{-1}\tilde{\mathbf{B}}^t \right)^{-1} \tilde{\mathbf{B}}\mathbf{N}^{-1}\mathbf{m} \quad (2.30)$$

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

$$\mathbf{N}_{\tilde{\mathbf{y}}} = \left( \tilde{\mathbf{B}}\mathbf{N}^{-1}\tilde{\mathbf{B}}^t \right)^{-1}. \quad (2.31)$$

Note, that the correlations of the noise between different maps already on the input are also introduced in the standard algorithm due to the pre-whitening/decorrelation procedure and, therefore, are not a new feature of the new approach proposed here. The signal correlation matrix for the new data,  $\tilde{\mathbf{y}}$  is given by,

$$\mathbf{C}_{\tilde{\mathbf{y}}} = \langle \tilde{\mathbf{y}}\tilde{\mathbf{y}}^t \rangle = \left( \tilde{\mathbf{B}}\mathbf{N}^{-1}\tilde{\mathbf{B}}^t \right)^{-1} \tilde{\mathbf{B}}\mathbf{N}^{-1} \langle \mathbf{m}\mathbf{m}^t \rangle \times \mathbf{N}^{-1}\tilde{\mathbf{B}}^t \left( \tilde{\mathbf{B}}\mathbf{N}^{-1}\tilde{\mathbf{B}}^t \right)^{-1} = \mathbf{1}. \quad (2.32)$$

Therefore, thanks to the rescaling applied in Equation 2.28, the new data set does not require any pre-whitening stage.

The normalization and frequency scaling are determined in a way analogous to the standard . However, it is necessary to take into account that the input maps for the ICA procedure will, in general, not correspond to any particular frequency, although a linear combination of them will give the estimation of the yielded components, as defined in Equation 2.30, i.e.,

$$\tilde{\mathbf{y}} = \mathbf{B}\mathbf{c}, \quad (2.33)$$

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

Combining Equations 2.24 and 2.28, we obtain,

$$\tilde{\mathbf{y}} = \mathbf{y} + \left( \tilde{\mathbf{B}}\mathbf{N}^{-1}\tilde{\mathbf{B}}^t \right)^{-1} \tilde{\mathbf{B}}\mathbf{N}^{-1}\mathbf{n}. \quad (2.34)$$

It means that:

$$\mathbf{s} = \tilde{\mathbf{B}}' \tilde{\mathbf{y}} = \tilde{\mathbf{B}}' \mathbf{W} \mathbf{c}, \quad (2.35)$$

and hence

$$\mathbf{A} = \tilde{\mathbf{B}}' \mathbf{W}. \quad (2.36)$$

The frequency scaling of the recovered components is now given by the generalized expression for the mixing matrix,  $\mathbf{A}$ , (Equation 2.36). This reduces to the standard form (see Maino et al. (2002)) whenever  $\tilde{\mathbf{B}}$  is a square unit matrix.

This modified version of            has been used in the analysis proposed in Chapter 6. Specifically, we took advantage of the new features of the algorithm to derived a map which contains all the foreground components in the *WMAP* data. This map has then been used to define a partition of the sky, which could reflect real variations of the properties of the sky.

## Chapter 3

# Foreground analysis of the *WMAP* three-year data with

*Note: the work presented in this chapter was done in collaboration with A.J. Banday and D.Maino and was published as Bottino et al. (2008). All the computations have been performed by M. Bottino.*

### 3.1 Introduction

In this chapter we will consider the detailed application of ICA component separation technique (already described in Chapter 2) to data from the *WMAP* satellite.

We focus our attention on the properties of diffuse Galactic foregrounds at high and intermediate latitude as traced by specific templates of the emission. This is necessary for two reasons. Firstly, it is likely that the physical conditions within the Galactic plane are different to those at higher latitudes, so that the assumption that the templates can be scaled to a given wavelength by a single factor (or spectral index) is almost certainly invalidated, thus a global analysis will be compromised. Secondly, the complex morphology of the Galactic plane results in non-negligible correlations between the different emission mechanisms, thus violating one of the assumptions of the ICA approach.

We first calibrated the performance of the method using Monte Carlo simulations and by comparison to the simpler template based method. Sections 3.5 and 3.6 attempt to improve our knowledge of the spectral and physical properties of the foreground components based on the scaling properties of the templates. An assessment of the efficiency of the method is then afforded by studying the power spectrum of the cleaned CMB data, whilst a final iterative application of ICA attempts to improve the removal of foreground contamination from the data.

### 3.2 and its use for foreground component studies

We refer to Chapter 2 for a detailed description of the ICA approach and its implementation in the code. Here, we just remind the reader of the fact that, although is a non linear algorithm, the returned components are defined as linear combinations of the input data, whose weights are the

elements of the matrix  $\mathbf{W}$  (see Equation 2.11):

$$y_j = \sum_{i=1}^M w_{ij} x_{v_i}. \quad (3.1)$$

Once the matrix  $\mathbf{W}$  is determined, the independent components are recovered via the relation:

$$x_{v_i} = \sum_{j=1}^N w_{ij}^{-1} y_j \quad (3.2)$$

with  $i = 1, \dots, N$ .

The coefficients associated with the foreground templates define the effective contamination of the CMB anisotropy signal by each foreground. After normalising the weights so that the factor associated with the microwave sky map is unity, we can then interpret the modified template weights as correlation coefficients between the foregrounds and a given channel of the microwave data. Therefore with  $\ell$ -derived CMB sky maps at different frequencies, we can derive the frequency dependence of the coupling coefficients associated with a given template or physical foreground component.

Such an analysis is directly analogous to the simple template fitting scheme commonly used in the field. Therefore, we also utilise the technique here as a convenient point of reference for our results. In general, the cross-correlation measure,  $\alpha$ , between a data vector  $\mathbf{d}$  and a template vector  $\mathbf{t}$  can be measured by minimising:

$$\chi^2 = (\mathbf{d} - \alpha \mathbf{t})^T \cdot \mathbf{M}_{SN}^{-1} \cdot (\mathbf{d} - \alpha \mathbf{t}) = \tilde{\mathbf{d}}^T \cdot \mathbf{M}_{SN}^{-1} \cdot \tilde{\mathbf{d}} \quad (3.3)$$

where  $\mathbf{M}_{SN}$  is the covariance matrix including both signal and noise for the template-corrected data vector  $\tilde{\mathbf{d}} \equiv \mathbf{d} - \alpha \mathbf{t}$ . In the case of  $N$  different foreground components, we have the simple system of linear equations  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , where

$$\begin{aligned} A_{kj} &= \mathbf{t}_k^T \cdot \mathbf{M}_{SN}^{-1} \cdot \mathbf{t}_j, \\ b_k &= \mathbf{t}_k^T \cdot \mathbf{M}_{SN}^{-1} \cdot \mathbf{d}, \\ x_k &= \alpha_k. \end{aligned} \quad (3.4)$$

The correlation method can also be extended to include various constraints on the data, e.g. fixed dust or free-free spectral indices. This is the case described by Hinshaw et al. (2007) in the analysis of the 3 year data of *WMAP*, but we impose no such constraints here.

Once the scaling factors are computed they can be used to clean the *WMAP* data from the estimated Galactic contaminations. The residuals then provide another characterization of the accuracy of the cleaning process. Indeed, residual foreground components are expected that are not well correlated with the adopted foreground templates. They are most easily interpreted as a consequence of the main limitation of the method as currently applied, namely the de-facto assumption of a fixed spectral index for each physical foreground component over the analysed sky coverage. We will show that

an ‘iterative’ application of the algorithm to multi-frequency pre-cleaned microwave data can provide additional insight into such residuals and an additional estimate of the CMB component. The outcome of such an iterative stage will include an improved CMB sky map, plus component maps that represent foreground residuals, although it is unlikely that they can be unambiguously assigned to specific physical components.

### 3.3 Data used in the analysis

The current analysis seeks to clean the *WMAP* data from the main Galactic foreground emissions by applying the algorithm to the observed *WMAP* sky signal at each frequency together with appropriate templates for each component. The current implementation of the algorithm assumes that each input sky map has the same spatial resolution. Thus we perform our analysis on sky maps convolved from their original resolution to an effective  $1^\circ$  Gaussian beam.

The *WMAP* satellite (Hinshaw et al., 2007) observes the sky with ten so-called differencing assemblies (DAs), with frequency dependent resolution of approximately  $0.23^\circ$  to  $0.93^\circ$ . The frequency ranges from  $\sim 23$  GHz (K-band) up to  $\sim 94$  GHz (W-band): there are two channels in the Q- and V-bands, and four channels in the W-band whilst there is only one channel at each of the K- and Ka-bands. The corresponding sky map data are available on the LAMBDA website<sup>1</sup> in a HEALPix<sup>2</sup> pixelisation scheme, with a pixel resolution parameter of  $N_{side} = 512$ .

We have smoothed each of the sky maps to an effective resolution of  $1^\circ$  (initially deconvolving the azimuthally symmetric beam profile for each DA), then we have combined the maps in the Q-, V- and W-bands using a simple average over the corresponding DAs in order to generate a single map for each frequency band. The effective central frequency of each DA depends on the spectrum of the emission being considered. For the band-averaged data, we simply adopt the values 23, 33, 41, 61 and 94 GHz for K- through W-band. Finally, we converted the *WMAP* data from thermodynamic temperature to brightness (antenna) temperature units, in order to make the data consistent with the templates.

Since regions close to the Galactic plane are the most seriously contaminated by foregrounds and the spectral and spatial nature of the integrated emission is complex, we exclude them according to the *Kp2* and *Kp0* masks provided by the *WMAP* team (Bennett et al., 2003). The Galactic part of the masks as well as the point-source exclusion regions are considered sufficient and no modifications are made.

The 408 MHz radio continuum all-sky map of Haslam et al. (1982) is used as synchrotron template as well as the difference between the K and Ka *WMAP* data, as suggested by Hinshaw et al. (2007). The last one in part compensates for possible errors introduced by using the Haslam template, presumably accounting for the change in morphology of the emission at microwave frequencies. However, the interpretation of the scalings is complicated by the fact that the difference map must also include a free-free component, and probably a contribution from the anomalous dust correlated emission. The all-sky  $H\alpha$ -map produced by Finkbeiner (2003) is adopted as tracer of the free-free emission, and the

<sup>1</sup> Legacy Archive for Microwave Background Data Analysis – <http://lambda.gsfc.nasa.gov/>.

<sup>2</sup><http://healpix.jpl.nasa.gov>.

Finkbeiner et al. (1999) FDS8 predicted emission at 94 GHz as the reference template for thermal and anomalous dust emission. All these maps are described in more detail in Section 1.4.

### 3.4 Monte Carlo simulations and calibration of method

As the first step of our analysis, we have tested using realistic simulations of the *WMAP* observations to calibrate the accuracy of the method for template fitting. Specifically, we performed 1000 simulations of the microwave sky at each of the 5 *WMAP* frequencies, each containing a realisation of the CMB signal, the Galactic foreground emission, and instrumental noise appropriate to the specific channel. The study was undertaken at an effective resolution of  $1^\circ$ .

Each simulation was used as an input to , together with the foreground emission templates, and the scaling factors computed. Appendix A provides full details of their statistical distribution as a function of both the applied Galactic cut ( $Kp2$  or  $Kp0$ ), and the non-linear function. In summary, the main conclusions are as follows.

- (i) The  $t$ -function is inappropriate for template fitting – the statistical distributions of the returned coupling coefficients are asymmetric and highly biased with respect to the input values.
- (ii) The  $p$ - and  $g$ - functions provide similar results – the distributions of the coupling coefficients are well-described by Gaussians with a weak bias of the mean that is essentially insignificant for our studies.
- (iii) There is evidence of cross-talk, specifically an anti-correlation, between the synchrotron and dust coefficients when either the Haslam or K-Ka templates are used to characterise the synchrotron emission. In the latter case, correlation between the template and free-free emission is also seen.
- (iv) The results from the simple  $\chi^2$  analysis indicate that the uncertainties are significantly larger than for the method, substantially so when compared to the  $p$ -function. Moreover, the method seems to demonstrate more cross-talk between components than seen for the analysis.

The uncertainties for the scaling factors presented in later sections of the chapter are derived directly from these simulations.

### 3.5 Analysis with the Haslam map as synchrotron template

For our initial analysis we consider template fits to the *WMAP* data for all five frequency bands using and the Haslam map as the model for the synchrotron emission. Table 3.1 summarises the results determined using both  $p$ - and  $g$ -functions for the  $Kp2$  and  $Kp0$  sky coverages. The corresponding  $Kp2$  results based on a simple  $\chi^2$  analysis (see Appendix A for details) are also shown, together with the template fit coefficients used by *WMAP* for their first year foreground correction (Bennett et al., 2003). We can make the following general observations based on these results.

	synchrotron		dust		free-free	
	$Kp2$	$Kp0$	$Kp2$	$Kp0$	$Kp2$	$Kp0$
	<b>- function p</b>					
$K$	$6.64 \pm 0.44$	$5.46 \pm 0.62$	$5.77 \pm 0.26$	$4.94 \pm 0.33$	$8.58 \pm 0.39$	$6.70 \pm 0.63$
$Ka$	$2.00 \pm 0.43$	$1.68 \pm 0.61$	$2.02 \pm 0.26$	$1.28 \pm 0.32$	$4.31 \pm 0.38$	$2.68 \pm 0.62$
$Q$	$1.01 \pm 0.42$	$0.87 \pm 0.60$	$1.08 \pm 0.26$	$0.42 \pm 0.31$	$2.89 \pm 0.38$	$1.36 \pm 0.61$
$V$	$0.26 \pm 0.40$	$0.26 \pm 0.55$	$0.60 \pm 0.24$	$0.03 \pm 0.28$	$1.39 \pm 0.36$	$0.01 \pm 0.57$
$W$	$0.05 \pm 0.35$	$0.05 \pm 0.49$	$0.95 \pm 0.21$	$0.46 \pm 0.25$	$0.67 \pm 0.32$	$-0.49 \pm 0.52$
	<b>- function g</b>					
$K$	$6.52 \pm 0.54$	$5.66 \pm 0.80$	$6.07 \pm 0.29$	$5.71 \pm 0.41$	$8.52 \pm 0.54$	$6.60 \pm 0.82$
$Ka$	$2.05 \pm 0.53$	$1.87 \pm 0.79$	$2.09 \pm 0.28$	$1.66 \pm 0.40$	$4.21 \pm 0.53$	$2.75 \pm 0.80$
$Q$	$1.09 \pm 0.52$	$1.08 \pm 0.78$	$1.11 \pm 0.28$	$0.69 \pm 0.40$	$2.79 \pm 0.53$	$1.47 \pm 0.79$
$V$	$0.39 \pm 0.49$	$0.56 \pm 0.72$	$0.58 \pm 0.27$	$0.19 \pm 0.36$	$1.31 \pm 0.50$	$0.19 \pm 0.76$
$W$	$0.18 \pm 0.43$	$0.31 \pm 0.65$	$0.93 \pm 0.23$	$0.59 \pm 0.32$	$0.61 \pm 0.43$	$-0.36 \pm 0.67$
	<b><math>\chi^2</math> analysis</b>					
$K$	$5.96 \pm 0.57$		$6.38 \pm 0.29$		$8.00 \pm 0.62$	
$Ka$	$1.83 \pm 0.56$		$2.25 \pm 0.29$		$3.71 \pm 0.60$	
$Q$	$0.94 \pm 0.54$		$1.24 \pm 0.28$		$2.32 \pm 0.58$	
$V$	$0.18 \pm 0.49$		$0.69 \pm 0.25$		$0.88 \pm 0.53$	
$W$	$0.01 \pm 0.37$		$1.02 \pm 0.19$		$0.23 \pm 0.41$	
	<b>Bennett et al. values (constrained: <math>\beta_s = 2.7</math>; <math>\beta_{ff} = 2.15</math>)</b>					
$K$	-		-		-	
$Ka$	-		-		-	
$Q$	1.01		1.04		(1.92)	
$V$	0.34		0.62		(0.82)	
$W$	0.11		0.87		(0.32)	

Table 3.1: Values of the coupling coefficients in antenna temperature units determined between the 3-year *WMAP* data and three foreground emission templates, at  $1^\circ$  resolution. The Haslam 408 MHz map is adopted as the synchrotron template. The  $\chi^2$  analysis is performed using the two non-linear functions  $p$  and  $g$  and for the  $Kp2$  and  $Kp0$  masks. The corresponding  $Kp2$  results for a simple  $\chi^2$  analysis are provided for comparison. In addition, we provide the values from the Bennett et al. (2003) fits to the Q,V and W-bands performed with constraints imposed on the synchrotron and free-free spectral indices. These are the coefficients adopted by *WMAP* for their first-year foreground corrections. The units are  $\mu\text{K}/\text{K}$  for synchrotron,  $\text{mK}/\text{mK}$  for dust and  $\mu\text{K}/\text{R}$  for free-free emission respectively.

The frequency dependence of the coefficients follows the generally expected trends, and in particular, the synchrotron and free-free coefficients decrease with increasing frequency. The situation with dust emission is more complex: the dust emission decreases from 94 GHz to 61 GHz as would be expected for thermal dust emission. However, there is a strongly increasing contribution from a dust correlated component with decreasing frequency below 61 GHz. This presumably constitutes further evidence for a new emission mechanism for dust, the precise physical nature of which remains unknown. Overall, however, there is a minimum contribution from the diffuse Galactic foregrounds at  $\sim 61$  GHz, which can be regarded as defining the optimal range for CMB temperature measurements.

The amplitude of the derived coefficients depends on the extent of the mask applied to the data. For all the foreground components, almost without exception, the coefficients derived with the *Kp0* sky coverage are less than for *Kp2*. This systematic trend presumably results from genuine spectral variations of the foregrounds on the sky. Indeed, this behaviour is particularly noticeable for the free-free emission, and may reflect the fact that the *Kp0* mask excludes strong emission regions near the Galactic plane, whilst the free-free emission remains weak at higher latitudes. These strong emission regions may also manifest real variations in scaling dependence relative to the  $H\alpha$  template due to variations in the dust absorption or in the temperature of the ionised gas in the medium latitude region present in the *Kp2* mask but removed by *Kp0*.

The numbers are also in good statistical agreement with our own simple  $\chi^2$  results, provided only for the *Kp2* sky coverage. Interestingly, the synchrotron and dust amplitudes are systematically lower (higher) for the  $\chi^2$  method. However, these results in general reflect the weak cross-talk seen between the fitted amplitudes in Appendix A. The free-free amplitudes are also typically  $1\text{-}\sigma$  higher for the analysis.

Similarly, the constrained fits from Bennett et al. (2003) are of comparable amplitude for the synchrotron and dust coefficients. These values were derived by fixing the synchrotron spectral index at a value of 2.7, and the free-free to 2.15 for frequencies above Q-band. The former constraint explains why the synchrotron amplitude remains higher at V- and W-band for the first-year *WMAP* corrections, although the Q-band amplitude is in excellent agreement with the numbers. It is curious to note that the constrained synchrotron index adopted is inconsistent with the Q-band amplitude, from which one can infer a spectral index of  $\sim 3$  between 408 MHz and 61 GHz. However, Bennett et al. (2003) specifically state that their template method is not particularly physical, but removes foregrounds outside the *Kp2* cut at the required level.

Finally, we would like to comment on the impact of dipole subtraction for the analysis. Specifically, the results we have presented here, including those of Bennett et al. (2003), do not subtract a best-fit dipole amplitude from the data and templates before fitting the template amplitudes. Since other independent analyses, e.g. Davies et al. (2006), do prefer to remove the dipole, we have tested its impact on the results. In fact, we find that the dust and free-free coefficients remain largely unchanged, whilst the synchrotron values become negative at high frequencies. It may be that this is connected with the projection of the North Polar Spur in the Haslam template onto the best-fit dipole computed for that template.

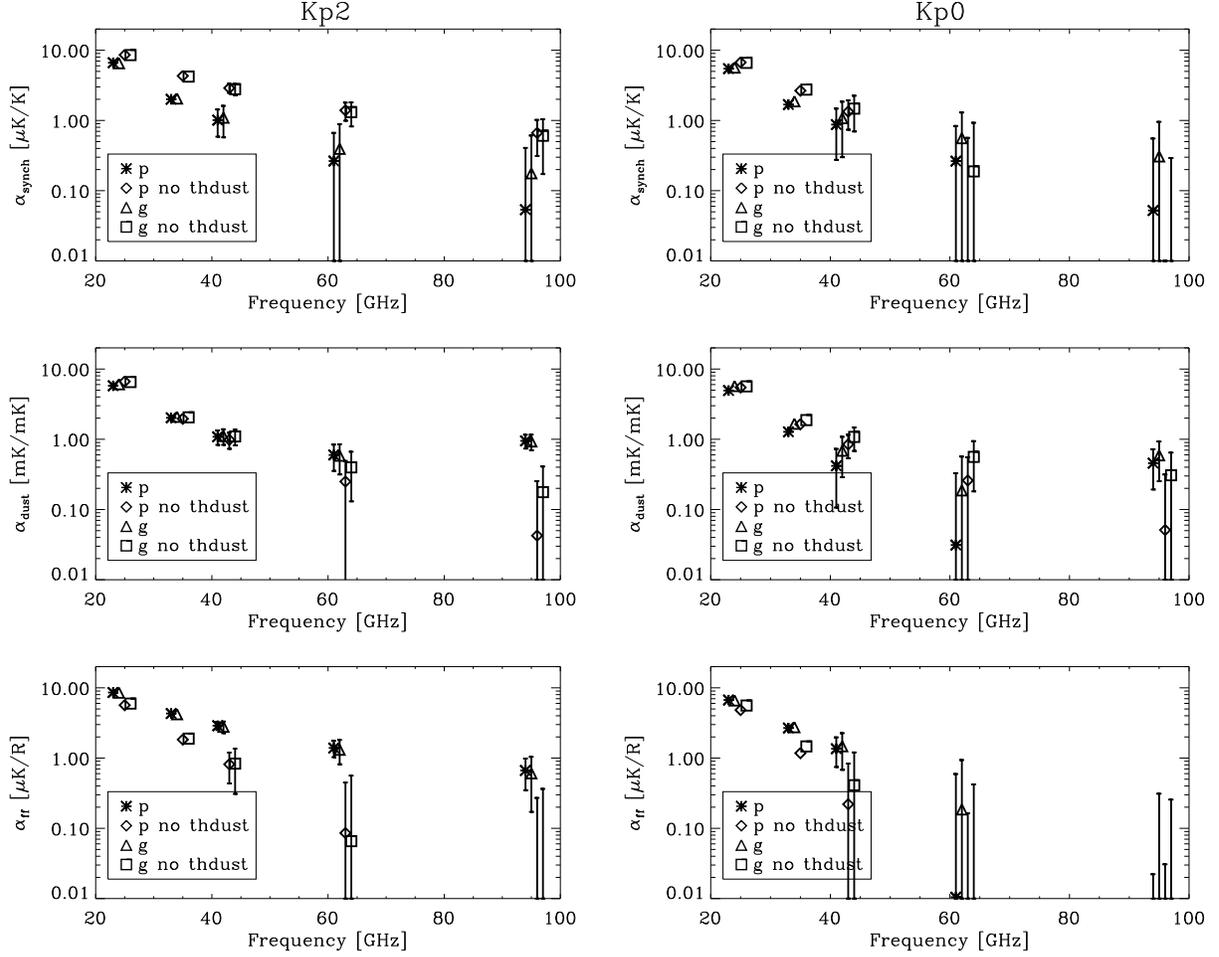


Figure 3.1: Comparison between scaling factors obtained from *WMAP* data with and without the thermal dust emission, at the resolution of  $1^\circ$ . At the lowest frequencies, the error bars are smaller or comparable to the size of the plot symbols. The only significant change is at W-band for the dust coefficients.

### 3.5.1 Spectral index of foreground emissions.

We utilise the derived frequency dependence of the scaling factors to parameterise the spectral behavior of the foreground emission components. For each component, we fit the corresponding coefficients with a power law model of the form  $A_{norm}(\nu/\nu_0)^{-\beta}$ .  $A_{norm}$  is the amplitude of the emission of a specific physical component at the reference frequency  $\nu_0$ , which we take as the K-band (23 GHz). The parameters values of the fit are shown in Table 3.2.

The results for the synchrotron emission generally indicate a steep spectral index consistent with a value in excess of  $\beta_s = 3.0$ . Curiously, the index is flatter for the *Kp0* mask, particularly when using the *g*-function. The amplitudes, however, fall with decreasing sky coverage – this is generally true for all of the emission components, suggesting that there remains within the *Kp2* to *Kp0* transition region emission that differs in nature from genuine high latitude emission. It is also interesting to note that the amplitude at 23 GHz implies a spectral index of  $\sim 3$  relative to the 408 MHz emission, so that in general there are hints of steepening at higher frequencies, as has been previously noted by the *WMAP* team.

	Synchrotron		Anomalous Component		Free-Free	
	<i>Kp2</i>	<i>Kp0</i>	<i>Kp2</i>	<i>Kp0</i>	<i>Kp2</i>	<i>Kp0</i>
	– function p					
$\beta$	$3.35 \pm 0.47$	$3.28 \pm 0.79$	$3.29 \pm 0.32$	$4.48 \pm 0.78$	$1.87 \pm 0.15$	$2.92 \pm 0.55$
$A_{norm}$	$6.64 \pm 0.43$	$5.46 \pm 0.62$	$5.68 \pm 0.26$	$4.85 \pm 0.33$	$8.56 \pm 0.38$	$6.79 \pm 0.63$
	– function g					
$\beta$	$3.10 \pm 0.53$	$2.82 \pm 0.80$	$3.35 \pm 0.35$	$4.14 \pm 0.75$	$1.92 \pm 0.22$	$2.72 \pm 0.65$
$A_{norm}$	$6.51 \pm 0.53$	$5.62 \pm 0.80$	$5.98 \pm 0.29$	$5.63 \pm 0.41$	$8.51 \pm 0.52$	$6.68 \pm 0.81$

Table 3.2: Spectral index  $\beta$  and normalisation factor  $A_{norm}$  obtained fitting values of the coupling coefficients for synchrotron, the anomalous component of dust and free-free emission, with different masks.

Since the *WMAP* frequency range does not allow a detailed study of the spectral behaviour of the thermal dust component, we consider the spectral properties of the anomalous dust correlated emission. Understanding the nature of this component remains of great importance for the study of CMB anisotropy and for motivating physical models of the emission. We use a single dust template to specify the morphology of both the thermal dust emission and the anomalous dust correlated component. Thus, in order to study the latter cleanly, we must apply a correction for the thermal dust contribution. Specifically, the scaling factors have been recomputed using  $\beta_a$  after subtracting the dust template from the data assuming that the FDS8 dust model (see Section 3.3) provides an exact model of this emission component. Under this assumption a persistent correlation with the dust template will be due to the anomalous component, spatially correlated with the thermal dust, but with a different physical nature. Figure 3.1 shows the scaling factors for the dust emission without the thermal component, compared with the values of the original analysis. In fact, the values differ significantly for the dust coefficients only at 61 and 94 GHz as might be expected since the thermal component falls off rapidly with decreasing frequency. Note also that whatever dust residuals remain at the highest frequency are statistically insignificant, providing strong support to the FDS8 model of thermal dust emission.

The values of the spectral index  $\beta_a$  of the anomalous dust component determined on the *Kp2* sky coverage are somewhat steeper than the value of 2.85 obtained by Davies et al. (2006). However, they have also noted spectral indices as steep as 3.8 in several dust dominated regions at mid- to high-latitude. As with the synchrotron component, the amplitude of the emission drops for the *Kp0* mask, and consequently the spectral index steepens noticeably. One might speculate that this is related to the differing properties of the anomalous component closer to the Galactic plane flattening the index of the *Kp2* results. In the context of the spinning dust models proposed by Draine & Lazarian (1998a) to explain the anomalous emission, the behaviour may reflect the properties of the spinning dust in different phases of the interstellar medium, e.g. a greater or lesser admixture of spinning dust in the warm ionised medium (WIM) depending on latitude.

Finally, we consider the properties of the free-free emission as represented by the coupling coefficients determined relative to the  $H\alpha$  template. Indeed, before determining the specific spectral behaviour, it is worth noting that the derived coefficients are not in accord with expectations for the

	Free-Free - Intensity			
	– function p		– function g	
	$Kp2$	$Kp0$	$Kp2$	$Kp0$
$\beta$	$0.13 \pm 0.15$	$-0.92 \pm 0.55$	$0.08 \pm 0.22$	$-0.72 \pm 0.65$
$A_{norm}$	$0.14 \pm 0.01$	$0.11 \pm 0.01$	$0.14 \pm 0.01$	$0.11 \pm 0.01$

Table 3.3: Spectral index  $\beta$  and normalisation factor  $A_{norm}$  obtained fitting values of the intensity coupling coefficients for free-free emission, with different masks.

$H\alpha$  to free-free conversion factor for which a value of 8000 K is conventionally adopted for the thermal electron temperature. In fact, the K- and Ka-band results for  $Kp2$  sky coverage are more consistent with temperatures in the range 5000-6000 K. Curiously, however, the coefficients at higher frequencies are increasingly consistent with a higher temperature value. In terms of a power-law fit for the frequency spectrum, this corresponds to a notably flatter spectral index than the typical value of 2.14 for free-free emission, with the best-fit value for the  $p$  function fits some  $2\sigma$  away from this canonical amplitude. However, the results when a  $Kp0$  mask is applied are significantly different, favouring an even steeper slope than expected for free-free, and with normalisation amplitudes corresponding to an electron temperature below 4000 K. This latter behaviour was also observed by Davies et al. (2006) for 5 regions specifically selected to be dominated by free-free emission.

In order to assess the consistency of our results with the theoretically motivated spectral index as compared to the apparently anomalous best-fit models, we have fitted the values of the coupling coefficients by an idealised model for the free-free emission,  $A_{norm}(\nu/\nu_0)^{-2.14}$ , and evaluated the goodness-of-fit of the model. The results are consistent with the  $\beta_{ff} = 2.14$  model, even in the case of the  $Kp0$  mask. However, in a set of 1000 simulations in which the simulated free-free emission followed the theoretically motivated spectrum, no cases were found for which there were such large changes in the fitted spectral slopes determined for  $Kp2$  or  $Kp0$  sky coverage as those observed with the data. This could reflect the ideal nature of the simulated emission, but irrespective of this, although our results are statistically consistent with the expected free-free scaling with frequency, there are inconsistencies that are difficult to reconcile.

Arguably then, the result for the  $Kp2$  mask could be considered an independent verification of the behaviour observed in Dobler & Finkbeiner (2008a). Motivated by their analysis, we augment our assessment of the free-free spectral behaviour by fitting the coefficients in intensity units. The results are presented in Table 3.3 and Figure 3.2. Dobler & Finkbeiner (2008a) found that the  $H\alpha$ -correlated emission showed a bump near to 50 GHz, although the exact behaviour was sensitive to the type of CMB estimator that they pre-subtract from the *WMAP* data before fitting template amplitudes. In our analysis, no such subtraction is necessary, since the coefficients are essentially determined as part of the process of deriving the best CMB estimate from the data. In fact, our  $Kp2$  amplitudes show more extreme behaviour at high frequencies, and indicate a systematically rising spectrum, although the best-fit is also consistent within errors with the slowly falling theoretically expected spectrum. However, our  $Kp0$  results are significantly steeper than either the  $Kp2$  or expected slopes. Referring

again to the results from the regional analysis of Davies et al. (2006), there is a wide range of spectral behaviour, with region 1 exhibiting a steeper rise in frequency than seen here for *Kp2*, region 5 being somewhat consistent with the observations here, but most regions falling off rapidly. Interestingly, the simple  $\chi^2$  analysis, for which a best-fit spectral index of  $0.2 \pm 0.3$  is found, is very consistent with theoretical expectations ( $\sim 0.14$ ).

The major difference between all of these results is how the method treats the CMB contribution. Dobler & Finkbeiner (2008a) subtract an estimate of the CMB based on a variant of the *WMAP* ILC method, Davies et al. (2006) explicitly include a CMB covariance term in fitting the data directly with three templates, our  $\chi^2$  analysis simply ignores the CMB contribution in fitting the templates, whilst the method is specifically attempting to construct the best estimate of the CMB sky and computing template coefficients accordingly. Clearly more work is needed to understand the relative merits and problems in these approaches.

Finally, Dobler & Finkbeiner (2008b) interpret the results of both their full-sky analysis and an assessment of the Gum nebula region in terms of an enhanced emissivity due to a spinning dust contribution from the WIM. It is difficult to make a statement either in support or against this conclusion based on our results, given the large difference between the fits on *Kp2* and *Kp0*, and the apparent dependence on the CMB subtraction method. Perhaps it reflects the different properties of the free-free emission close to the plane, or the presence indeed of spinning dust in the WIM, or simply cross-talk between different physical components that confuses the spectral analysis. However, we remind the reader that our fits for both *Kp2* and *Kp0* sky coverage are statistically consistent with the expected scaling for free-free emission with frequency,  $\beta_{ff} = 2.14$ .

### 3.5.2 Single Year Analysis

In order to test the stability of the coupling coefficients between the foreground templates and the *WMAP* data to potential systematics, e.g. annual instrumental or calibration variations, we have performed the same analysis for the yearly sky maps. These results are plotted in Figure 3.3 together with the 3-year data coefficients. We find no evidence of significant annual variations in the coefficients.

## 3.6 Analysis with the K-Ka map as synchrotron template

One of the limitations of the previous foreground analysis is that the Haslam map used as template for the synchrotron emission is unlikely to be completely representative of the foreground morphology at the *WMAP* frequencies due to spectral index variations across the sky. Therefore, we repeated the analysis using the K-Ka map as an internal template for the synchrotron emission. We then computed the scaling factors for the Q, V and W channels. The results are shown in Table 3.4. As before, the uncertainties of the coupling coefficients have been computed from 1000 simulations. Once again, the  $p$ -function seem to be statistically more robust than the corresponding  $g$ -function ones.

Table 3.4 also contains results returned by the  $\chi^2$  minimisation method, both our own and those

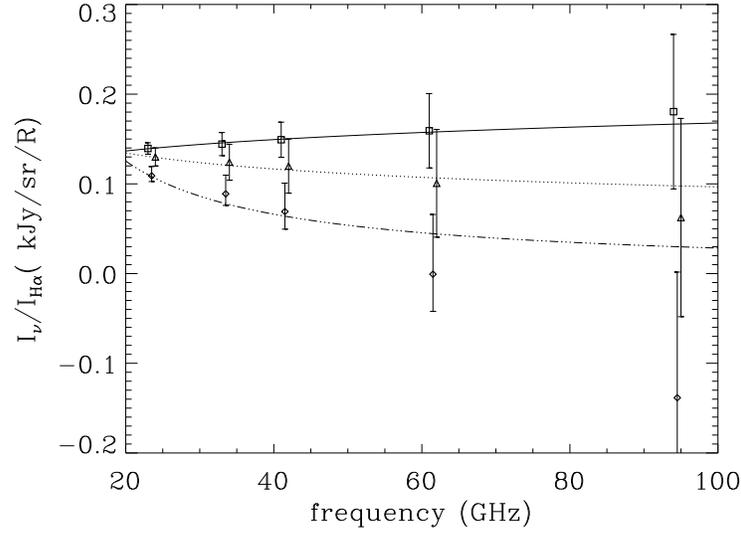


Figure 3.2: The coupling coefficients in intensity units for free-free emission as traced by the  $H\alpha$  template. Squares represent the derived amplitudes for  $Kp2$  sky coverage from the  $p$ -function analysis, diamonds are for the corresponding analysis on  $Kp0$ , whilst triangles show the results for a simple  $\chi^2$  analysis on  $Kp2$ . Best-fit curves are also shown. The  $Kp2$  results show an anomalous rising spectrum for the free-free emissivity, whereas the  $Kp0$  and  $\chi^2$   $Kp2$  are steeper than expected, the former significantly so.

from *WMAP* (Hinshaw et al., 2007). For our analysis, we did not impose any constraints on the spectral behaviour of the various physical components. In contrast, Hinshaw et al. (2007) explicitly account for the free-free signal present in the K-Ka template, and impose constraints on the free-free ( $\beta_{ff} = 2.14$ ) and thermal dust ( $\beta_d = 2$ ) spectral indices. The coefficients derived for the K-Ka and dust templates are in excellent agreement for all three methods. The free-free values, however, are notably different in all three cases. The difference to the *WMAP* results could, at least in part, be due to their application of a large correction for dust absorption, although comparing coefficients at all three frequencies the explanation is clearly more complex. A more useful observation in regard to the free-free coefficients is that they are universally lower than in the corresponding Haslam case for the Q-band but almost identical for W-band. This directly reflects the fact that the K-Ka template contains not only synchrotron but also free-free emission. The latter also indicates that the results continue to prefer a flatter slope for the free-free emission than expected from theory, and as derived by the  $\chi^2$  analysis.

The most interesting results are related to the amplitudes of the coupling coefficients for the dust. In particular, note the low value of the dust contamination in the Q-band in contrast with previous findings for the Haslam case. This could be interpreted as indicating the absence of an anomalous dust correlated component, in apparent contradiction to previous results. However, it is also consistent with a picture in which the K-Ka template contains contributions not only from the synchrotron and free-free emission, but also the anomalous component, which the Haslam map certainly does not. At W-band and for the  $Kp2$  sky coverage, the K-Ka contribution is essentially zero, so that there is no effective contribution of

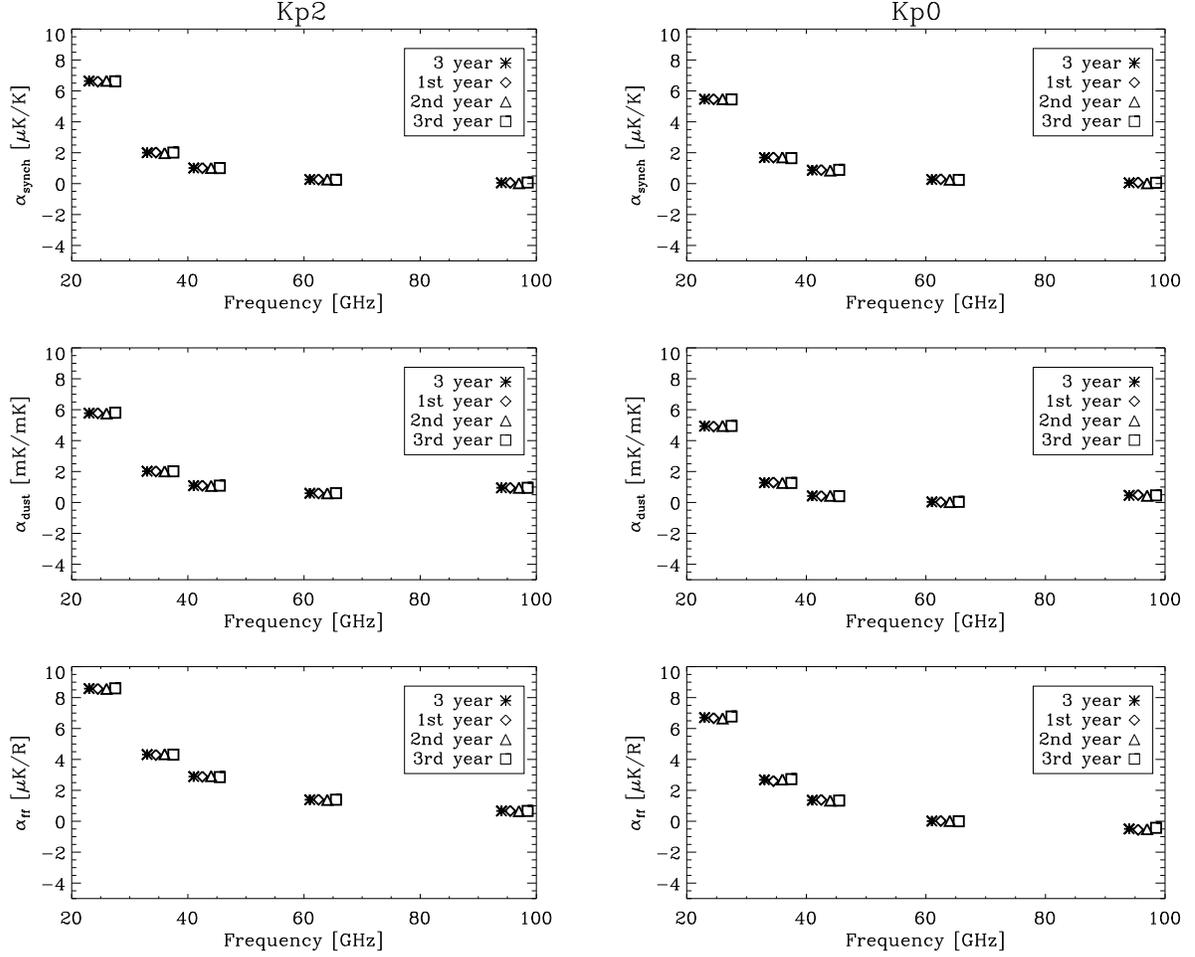


Figure 3.3: Comparison between scaling factors determined for the yearly *WMAP* sky maps. The results for the 3-year *WMAP* data are also shown for reference. All results are determined using the  $p$ -function.

this embedded anomalous component, and the dust coefficients should be consistent with the thermal dust emission. Indeed, there is again strong support for the FDS8 model of the thermal dust emission, although for the *Kp0* cut the amplitude drops as before.

In order to understand whether there is a self-consistent picture with the Haslam results, we can perform some simple numerical comparisons. If we assume that the *WMAP* data contains foregrounds comprised of synchrotron, free-free and anomalous dust contributions as described in Table 3.1, then we can infer the extent of these components present in the K-Ka template, and then predict the amplitudes expected for the free-free and dust contributions in Table 3.4 allowing for that fraction contained in the K-Ka map. The consistency is excellent for both  $p$  and  $g$  function results on the *Kp2* sky cut, and certainly good for the *Kp0* coverage. We have also made a more detailed study using simulations, with the same results. Of course, this does not imply that the K-Ka template is not a significant improvement on the Haslam map, but that the gross features of the two foreground models are in agreement. The quality of the two synchrotron templates will be considered further in the following sections.

	Synchrotron		Dust		Free-free	
	$Kp2$	$Kp0$	$Kp2$	$Kp0$	$Kp2$	$Kp0$
<b>- function p</b>						
$Q$	$0.24 \pm 0.04$	$0.21 \pm 0.07$	$0.17 \pm 0.32$	$-0.13 \pm 0.41$	$1.92 \pm 0.42$	$0.48 \pm 0.70$
$V$	$0.05 \pm 0.04$	$0.03 \pm 0.06$	$0.50 \pm 0.29$	$0.40 \pm 0.30$	$1.19 \pm 0.40$	$-0.16 \pm 0.66$
$W$	$-0.01 \pm 0.03$	$-0.04 \pm 0.05$	$1.00 \pm 0.26$	$0.77 \pm 0.34$	$0.73 \pm 0.35$	$-0.41 \pm 0.58$
<b>- function g</b>						
$Q$	$0.24 \pm 0.09$	$0.16 \pm 0.15$	$0.16 \pm 0.52$	$0.03 \pm 0.80$	$1.76 \pm 0.66$	$0.84 \pm 0.98$
$V$	$0.06 \pm 0.08$	$-0.004 \pm 0.14$	$0.23 \pm 0.76$	$0.39 \pm 0.50$	$1.05 \pm 0.63$	$0.16 \pm 0.93$
$W$	$0.01 \pm 0.072$	$-0.05 \pm 0.12$	$0.94 \pm 0.44$	$0.80 \pm 0.66$	$0.56 \pm 0.54$	$-0.20 \pm 0.81$
<b><math>\chi^2</math> analysis (no constraints)</b>						
$Q$	$0.21 \pm 0.09$		$0.42 \pm 0.50$		$1.44 \pm 0.77$	
$V$	$0.03 \pm 0.08$		$0.57 \pm 0.46$		$0.75 \pm 0.70$	
$W$	$-0.01 \pm 0.06$		$1.10 \pm 0.36$		$0.29 \pm 0.54$	
<b>Hinshaw et al.</b>						
$Q$	0.23		0.19		0.99	
$V$	0.05		0.41		0.63	
$W$	0.00		0.98		0.32	

Table 3.4: Values of coupling coefficients in antenna temperature units between the 3-year *WMAP* Q, V, and W-band data and three foreground emission templates at  $1^\circ$  resolution. The K-Ka map is adopted as the synchrotron template. The  $\chi^2$  analysis is performed using the two non-linear functions  $p$  and  $g$  and for the  $Kp2$  and  $Kp0$  masks. The corresponding  $Kp2$  results for a simple  $\chi^2$  analysis are provided for comparison. In addition, we provide the band-averaged values from the Hinshaw et al. (2007) fits performed with constraints imposed on the free-free and thermal dust spectral indices. Note that the *WMAP* team used an  $H\alpha$  template corrected for dust absorption assuming  $f_d = 0.5$  for their analysis. The units are mK/mK for synchrotron, mK/mK for dust and  $\mu\text{K/R}$  for free-free emission respectively.

### 3.6.1 Spectral properties of foreground emissions

A study of the spectral index of the returned K-Ka coefficients would be unphysical, or at least difficult to interpret, since the template is a mix of synchrotron, free-free and anomalous dust emission. Moreover, the interpretation of the spectral behaviour of the returned dust and free-free coefficients is similarly problematic, requiring a detailed understanding of the fraction of the signal associated with the K-Ka template. However, it is worth reiterating, as shown above, that the template fits are in excellent agreement with those in which the Haslam map is used as the synchrotron template, and therefore the spectral behaviour of the components can be interpreted in the same manner.

## 3.7 Evaluation of the Q-, V- and W-band sky maps after foreground cleaning

One of the main aims of the foreground analysis performed thus far with  $\chi^2$  is to promote the possibility of cleaning the *WMAP* data sufficiently to allow its use for cosmological purposes. Therefore, adopting the coupling coefficients derived for various foreground templates, we have studied *WMAP* sky maps cleaned of the Galactic emission.

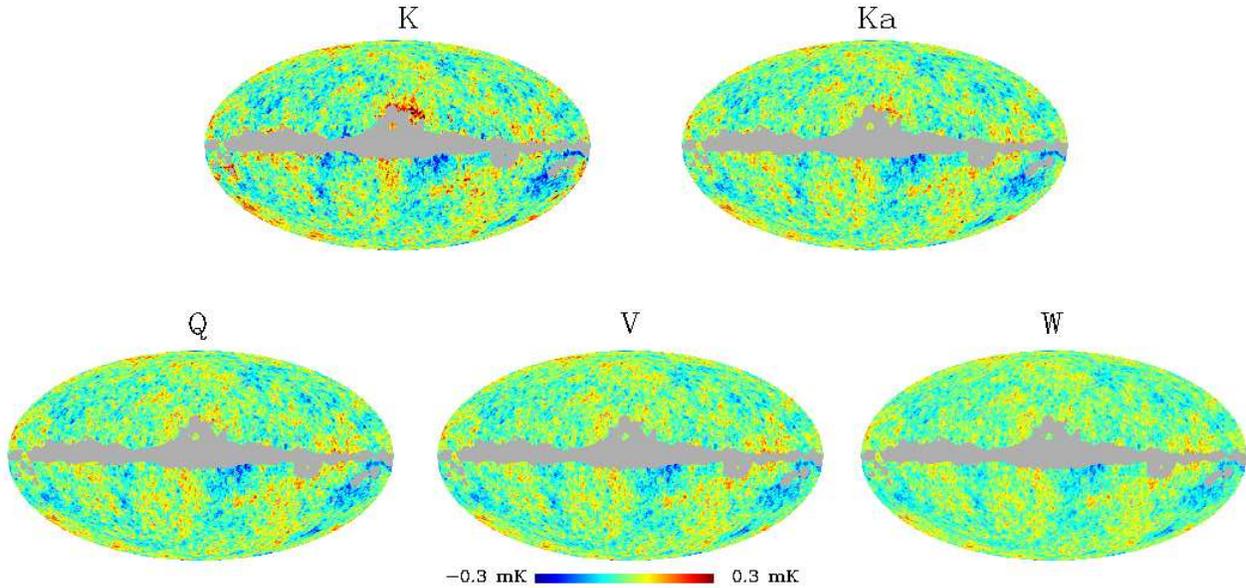


Figure 3.4: *WMAP* data for K-, Ka-, Q-, V- and W-bands cleaned by subtracting foreground templates scaled by the coupling coefficients determined by  $\chi^2$  with the  $p$ -function for the  $Kp2$  sky coverage. Here, the synchrotron template is the Haslam map. The maps are shown in a conventional mollweide projection in a Galactic frame-of-reference, with the north pole at the top of the image and the Galactic Center in the middle with longitude increasing to the left. The regions in grey correspond to the bright Galactic emission and point sources excised from the analysis by the  $Kp2$  mask.

Figure 3.4 shows the residual sky signal of the five *WMAP* channels that have been cleaned using the Haslam data as the synchrotron template. We show only the results based on the  $p$ -function analysis, since there is little visual difference relative to the  $g$ -function. There are clearly residuals close to the Galactic plane for the K and Ka sky maps, particularly near the edge of the  $Kp2$  mask in the vicinity of the Galactic Centre. This excess foreground emission falls off rapidly with frequency, and only a hint can be seen in the Q-band data. Nevertheless, the dominant frequency-independent CMB structure is clear at high latitudes. Figure 3.5 shows the equivalent results for the Q-, V- and W-bands when the K-Ka synchrotron template is adopted. There is little visual evidence of residual foregrounds even close to the masked regions.

As an interesting comparison, we have investigated the difference between the cleaned maps obtained using either the Haslam or the K-Ka synchrotron templates, scaled by the coupling coefficients determined either by our  $\chi^2$  analysis or by the *WMAP* team. Figure 3.6 shows the results for the Q-, V- and W-bands (in neither the first nor second data release did *WMAP* provide cleaned K- or Ka-data, presumably because they do not utilise them for cosmological analysis). The main difference apparent at Q-band is the presence of a bright structure around the mask concentrated near the Galactic Centre both above and below the plane. This is a clear delineation of the K-band residual noted above when using the Haslam template, and is most easily interpreted as Galactic emission that is not traced by the Haslam synchrotron template. It therefore supports the *WMAP* assertion that adopting the K-

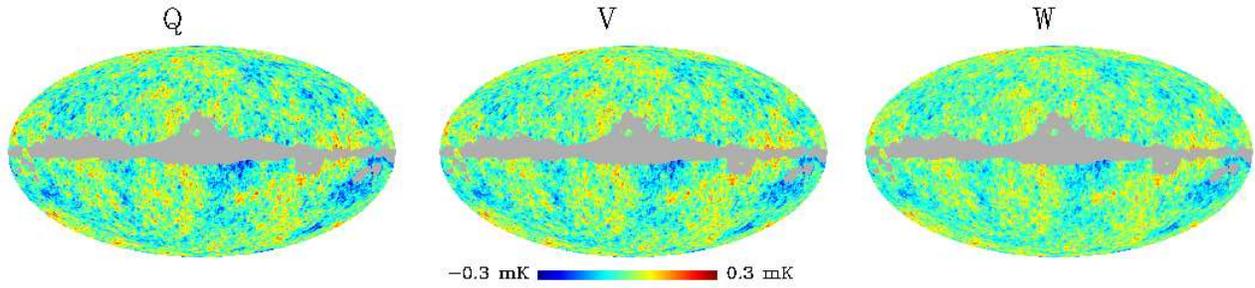


Figure 3.5: *WMAP* data for Q-, V- and W-bands cleaned by subtracting foreground templates scaled by the coupling coefficients determined by  $p$ -function. Here, the synchrotron template is the K-Ka map.

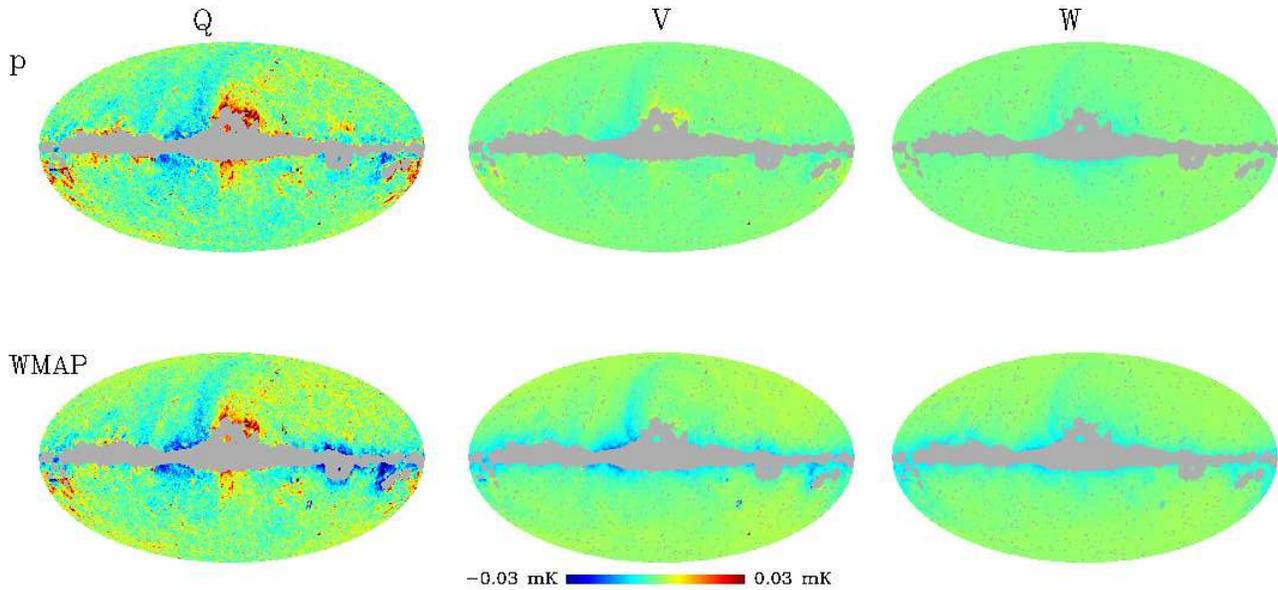


Figure 3.6: Difference between the cleaned maps using the Haslam map and the K-Ka map at 41, 61 and 94 GHz. The top row shows the difference when the  $p$ -function coefficients are used, the lower line shows the equivalent difference maps from when applying the *WMAP* first-year and three-year foregrounds corrections.

Ka map as a template provides a more complete model of the foreground emission at low frequency where other contributions beyond the synchrotron and free-free emission may exist. This structure is seen clearly in both the  $p$ -function and *WMAP* residual plots. Also seen is evidence of a negative residual associated with a well-known feature in the low frequency template – the North Polar Spur – plus other residuals around the edge of the Galactic plane part of the *Kp2* mask, although whether these are more associated with the Haslam or K-Ka analysis is unclear.

The residuals at V- and W-bands show much fainter features, and particularly for the  $p$ -function analysis. Indeed, in this case the plot indicates that the total foreground level determined using either the Haslam or K-Ka maps, plus the usual tracers of free-free and dust emission, are in very good

agreement. This is consistent with the discussion in the previous section. For the data corrected using the *WMAP* template coefficients, the residuals seem to be larger in all cases.

The comparison here cannot easily be used to infer whether the  $\text{WMAP}$  template fits are to be preferred to the *WMAP* ones, but we will attempt to address this issue in the next section. What should be clear is that it is always good practise for a cosmological analysis to study the cleaned maps at different frequencies and test for evidence of a frequency dependence that is the signature of foreground residuals.

In fact, we do exactly that by examining the power spectra of the cleaned maps determined using the MASTER algorithm (Hivon et al., 2002). Figure 3.7 indicates that the maps are essentially insensitive to the residual foreground features that remain in the data due to the different fitting methods and templates, although the quadrupole amplitude is generally suppressed, and notably so when using the Haslam template to trace synchrotron emission. Thus the cleaned maps, particularly at V- and W-bands, can be considered adequate for cosmological analyses, at least when based on the angular power spectrum.

### 3.8 An ‘iterative’ blind component separation study using *WMAP* data pre-cleaned using templates

As demonstrated in the previous section, the *WMAP* data cleaned by foreground templates still show evidence of contamination due to either the inability of the templates to describe real spectral variations in the foreground emission on the sky, or other emission mechanisms not identified with the standard Galactic components/templates. Inspired by the analysis of Patanchon et al. (2005), we have performed an ‘iterative’ blind analysis on the *WMAP* data, *after* pre-cleaning with the foreground templates, in order to study the residuals in more detail.

Various combinations of the cleaned *WMAP* channels are used as input to the  $\text{WMAP}$  algorithm. This then returns a number of components equal to the number of the input sky maps. One of these is identified as a CMB map, cleaned as far as possible from any *residual* emission (see Section 3.8.2). Without exception, at most one of the other returned components can be clearly identified with Galactic residuals, the remainder generally being some combination of anisotropic noise (reflecting the underlying observation pattern of the *WMAP* satellite) and weak residual dipole emission. Figure 3.8 shows maps of the candidate foreground residuals for a number of *WMAP* and template input permutations, as described in the figure caption and discussed in Section 3.8.1

#### 3.8.1 Foreground residuals

The upper row of Figure 3.8 shows the output maps from  $\text{WMAP}$  that we identify with foreground residuals when using the Haslam template for synchrotron emission.

The leftmost figure presents results from analysing the five *WMAP* frequency bands after cleaning using the coefficients previously derived in this chapter (see Table 3.1). There are some bright features along the Galactic plane near to the Galactic Centre and the presence of a structure related to the

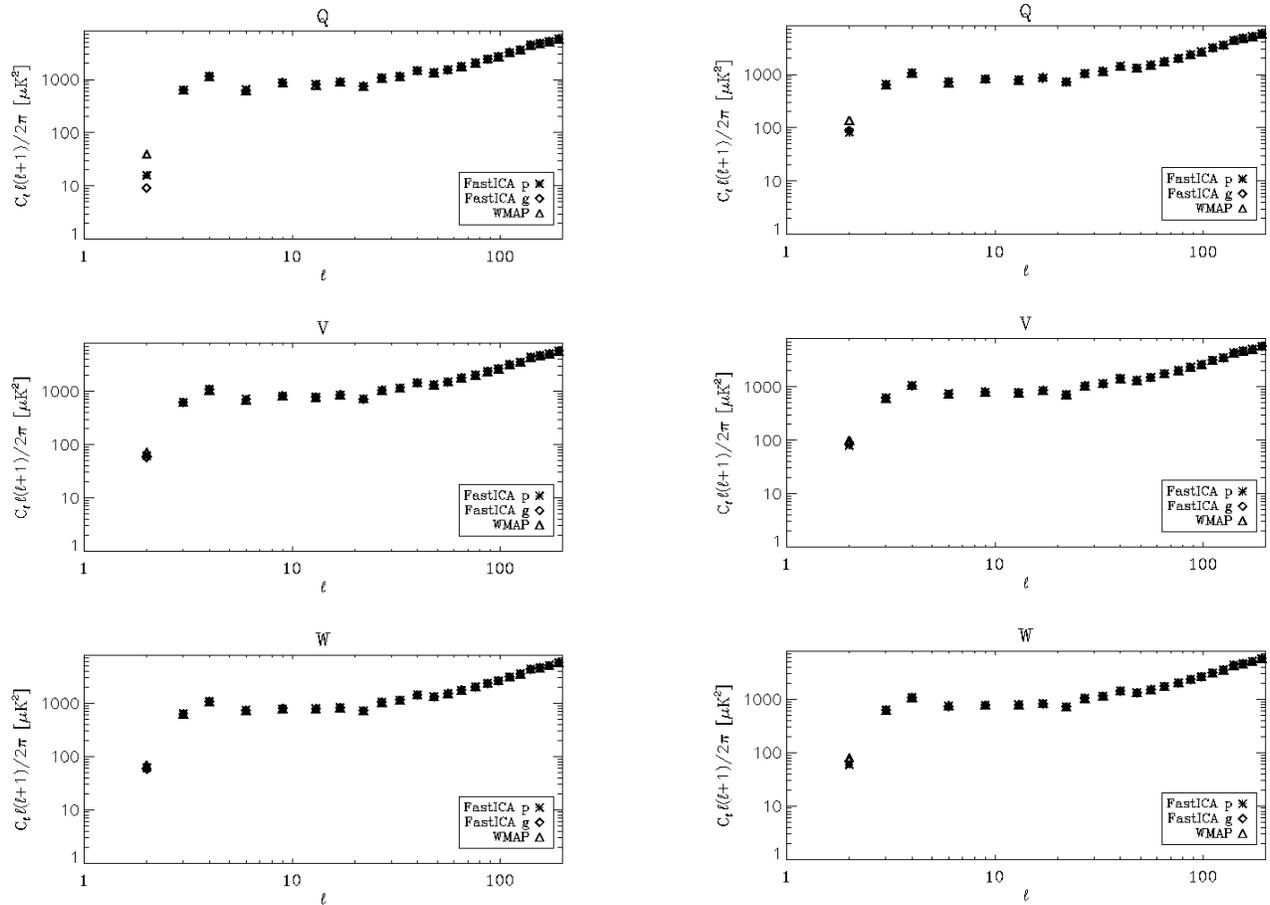


Figure 3.7: Binned power spectra of the Q-, V- and W-maps cleaned from the foreground emission. Their contamination is described by the coupling coefficients determined by  $\beta$  with both the  $p$ - and  $g$ -functions (asterisk and diamond respectively) and by the *WMAP* analysis (triangle). The left figure corresponds to the case when the Haslam map is used as the synchrotron template, the right figure to the corresponding analysis using K-Ka for the low-frequency emission.

North Polar Spur. These are a clear indication of the deficiencies of the Haslam map as a synchrotron template, particularly for cleaning the K- and Ka-band data, with which the residuals are predominantly associated. Some of the positive and negative features presumably reflect either under- or over-subtraction of the true foreground signal by the templates as a consequence of spectral variations.

The middle figure shows the equivalent analysis using just the Q-, V- and W-band pre-cleaned sky maps. The residuals are significantly reduced, and no evidence of the North Polar Spur is now observed. The dominant feature, as before, is the structure close to the Galactic centre. This was originally identified and referred to as the ‘free-free haze’ by Finkbeiner (2004a) (although its origin as free-free emission is unlikely and it is now referred to as the ‘*WMAP* haze’), and was already visible in the foreground residuals plot of Bennett et al. (2003). Moreover, the map is strikingly (though perhaps unsurprisingly) similar to the map obtained by Patanchon et al. (2005), using Spectral Matching Independent Component Analysis (SMICA). They have associated these residual structures with the Ophiuchus complex, the Gum nebula, the Orion-Eridanus bubble and the Taurus region. The right-

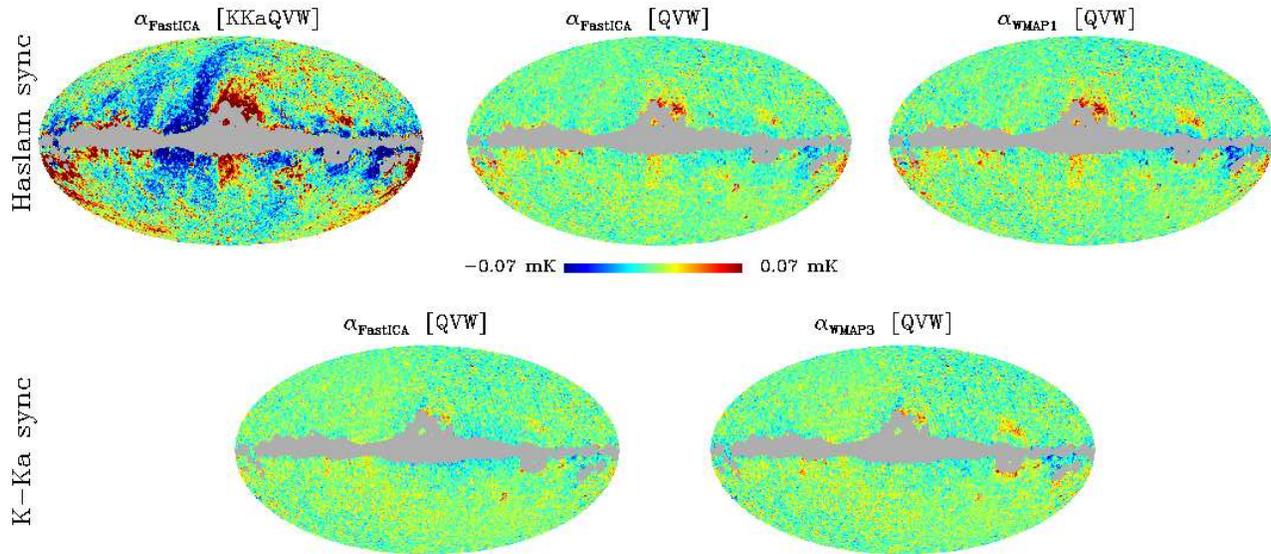


Figure 3.8: Maps of the residual components obtained by  $\alpha$  when applied to the *WMAP* data after pre-cleaning using template fits with coupling coefficients  $\alpha$  derived either using a  $p$ -function and the *Kp2* mask or by the *WMAP* team in Bennett et al. (2003) (*WMAP1*) and Hinshaw et al. (2007) (*WMAP3*). Several cases have been considered: i) input maps K-, Ka-, Q-, V- and W-bands plus Haslam synchrotron template,  $\alpha$  coefficients; ii) input maps Q-, V- and W-bands plus Haslam synchrotron template,  $\alpha$  coefficients; iii) input maps Q-, V- and W-bands plus Haslam synchrotron template, *WMAP* constrained fit coefficients; iv) input maps Q-, V- and W-bands plus K-Ka synchrotron template,  $\alpha$  coefficients; v) input maps Q-, V- and W-bands plus K-Ka synchrotron template, *WMAP* constrained fit coefficients.

hand plot is the same analysis when the three input maps are cleaned using the *WMAP* foreground model of Bennett et al. (2003). The residual features are in excellent agreement with those from the middle plot, although the amplitudes are slightly larger, suggesting that the  $\alpha$  template coefficients provide a modestly better cleaning of the *WMAP* frequency data.

The residual maps from input data cleaned with the K-Ka synchrotron model are shown in the second row of Figure 3.8. The maps can be compared directly to the equivalent Haslam cases in the row above. It is particularly striking that the residuals are substantially suppressed when using the K-Ka template, especially close to the Galactic Centre. This reinforces the latter's superiority for foreground separation purposes, although the fact that it is a mixture of physical emission components makes the interpretation of the results more difficult. As before, the residuals are larger when using the *WMAP* template coefficients, notably so near the Gum nebula – a source of strong free-free emission.

### 3.8.2 CMB component

In order to quantify the improvement connected to the iterative processing step, we have studied the CMB components returned by  $\alpha$  when working on the data pre-cleaned with a template analysis stage. They are very similar, as shown in Figure 3.9. This is also confirmed by the plot of the power spectrum for the CMB map derived from the different cases of the foreground analysis. Figure 3.10

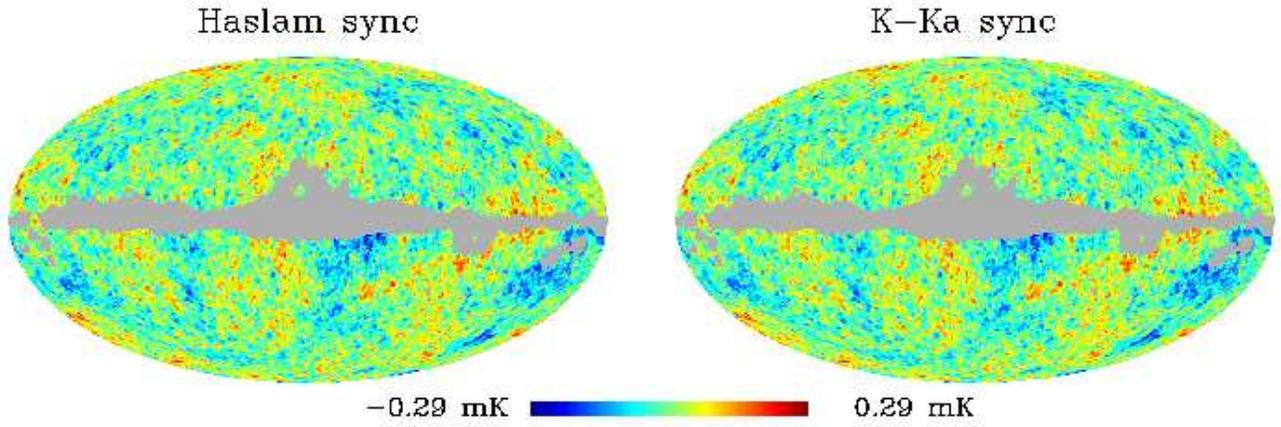


Figure 3.9: Map of the CMB component obtained by  $\text{FastICA}$  with the  $p$ -function using *WMAP* Q-, V- and W-band data pre-cleaned using coupling coefficients estimated with the  $Kp2$  mask, and either the Haslam map as the synchrotron template (left) or the K-Ka map (right). The temperature is measured in mK.

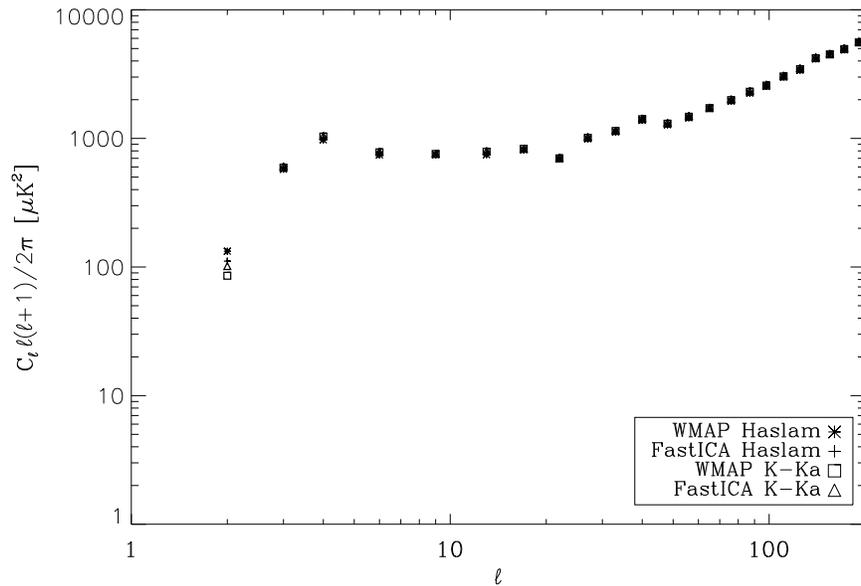


Figure 3.10: Power spectra of the returned CMB components applying  $\text{FastICA}$  to the set of Q-, V- and W-band data from the foreground emissions, using various combinations of foreground templates and fit coefficients.

compares the binned power spectrum of the CMB maps returned from the Q-, V- and W-band data cleaned using the coupling coefficients of the  $\text{IC}^2$  analysis or the *WMAP* values, with either the Haslam or K-Ka maps as synchrotron template. The discrepancy is less than  $150 \mu K^2$ .

Given this, one might argue that the iterative step is unnecessary for *WMAP* analysis, but we fully expect that it will be essential for studying the *Planck* data in detail. Moreover, in this era of precision cosmology, it is the deviations from the theoretically motivated model spectra which are of interest and the search for evidence of primordial non-Gaussianity requires that false detections due to foreground residuals be eliminated.  $\text{IC}^2$  can play a role in the initial cleaning of the data, and then in defining regions on the sky that merit special treatment, be it simply masking them from future cosmological analysis due to the apparently complex nature of the foregrounds there, or the re-application of the component separation techniques on the regions to improve the final CMB sky estimate.

### 3.9 Discussion

We have utilised a blind component separation algorithm –  $\text{IC}^2$  – to address the issue of the Galactic foreground emission in the *WMAP* 3-year data. In particular, we have adapted the technique to allow three Galactic foreground templates (that trace synchrotron, free-free and thermal dust emission) to be fitted to the data on a frequency-by-frequency basis. Specifically, we use the Finkbeiner (2003)  $H\alpha$ -map as a template for the free-free emission, the Finkbeiner et al. (1999) FDS8 model for thermal dust emission, and for synchrotron emission two alternatives – the 408 MHz radio continuum all-sky map of Haslam et al. (1982) as utilised in the first year *WMAP* analysis, or the difference of the *WMAP* K- and Ka-band data preferred in their three year analysis.

Detailed simulations indicate that the algorithm behaves in an unbiased way only for two ( $p, g$ ) of the three non-linear functions typically applied by  $\text{IC}^2$ . In these cases, the  $p$ -function also appears to be marginally preferred as a consequence of its smaller error bars, both relative to the  $g$ -function analysis and the simple  $\chi^2$  method that is usually adopted for template fitting.

When applied to the *WMAP* data, we have quantified the foreground contamination in term of coupling coefficients between the data and the foreground templates. The derived values are sensitive to both the extension of the mask applied to exclude the Galactic plane and on the non-linear function used by the algorithm. The mask dependence is likely to reflect genuine variations in the spectral behaviour of the foregrounds with sky location, as seen in previous local studies of template fits by Davies et al. (2006). The non-linear function dependence is plausibly linked to sensitivity to the actual statistical nature of the foreground components (as represented by the templates), and suggests that it may be useful to study more appropriate approximations to the neg-entropy.

We have considered the spectral behaviour of the derived scaling factors when the Haslam et al. (1982) data is used as the synchrotron template. We evaluated the spectral index for the synchrotron emission, the anomalous dust-correlated component, and the free-free emission. In the first two cases, we found steeper, though statistically consistent, spectral behaviour as compared to previous analysis, e.g. Davies et al. (2006). For the free-free emission, we confirmed previous observations about

inconsistencies in the fitted scaling to convert the  $H\alpha$  emissivity to a free-free contribution at K- and Ka-bands. Specifically, the derived factors correspond to thermal electron temperatures of  $\sim 4000 K$  compared to the expected value of  $8000 K$ .

In addition, we find independent verification of a flat spectral dependence in the free-free emissivity as reported by Dobler & Finkbeiner (2008a). However, this behaviour is seen only for the  $Kp2$  sky coverage. With the  $Kp0$  mask, the behaviour is even steeper than the expected  $\nu^{-2.14}$  dependence, suggesting that the anomalous free-free behaviour is associated with bright structures close the Galactic plane.

We have also proposed that the  $\chi^2$  algorithm can be applied ‘iteratively’, that is, used to analyse a set of multi-frequency maps pre-cleaned using templates for which the coefficients have themselves been derived by a  $\chi^2$  analysis. This is particularly useful in order to determine the presence of residual foregrounds that arise either due to a mismatch between their spectral dependence and the average high-latitude value determined by the template fits, or because they are new components that are not traced by the adopted templates. In this manner, we have confirmed the existence of a component spatially distributed along the Galactic plane, with pronounced emission near the Galactic center. This is the emission previously noted in the SMICA analysis of Patanchon et al. (2005), and is the ‘*WMAP* haze’ of Finkbeiner (2004a). However, the amplitude and extent of the emission is less when the K-Ka map is adopted as the synchrotron template. This is not surprising if the foreground component is a genuinely new contribution to the sky emission that is untraced by the Haslam data. In this context, the K-Ka map must be considered the better template for cleaning the *WMAP* data from foregrounds. However, the interpretation of the results in terms of Galactic emission components is complicated by the fact that the K-Ka template contains a mixture of several, physically distinct, emission mechanisms.

Finally, the analysis that we have performed with the  $\chi^2$  algorithm is based on the unrealistic hypothesis that the spectral behaviour of the various foreground emission components is unchanging over the sky. In Chapter 6, in order to take into account the spatial variation of the foreground spectra, we have attempted to apply  $\chi^2$  on smaller patches of the sky where the assumption of uniform spectral behaviour is more realistic.



## Chapter 4

# New insights into foreground analysis of the *WMAP* five-year data using

*Note: the work presented in this chapter was done in collaboration with A.J. Banday and D.Maino and was also published as Bottino et al. (2010). All the computations have been performed by M. Bottino.*

### 4.1 Introduction

The release of five years of observations from the *WMAP* satellite allows a quantitative improvement in studies of the CMB for cosmological purposes. However, such an improvement requires an analogous refinement in both our understanding of local astrophysical foregrounds and in the methods employed for the separation of these components from the CMB emission.

The aim of the work presented in this chapter is to undertake the foreground analysis of the *WMAP* 5-year data (Hinshaw et al., 2009) with the  $\text{IC}^2$  algorithm, as previously applied to the *WMAP* 3-year data (see Chapter 3). As before, we focus on the diffuse Galactic foreground components and their spectral and spatial properties.

However, novel aspects of the analysis are now introduced:

- consideration of the new KQ85 and KQ75 masks introduced by the *WMAP* science team with the 5-year data release
- an investigation into the connection between the mask applied to derive the coupling coefficients and the residuals resulting from the iterative multi-frequency analysis when these coefficients are applied, itself as a function of applied mask
- an attempt to understand the physical nature of the residuals revealed by the iterative analysis.
- the definition of a minimal mask that supports accurate foreground removal

As previously done, we compare our analysis to the simpler  $\chi^2$ -based template fitting scheme, commonly used in the field, as a convenient point of reference for our results (see Chapter 3 for details).

To be consistent with the analysis, we did not impose any constraints on the data, although this is possible as described in Hinshaw et al. (2007) and Gold et al. (2009).

## 4.2 Data used in the analysis

Our primary data set consists of the *WMAP* 5-year maps improved in sensitivity and precision as described in Hinshaw et al. (2009).

As in the analysis proposed in Chapter 3, we assume that the data can be well described by a superposition of the CMB anisotropy plus foreground emission traced by three templates. These templates, suitably scaled to each frequency, are considered to represent four different physical components: synchrotron, free-free and thermal dust emission, together with the anomalous dust component – a dust correlated emission whose detailed nature is not well known.

The templates are based on maps produced from independent observations of the sky at frequencies where only one emission mechanism dominates. In practice, we retain the set of templates used for the *WMAP* 3-year analysis (see Chapter 3), in order to be able to relate all changes in the analysis results to the improvements in the *WMAP* data. Therefore, we used the 408 MHz radio continuum all-sky map of Haslam et al. (1982) and the difference between the K- and Ka-band data of *WMAP* (Hinshaw et al., 2007) as synchrotron templates. The all-sky  $H\alpha$ -map, produced by Finkbeiner (2003) by assembling data from several surveys, is the template utilised for the free-free emission. The map was not corrected for dust absorption, assuming it to be negligible as suggested by Banday et al. (2003). Nevertheless, we tested the dependence of our results to this correction applied on the template. Moreover, we have also tested an alternative  $H\alpha$ -map provided by Dickinson et al. (2003). This has demonstrated some differences in correlation properties with the *WMAP* data in previous analyses (Davies et al., 2006), but we do not find any sensitivity to the choice of  $H\alpha$  template in our work. Finally, we adopted the model for thermal dust emission at 94 GHz developed by Finkbeiner et al. (1999) (FDS8), which has been also used for the anomalous dust emission. More detailed descriptions of these foregrounds and the corresponding templates may be found in Chapter 1. However, we note that the synchrotron emission (in antenna temperature) is generally described by a power-law spectrum  $T_s(\nu) \sim \nu^{-\beta_s}$ , similarly the free-free by  $T_{ff} \sim \nu^{-\beta_{ff}}$ , and the thermal dust approximated as  $T_d \sim \nu^{\beta_d}$  over the *WMAP* range of frequencies.

The templates are used to fit the spectral behaviour of the *WMAP* data at each frequency of observation from  $\sim 23$  GHz (K-band) up to  $\sim 94$  GHz (W-band). Where multiple samples are available at a given frequency, we have coadded them into a single sky map using uniform and equal weights. As before, we performed our analysis on sky maps convolved from their original resolution to an effective  $1^\circ$  Gaussian beam. Finally, we converted the *WMAP* data from thermodynamic temperature to brightness (antenna) temperature units, which is more natural for an analysis of foreground spectral behaviour.

Since regions close to the Galactic plane are the most seriously contaminated by foregrounds and the spectral and spatial nature of the integrated emission is complex, they should generally be excluded

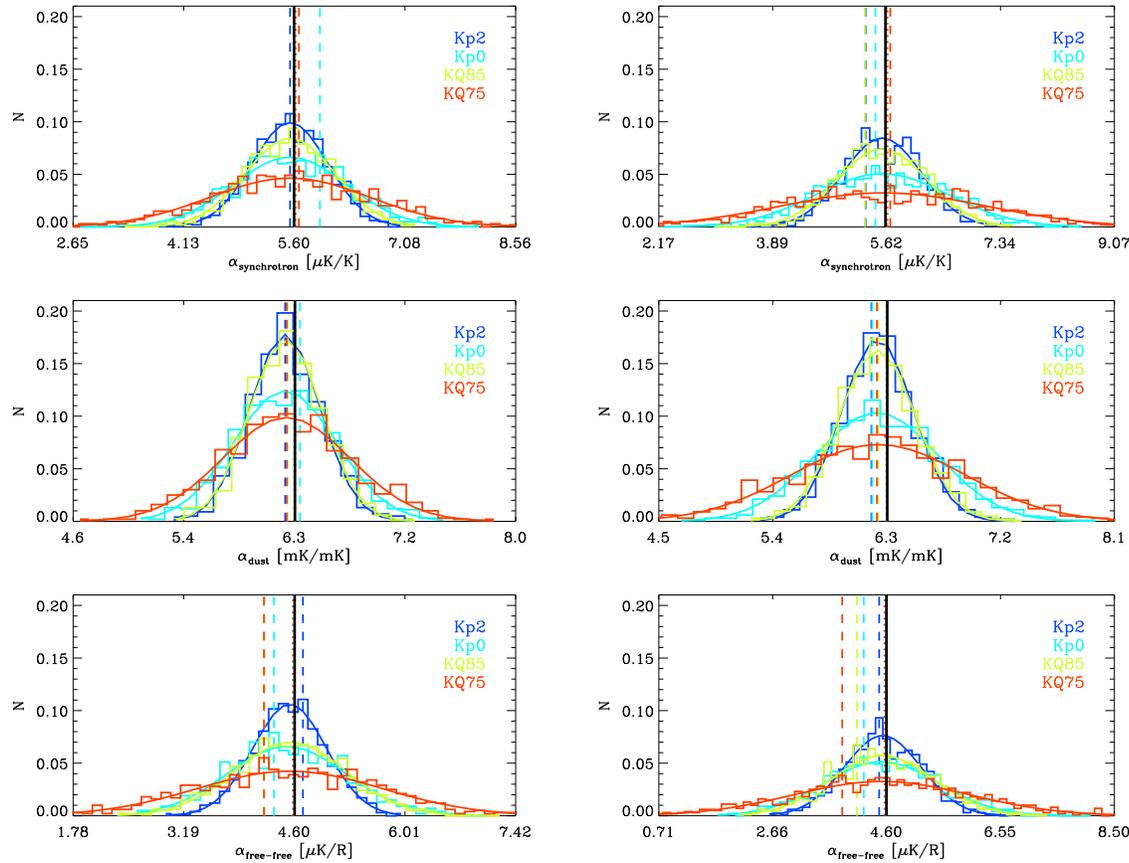


Figure 4.1: Histogram of the coupling coefficient distributions at K-band obtained from the simulations using the  $p$  (left) and  $g$  (right) functions and all the masks. Different color are associated to different masks in order to compare the results: blue for  $Kp2$ , cyan for  $Kp0$ , green for  $KQ85$  and red for  $KQ75$ . The dotted and dashed line show respectively the mean and the mode of the distributions, to be compared with the expected values (black line). For all the Galactic components except the dust emission, there is a strong dependence of the distribution width on the extension of the plane cut. This is particularly visible for the free-free emission where the scaling factors obtained with the  $KQ75$  mask are characterized by a very flat distribution.

from analysis. Indeed the *WMAP* science team itself has proposed two new different masks than the previously used  $Kp2$  and  $Kp0$  cuts. They refer to the new masks as  $KQ85$  and  $KQ75$  since they are produced from the maps at the K and Q band and they exclude respectively 15% and 25% of the sky (Hinshaw et al., 2009). We employed all 4 masks together with a mask for the point sources, without any modification of the angular resolution. However, in the latter part of the chapter we also experiment with a full-sky analysis as shown in Section 4.8.

### 4.3 Monte Carlo simulations

Monte Carlo simulations are used in this context for several reasons. First, they provide useful criteria and figures-of-merit that aid the selection of the non-linear function ( $p$ ,  $g$  or  $t$ ) used by the

algorithm to implement the component separation. Second, they provide a method to evaluate the level of correlation with and among the foregrounds. Finally we used these simulations to compute the uncertainties for the scaling factors presented in later sections of the chapter.

We find that essentially all of the previous results from the 3-year analysis are confirmed. The reconstruction of the input coefficients is satisfactory for both the  $p$  and  $g$  functions, which provide similar results. Indeed, the distribution of the coefficients is a symmetric Gaussian function in both cases: the mean value is very close to the expected one and the mode differs only slightly. However, the  $g$  function shows generally a broader distribution compared to that for  $p$ , most likely as a consequence of the fact that they are sensitive to different statistical features of the spatial distribution of the sky radiation. For the same reason the  $t$ -function remains inappropriate for template fitting: the statistical distributions of the returned coupling coefficients are asymmetric and highly biased with respect to the input values.

The distribution of the scaling factors generally depends on the mask used in the analysis (see Figure 4.1): the more extensive the cut is, the larger is the dispersion of the values around the mean. It is particularly evident for the free-free emission where the  $KQ75$  mask is associated with the largest distribution, though it is still symmetric and not biased. We also find that the anti-correlation between the synchrotron and dust coefficients decreases for larger masks, as is the case for the correlation between the K-Ka template and the free-free.

The results from the simple  $\chi^2$  analysis indicate that the uncertainties for the synchrotron component are smaller than found using the                    method, while they are larger for the other components. This method dependent cross-talk likely reflects the different statistical measures of the data employed by the                    and  $\chi^2$  analyses.

#### 4.4 Coupling coefficients between data and templates

As noted above, the                    component separation allows us to compute coupling coefficients between the data at each frequency of observation, and the three templates of the Galactic foreground components. We calculated these values using the  $p$  and the  $g$  non-linear functions and the four masks provided by the *WMAP* science team. Then, we compared the results with those from the simple  $\chi^2$  analysis. In Tables 4.1 and 4.2, we report the results obtained with two different sets of templates, where either the Haslam or the K-Ka map was used as synchrotron model. In the same Tables, we show also the results from Bennett et al. (2003), Hinshaw et al. (2007) and Gold et al. (2009): these values are the *WMAP* results respectively for the 1-, 3- and 5-year foreground analysis corrections, with the  $Kp2$  or  $KQ85$  masks. In all three cases, some constraints were placed on the spectral behaviour of the components. In the first case the values were derived by fixing the synchrotron spectral index at a value of 2.7, and the free-free to 2.15 for frequencies above Q-band. In the latter cases, instead, they explicitly take into account the free-free signal present in the K-Ka template, and impose constraints on the free-free ( $\beta_{ff} = 2.14$ ) and thermal dust ( $\beta_d = 2$ ) spectral indices. However, we did not consider any constraints on the coefficients.

	synchrotron		dust		free-free	
	$Kp2$	$Kp0$	$Kp2$	$Kp0$	$Kp2$	$Kp0$
<b>- function p</b>						
$K$	$6.20 \pm 0.49$	$5.47 \pm 0.75$	$5.88 \pm 0.29$	$4.96 \pm 0.40$	$8.57 \pm 0.47$	$6.66 \pm 0.75$
$Ka$	$1.91 \pm 0.49$	$1.70 \pm 0.74$	$2.04 \pm 0.28$	$1.29 \pm 0.38$	$4.35 \pm 0.46$	$2.74 \pm 0.74$
$Q$	$0.96 \pm 0.48$	$0.87 \pm 0.74$	$1.09 \pm 0.28$	$0.44 \pm 0.37$	$2.91 \pm 0.46$	$1.34 \pm 0.73$
$V$	$0.26 \pm 0.46$	$0.24 \pm 0.70$	$0.61 \pm 0.26$	$0.05 \pm 0.35$	$1.34 \pm 0.43$	$-0.07 \pm 0.70$
$W$	$0.06 \pm 0.40$	$0.07 \pm 0.61$	$0.97 \pm 0.23$	$0.46 \pm 0.31$	$0.64 \pm 0.38$	$-0.52 \pm 0.61$
<b>- function g</b>						
$K$	$6.42 \pm 0.58$	$5.63 \pm 0.95$	$6.14 \pm 0.30$	$5.76 \pm 0.48$	$8.48 \pm 0.65$	$6.52 \pm 0.97$
$Ka$	$2.04 \pm 0.56$	$1.91 \pm 0.94$	$2.10 \pm 0.29$	$1.68 \pm 0.47$	$4.26 \pm 0.64$	$2.80 \pm 0.95$
$Q$	$1.10 \pm 0.56$	$1.12 \pm 0.92$	$1.11 \pm 0.29$	$0.71 \pm 0.47$	$2.79 \pm 0.63$	$1.45 \pm 0.94$
$V$	$0.40 \pm 0.53$	$0.54 \pm 0.87$	$0.60 \pm 0.27$	$0.24 \pm 0.44$	$1.21 \pm 0.60$	$0.03 \pm 0.89$
$W$	$0.19 \pm 0.46$	$0.33 \pm 0.77$	$0.95 \pm 0.24$	$0.61 \pm 0.39$	$0.56 \pm 0.53$	$-0.41 \pm 0.78$
$\chi^2$ analysis						
$K$	$5.98 \pm 0.42$	$5.42 \pm 0.57$	$6.44 \pm 0.31$	$6.34 \pm 0.55$	$7.93 \pm 0.74$	$5.68 \pm 1.20$
$Ka$	$1.84 \pm 0.42$	$1.62 \pm 0.56$	$2.26 \pm 0.30$	$2.12 \pm 0.54$	$3.77 \pm 0.73$	$2.11 \pm 1.18$
$Q$	$0.92 \pm 0.41$	$0.79 \pm 0.56$	$1.24 \pm 0.30$	$1.12 \pm 0.53$	$2.31 \pm 0.72$	$0.80 \pm 1.17$
$V$	$0.24 \pm 0.39$	$0.17 \pm 0.53$	$0.71 \pm 0.28$	$0.62 \pm 0.51$	$0.75 \pm 0.68$	$-0.59 \pm 1.11$
$W$	$0.04 \pm 0.34$	$-0.01 \pm 0.46$	$1.04 \pm 0.25$	$0.94 \pm 0.45$	$0.17 \pm 0.60$	$-0.95 \pm 0.97$
<b>Bennett et al. (constrained: <math>\beta_s = 2.7; \beta_{ff} = 2.15</math>)</b>						
$Q$	1.01		1.04		(1.92)	
$V$	0.34		0.62		(0.82)	
$W$	0.11		0.87		(0.32)	
	$KQ85$	$KQ75$	$KQ85$	$KQ75$	$KQ85$	$KQ75$
<b>- function p</b>						
$K$	$6.57 \pm 0.58$	$6.05 \pm 1.09$	$5.62 \pm 0.29$	$4.78 \pm 0.51$	$8.32 \pm 0.72$	$8.14 \pm 1.15$
$Ka$	$2.28 \pm 0.57$	$2.58 \pm 1.08$	$2.04 \pm 0.28$	$1.66 \pm 0.51$	$4.47 \pm 0.71$	$4.51 \pm 1.14$
$Q$	$1.36 \pm 0.56$	$1.85 \pm 1.06$	$1.17 \pm 0.28$	$0.95 \pm 0.50$	$3.12 \pm 0.70$	$3.23 \pm 1.12$
$V$	$0.64 \pm 0.54$	$1.29 \pm 1.01$	$0.76 \pm 0.27$	$0.67 \pm 0.47$	$1.66 \pm 0.67$	$1.92 \pm 1.06$
$W$	$0.38 \pm 0.47$	$0.94 \pm 0.89$	$1.11 \pm 0.24$	$1.01 \pm 0.42$	$0.95 \pm 0.59$	$1.16 \pm 0.93$
<b>- function g</b>						
$K$	$6.45 \pm 0.65$	$6.34 \pm 1.47$	$6.00 \pm 0.31$	$6.37 \pm 0.69$	$7.72 \pm 0.86$	$6.18 \pm 1.56$
$Ka$	$2.25 \pm 0.64$	$3.02 \pm 1.45$	$2.09 \pm 0.31$	$2.39 \pm 0.68$	$3.85 \pm 0.85$	$2.93 \pm 1.53$
$Q$	$1.33 \pm 0.63$	$2.31 \pm 1.43$	$1.16 \pm 0.30$	$1.47 \pm 0.67$	$2.49 \pm 0.84$	$1.73 \pm 1.51$
$V$	$0.65 \pm 0.60$	$1.76 \pm 1.35$	$0.69 \pm 0.29$	$1.05 \pm 0.63$	$1.03 \pm 0.80$	$0.53 \pm 1.44$
$W$	$0.40 \pm 0.53$	$1.34 \pm 1.19$	$1.03 \pm 0.25$	$1.29 \pm 0.56$	$0.44 \pm 0.70$	$-0.01 \pm 1.26$
$\chi^2$ analysis						
$K$	$5.91 \pm 0.46$	$5.33 \pm 0.70$	$6.36 \pm 0.34$	$6.67 \pm 0.75$	$7.06 \pm 1.00$	$4.99 \pm 1.85$
$Ka$	$1.86 \pm 0.45$	$1.63 \pm 0.69$	$2.29 \pm 0.33$	$2.47 \pm 0.73$	$3.25 \pm 0.99$	$1.88 \pm 1.82$
$Q$	$0.97 \pm 0.45$	$0.83 \pm 0.68$	$1.32 \pm 0.33$	$1.49 \pm 0.72$	$1.90 \pm 0.98$	$0.73 \pm 1.79$
$V$	$0.30 \pm 0.42$	$0.22 \pm 0.64$	$0.82 \pm 0.31$	$1.01 \pm 0.69$	$0.45 \pm 0.93$	$-0.48 \pm 1.70$
$W$	$0.09 \pm 0.37$	$0.03 \pm 0.57$	$1.13 \pm 0.28$	$1.27 \pm 0.60$	$-0.06 \pm 0.81$	$-0.88 \pm 1.50$

Table 4.1: Values of the coupling coefficients in antenna temperature units determined between the 5-year *WMAP* data and three foreground emission templates, at  $1^\circ$  resolution. The Haslam 408 MHz map is adopted as the synchrotron template. The analysis is performed using the two non-linear functions  $p$  and  $g$  and four masks of the Galactic plane. The corresponding results for a simple  $\chi^2$  analysis are provided for comparison. In addition, we provide the values from the *WMAP* 1-year (Bennett et al., 2003) fits to the Q, V and W-bands performed with constraints imposed on the synchrotron and free-free spectral indices. The units are  $\mu\text{K}/\text{K}$  for synchrotron,  $\text{mK}/\text{mK}$  for dust and  $\mu\text{K}/\text{R}$  for free-free emission respectively.

	synchrotron		dust		free-free	
	$Kp2$	$Kp0$	$Kp2$	$Kp0$	$Kp2$	$Kp0$
<b>- function p</b>						
$Q$	$0.18 \pm 0.08$	$0.15 \pm 0.12$	$0.40 \pm 0.40$	$0.12 \pm 0.58$	$2.16 \pm 0.60$	$0.67 \pm 0.89$
$V$	$-0.00 \pm 0.08$	$-0.03 \pm 0.12$	$0.63 \pm 0.38$	$0.38 \pm 0.55$	$1.33 \pm 0.57$	$-0.05 \pm 0.84$
$W$	$-0.05 \pm 0.07$	$-0.07 \pm 0.10$	$1.16 \pm 0.34$	$0.92 \pm 0.48$	$0.82 \pm 0.50$	$-0.32 \pm 0.74$
<b>- function g</b>						
$Q$	$0.25 \pm 0.10$	$0.16 \pm 0.17$	$0.15 \pm 0.47$	$0.13 \pm 0.80$	$1.76 \pm 0.79$	$0.82 \pm 1.17$
$V$	$0.06 \pm 0.09$	$-0.02 \pm 0.16$	$0.41 \pm 0.45$	$0.41 \pm 0.76$	$0.94 \pm 0.75$	$0.05 \pm 1.11$
$W$	$0.01 \pm 0.08$	$-0.06 \pm 0.14$	$0.96 \pm 0.40$	$0.93 \pm 0.67$	$0.50 \pm 0.66$	$-0.24 \pm 0.98$
<b><math>\chi^2</math> analysis</b>						
$Q$	$0.22 \pm 0.09$	$0.20 \pm 0.12$	$0.37 \pm 0.41$	$0.33 \pm 0.68$	$1.43 \pm 0.85$	$0.11 \pm 1.21$
$V$	$0.03 \pm 0.08$	$0.01 \pm 0.11$	$0.59 \pm 0.39$	$0.53 \pm 0.64$	$0.76 \pm 0.81$	$-0.47 \pm 1.15$
$W$	$-0.01 \pm 0.07$	$-0.02 \pm 0.10$	$1.10 \pm 0.34$	$1.04 \pm 0.56$	$0.29 \pm 0.71$	$-0.81 \pm 1.01$
<b>Hinshaw et al. (constrained: <math>\beta_{ff} = 2.14</math>; <math>\beta_s = 2</math>)</b>						
$Q$	0.23		0.19		0.99	
$V$	0.05		0.41		0.63	
$W$	0.00		0.98		0.32	
	$KQ85$	$KQ75$	$KQ85$	$KQ75$	$KQ85$	$KQ75$
	<b>- function p</b>					
$Q$	$0.28 \pm 0.10$	$0.44 \pm 0.13$	$0.27 \pm 0.42$	$-0.29 \pm 0.66$	$2.03 \pm 0.84$	$1.54 \pm 1.16$
$V$	$0.10 \pm 0.09$	$0.25 \pm 0.13$	$0.51 \pm 0.40$	$0.03 \pm 0.62$	$1.25 \pm 0.79$	$0.89 \pm 1.09$
$W$	$0.04 \pm 0.08$	$0.15 \pm 0.11$	$1.08 \pm 0.35$	$0.66 \pm 0.55$	$0.78 \pm 0.70$	$0.48 \pm 0.96$
<b>- function g</b>						
$Q$	$0.30 \pm 0.12$	$0.35 \pm 0.36$	$0.06 \pm 0.53$	$0.11 \pm 1.56$	$1.34 \pm 0.99$	$0.39 \pm 1.59$
$V$	$0.12 \pm 0.11$	$0.16 \pm 0.36$	$0.31 \pm 0.50$	$0.43 \pm 1.10$	$0.55 \pm 0.94$	$-0.24 \pm 1.51$
$W$	$0.06 \pm 0.10$	$0.07 \pm 0.31$	$0.87 \pm 0.44$	$1.04 \pm 0.97$	$0.19 \pm 0.83$	$-0.45 \pm 1.33$
<b><math>\chi^2</math> analysis</b>						
$Q$	$0.23 \pm 0.10$	$0.21 \pm 0.14$	$0.43 \pm 0.44$	$0.66 \pm 0.90$	$1.05 \pm 1.06$	$0.07 \pm 1.73$
$V$	$0.04 \pm 0.09$	$0.03 \pm 0.13$	$0.65 \pm 0.42$	$0.84 \pm 0.86$	$0.46 \pm 1.00$	$-0.35 \pm 1.64$
$W$	$-0.00 \pm 0.08$	$-0.01 \pm 0.11$	$1.14 \pm 0.37$	$1.29 \pm 0.75$	$0.00 \pm 0.88$	$-0.79 \pm 1.44$
<b>Gold et al. (constrained: <math>\beta_{ff} = 2.14</math>; <math>\beta_s = 2</math>)</b>						
$Q$	0.23		0.19		0.95	
$V$	0.05		0.43		0.60	
$W$	0.00		1.01		0.32	

Table 4.2: Values of the coupling coefficients in antenna temperature units determined between the 5-year *WMAP* data in the Q, V and W band and three foreground emission templates, at  $1^\circ$  resolution. The K-Ka map is adopted as the synchrotron template. The  $\chi^2$  analysis is performed using the two non-linear functions  $p$  and  $g$  and four masks of the Galactic plane. The corresponding results for a simple  $\chi^2$  analysis are provided for comparison. In addition, we provide where appropriate the values from the *WMAP* 3-year (Hinshaw et al., 2007) and 5-year (Gold et al., 2009) fits to the Q,V and W-bands performed with constraints imposed on the synchrotron and free-free spectral indices. The units are mK/mK for synchrotron, mK/mK for dust and  $\mu\text{K/R}$  for free-free emission respectively.

We note that the results obtained using the  $Kp2$  and  $Kp0$  masks are very consistent with those from our three-year analysis, with values varying at most by  $0.5 \sigma$  with the previous ones. The frequency behavior is consistent with the theoretical expectations. The synchrotron and free-free emissions decrease with increasing frequency, whereas the dust coefficients are consistent with the superposition of an anomalous component with a falling contribution until approximately 61 GHz where a rising contribution from thermal emission is seen.

However, consideration of the results derived with the two new masks indicates that it is no longer possible to define a trend: a larger mask does not necessarily imply lower scaling factors, as was the case for the free-free coefficients obtained with the  $Kp2$  and the  $Kp0$  masks. On inspection of the results obtained with the Haslam map as synchrotron template, although the  $KQ75$  mask is the largest cut provided, we generally obtain values comparable to or higher than those returned using the  $Kp2$  and  $KQ85$  masks. It is certainly a confirmation of the fact that there exist spectral or physical variations of the foregrounds on the sky. These are particularly significant for the free-free emission and are presumably connected to specific regions near the Galactic plane, which are completely removed by the  $KQ75$  mask, but retained by the  $Kp2$  and  $KQ85$  cuts.

We studied the dependence of the analysis on the correction applied to the  $H\alpha$  map to correct for the dust absorption. The coupling coefficients show differences which are not very significant, but consistent with expectation, i.e. as the absorption correction raises the  $H\alpha$  intensity, so the coefficients decrease. What is interesting then, is that the inconsistency between the values derived with the  $Kp0$  cut compared to the other masks still remains, implying that it is not easily associated with variations in the dust absorption. Thus it could be related to variations in the temperature of the ionized gas in the medium latitude regions. In fact, using the free-free coefficients we infer values of the electron temperature that change with the mask and that are generally lower than the expected value of 8000 K. This is true in particular for the  $Kp0$  mask.

The mask dependence is different for the K-Ka analysis. Each of the foreground components shows a specific trend with the cuts. In particular, the dust and free-free coefficients obtained with the  $KQ75$  mask are now lower than those with the  $Kp2$  and  $KQ85$  masks. However, this is not surprising, since the K-Ka map is a mix of different emissions. It has contributions from synchrotron and free-free, as well as from the anomalous component: none of which are present in the Haslam map (at high latitudes – there is a small contribution from free-free emission in the Galactic plane). Moreover, this can also be a possible explanation for the amplitudes of the dust coupling coefficients in the Q-band which are low compared to the Haslam case. Specifically, the anomalous component is now traced largely by the K-Ka template, rather than the dust. Indeed, this may also suggest that the thermal dust template is not morphologically identical to the anomalous dust emission, although well-correlated.

We have also used a simple  $\chi^2$  analysis as a standard method to be compared with . Whereas with the *WMAP* 3-year data we considered only the results for the  $Kp2$  mask, here we extend the comparison to all of the sky coverages. This choice was motivated by the unusual dependence on the extension of the mask observed in our analysis. With the Haslam map as synchrotron template, the anomaly is confirmed by the  $\chi^2$  analysis. In fact the numbers are generally in good statistical agreement with the  $\chi^2$  results, even though the synchrotron and dust amplitudes are systematically

lower (higher) for the  $\chi^2$  method. However, this in general reflects the weak cross-talk seen between the fitted amplitudes. Alternately, when the K-Ka map is adopted instead, the  $\chi^2$  coupling coefficients of the synchrotron and dust emission are stable with respect to the mask used. Conversely, the free-free coefficients are found to be strongly dependent on the cut. Moreover, the values obtained with the *Kp0* and *KQ75* masks are consistent with zero. Nevertheless, all the coefficients are consistent with the results.

Finally, we compared our results with those obtained by Bennett et al. (2003), Hinshaw et al. (2007) and Gold et al. (2009) with constrained fits. The values are comparable for the synchrotron and dust emissions, with both the synchrotron templates. The free-free values, however, are notably different in all the cases, in part due to the fact that they corrected the  $H\alpha$  map for dust absorption. However, even if we repeat our analysis using the *WMAP* value for the dust absorption correction, the coefficients remain larger than *WMAP* derive. Therefore, it cannot be the only explanation for the difference.

#### 4.4.1 Spectral index of foreground emissions

We parameterise the spectral behavior of the synchrotron emission (as traced by the Haslam template), the anomalous dust component, and free-free emission with particular emphasis on the latter. The *WMAP* frequency range does not allow a detailed study of the spectral behaviour of the thermal component of dust. We fitted the coefficients of each component with a power law model of the form  $A_{norm}(\nu/\nu_0)^{-\beta}$ .  $A_{norm}$  is the amplitude of the emission of a specific physical component at the reference frequency  $\nu_0$ , which we take as the K-band (23 GHz). In the case of the anomalous component, in order to isolate its contribution we recomputed the coefficients from the sky maps after correcting them for a thermal dust contribution assuming the FDS8 dust model. This assumption is consistent with the W-band correlation results in Table 4.1.

Since, the *Kp2* and *Kp0* results do not show any significant difference with respect to those computed with the *WMAP* 3-year data, we do not expect the spectral index to change and this is indeed seen to be the case in Table 4.3. For the synchrotron emission, using the *Kp2* mask the spectral index is steeper than  $\beta_s = 3.0$ , even though still consistent, while it is flatter for the *Kp0* mask, particularly when using the  $g$ -function. The values of the spectral index  $\beta_a$  of the anomalous dust component are larger than 3 especially if the *Kp0* mask is adopted. Therefore, they are steeper than the value of 2.85 obtained by Davies et al. (2006), although higher values of the spectral index are expected if we consider regions of the sky at mid- to high-latitude. In fact, Davies et al. (2006) have also noted spectral indices as steep as 3.8 in several dust dominated regions.

However, the values obtained with the new masks highlight some unexpected physical properties of the foreground components. In fact, with the *KQ85* and *KQ75* masks, the synchrotron and anomalous dust spectral indices are flatter than those derived from the older *Kp2* and *Kp0* masks, and in some cases considerably so. Indeed, whilst all values are consistent within the errors (derived from simulations), we are unable to explain these differences given that the new masks can be considered as small modifications to the older ones. The major difference relates to the omission now of various features associated with the free-free emission. That such apparently small changes can affect the

	Synchrotron		Anomalous Component		Free-Free	
	– function p					
	<i>Kp2</i>	<i>Kp0</i>	<i>Kp2</i>	<i>Kp0</i>	<i>Kp2</i>	<i>Kp0</i>
$\beta$	$3.25^{+1.18}_{-0.49}$	$3.21^{+1.23}_{-0.54}$	$3.32^{+0.48}_{-0.26}$	$4.72^{+0.31}_{-0.36}$	$1.88^{+0.62}_{-0.50}$	$2.89^{+1.28}_{-0.38}$
$A_{norm}$	$6.20 \pm 0.47$	$5.46 \pm 0.72$	$5.79 \pm 0.20$	$4.87 \pm 0.28$	$8.57 \pm 0.44$	$6.76 \pm 0.69$
	<i>KQ85</i>	<i>KQ75</i>	<i>KQ85</i>	<i>KQ75</i>	<i>KQ85</i>	<i>KQ75</i>
$\beta$	$2.70^{+1.11}_{-0.46}$	$1.80^{+0.90}_{-0.75}$	$3.09^{+0.39}_{-0.33}$	$3.29^{+0.25}_{-0.50}$	$1.65^{+0.23}_{-0.58}$	$1.51^{+0.41}_{-0.41}$
$A_{norm}$	$6.52 \pm 0.56$	$5.84 \pm 1.02$	$5.52 \pm 0.20$	$4.67 \pm 0.37$	$8.27 \pm 0.66$	$8.05 \pm 0.91$
	– function g					
	<i>Kp2</i>	<i>Kp0</i>	<i>Kp2</i>	<i>Kp0</i>	<i>Kp2</i>	<i>Kp0</i>
$\beta$	$3.06^{+1.35}_{-0.56}$	$2.76^{+1.07}_{-0.75}$	$3.35^{+0.42}_{-0.31}$	$4.09^{+0.54}_{-0.38}$	$1.94^{+0.31}_{-0.52}$	$2.74^{+1.28}_{-0.24}$
$A_{norm}$	$6.41 \pm 0.55$	$5.59 \pm 0.90$	$6.05 \pm 0.21$	$5.68 \pm 0.35$	$8.50 \pm 0.60$	$6.62 \pm 1.09$
	<i>KQ85</i>	<i>KQ75</i>	<i>KQ85</i>	<i>KQ75</i>	<i>KQ85</i>	<i>KQ75</i>
$\beta$	$2.69^{+1.11}_{-0.46}$	$1.41^{+0.90}_{-0.75}$	$3.23^{+0.45}_{-0.28}$	$2.79^{+0.71}_{-0.16}$	$1.99^{+0.30}_{-0.50}$	$2.28^{+0.72}_{-0.13}$
$A_{norm}$	$6.40 \pm 0.63$	$6.00 \pm 1.34$	$5.90 \pm 0.22$	$6.23 \pm 0.51$	$7.75 \pm 0.81$	$6.25 \pm 1.03$
	$\chi^2$					
	<i>Kp2</i>	<i>Kp0</i>	<i>Kp2</i>	<i>Kp0</i>	<i>Kp2</i>	<i>Kp0</i>
$\beta$	$3.26^{+1.39}_{-0.48}$	$3.37^{+1.23}_{-0.54}$	$3.20^{+0.36}_{-0.38}$	$3.39^{+0.34}_{-0.32}$	$2.20^{+0.66}_{-0.55}$	$3.44^{+0.72}_{-0.24}$
$A_{norm}$	$5.98 \pm 0.41$	$5.43 \pm 0.54$	$6.34 \pm 0.22$	$6.25 \pm 0.40$	$7.99 \pm 0.68$	$5.78 \pm 1.01$
	<i>KQ85</i>	<i>KQ75</i>	<i>KQ85</i>	<i>KQ75</i>	<i>KQ85</i>	<i>KQ75</i>
$\beta$	$3.14^{+1.11}_{-0.46}$	$3.26^{+0.86}_{-0.86}$	$3.06^{+0.30}_{-0.45}$	$2.86^{+0.71}_{-0.16}$	$2.38^{+0.30}_{-0.50}$	$-3.39^{+0.41}_{-0.41}$
$A_{norm}$	$5.90 \pm 0.44$	$5.33 \pm 0.61$	$6.26 \pm 0.24$	$6.55 \pm 0.55$	$7.14 \pm 0.89$	$5.07 \pm 1.36$

Table 4.3: Spectral index  $\beta$  and normalisation factor  $A_{norm}$  obtained fitting values of the coupling coefficients for synchrotron (as traced by the Haslam template), the anomalous component of dust and free-free emission, with different masks. The normalization factor  $A_{norm}$  has units equal to mK/mK for synchrotron, mK/mK for dust and  $\mu\text{K}/\text{R}$  for free-free emission respectively.

analysis to such an extent might be considered problematic, although it may ultimately be telling us something about the properties of the foregrounds close to the mask boundaries.

With regards to the free-free emission, the previous analysis in Chapter 3 found suggestions of anomalous behaviour that remain to be explained. Our fits for all of the sky coverages are statistically consistent with the expected scaling for free-free emission with frequency,  $\beta_{ff} = 2.14$ . Nonetheless, depending on the extension of the cut the spectral behaviour of the coefficients show different trends, as already pointed out in the earlier work. Comparing the spectral behaviour recovered with the different Galactic cuts, the index is always flatter than the canonical value of 2.14 except with the *Kp0* mask which is steeper. Previously, we had interpreted this result as the consequence of different properties of the free-free emission close to the plane (perhaps a mixture of regions with different electron temperatures), the presence of spinning dust in the Warm Ionized Medium (WIM), or simply cross-talk between different physical components that confuses the spectral analysis.

Here, we consider the second option in more detail by adopting a model proposed by Dobler & Finkbeiner (2008b). For consistency with their analysis, we consider the template fit coefficients after conversion into intensity units. Practically, the coefficients relative to the  $\text{H}\alpha$  template are described as

Free-Free				
– function p				
	<i>Kp2</i>	<i>Kp0</i>	<i>KQ85</i>	<i>KQ75</i>
$F_0$	$8.21 \pm 0.63$	$7.73 \pm 1.01$	$7.52 \pm 0.97$	$7.05 \pm 1.55$
$D_0$	$0.05 \pm 0.10$	$0.05 \pm 0.16$	$0.05 \pm 0.15$	$0.04 \pm 0.24$
$C_0$	$0.31 \pm 0.38$	$-1.13 \pm 0.61$	$0.75 \pm 0.59$	$1.07 \pm 0.94$
– function g				
	<i>Kp2</i>	<i>Kp0</i>	<i>KQ85</i>	<i>KQ75</i>
$F_0$	$8.24 \pm 0.87$	$7.43 \pm 1.30$	$7.59 \pm 1.16$	$6.49 \pm 2.09$
$D_0$	$0.05 \pm 0.13$	$0.06 \pm 0.20$	$0.05 \pm 0.18$	$0.06 \pm 0.32$
$C_0$	$0.19 \pm 0.53$	$-0.99 \pm 0.79$	$0.08 \pm 0.70$	$-0.37 \pm 1.26$
$\chi^2$				
	<i>Kp2</i>	<i>Kp0</i>	<i>KQ85</i>	<i>KQ75</i>
$F_0$	$8.16 \pm 1.00$	$7.24 \pm 1.62$	$7.54 \pm 1.35$	$6.37 \pm 2.48$
$D_0$	$0.06 \pm 0.15$	$0.07 \pm 0.25$	$0.06 \pm 0.21$	$0.07 \pm 0.38$
$C_0$	$-0.29 \pm 0.60$	$-1.65 \pm 0.98$	$-0.54 \pm 0.82$	$-1.47 \pm 1.50$

Table 4.4: Values of the parameters obtained fitting the free-free coupling coefficients with the model proposed by Dobler & Finkbeiner (2008b) (see Equation 4.1). The coefficients are those derived using both the  $\chi^2$  analysis.  $C_0$  and  $D_0$  are consistent with zero, the later result demonstrating that we do not find any evidence of a spinning dust admixture with the free-free emission. The units of the parameters are kJy/sr/R.

a mix of free-free emission ( $F_0$ ) and radiation from the WIM ( $D_0$ ) that together generate the spectral bump that Dobler & Finkbeiner (2008b) observe. The final term ( $C_0$ ) with a CMB-like spectrum is a consequence of the initial subtraction of an estimate of the CMB sky before undertaking the template fitting. Thus the derived free-free coefficients are parameterised by the following relation:

$$I_{mod} = F_0 \left( \frac{\nu}{23\text{GHz}} \right)^{-0.15} + D_0(\text{DL98}, \nu_{peak} = 40\text{GHz}) + C_0 \left( \frac{\nu}{23\text{GHz}} \right)^2 a(\nu) \quad (4.1)$$

where DL98 is the WIM model of spinning dust due to Draine & Lazarian (1998a), but shifted in order to have the maximum emission at 40GHz, and where  $a(\nu)$  represents the conversion factor from thermodynamic to antenna temperature at a given frequency. In fact Dobler & Finkbeiner (2008b) utilise several variants of the CMB sky estimated by internal linear combination (ILC) methods. Since an ILC map will necessarily contain foreground residuals, the ILC-corrected data contain modified amounts of the expected foreground levels, and the  $C_0$  term effectively attempts to correct for this bias. Since we did not pre-process the data in the same way, we do not formally need to include this term in the analysis, but do so in the expectation that the coefficient  $C_0$  will be consistent with zero. This is, in fact, the case, and more importantly, as shown in Table 4.4, the  $D_0$  dust coefficient is also consistent with zero – we do not find any evidence of a spinning dust admixture with the free-free emission. Interestingly, the simple  $\chi^2$  analysis yields a spectral index that is very consistent with theoretical expectations ( $\sim 0.14$ ), especially with the *Kp2* and *KQ85* masks.

As noted before, the different treatment of the CMB component in the derivation of the coefficients, together with the different fitting methodologies, seem to play important roles in attempting a

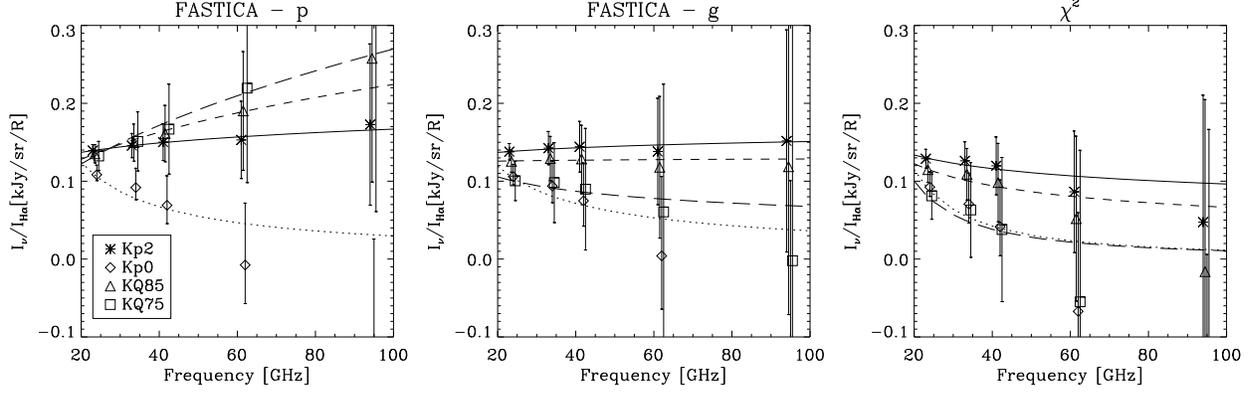


Figure 4.2: The coupling coefficients in intensity units (kJy/sr/R) for free-free emission as traced by the  $H\alpha$  template. We show the results from the  $p$  function analysis in the leftmost plot, from the  $g$  function in the middle plot, and the values from the  $\chi^2$  analysis are shown on the right. Asterisks represent the derived amplitudes for the  $Kp2$  sky coverage, diamonds are for  $Kp0$ , the triangles for  $KQ85$  and the squares for  $KQ75$ . Best-fit curves are also shown. The  $p$  results show an anomalous rising spectrum for the free-free emissivity with all the masks but the  $Kp0$ , whereas with the  $g$  function and the  $\chi^2$  analysis the results are more consistent with expectations. The  $Kp0$  mask shows consistently a steeper spectrum than expected. The  $\chi^2$  results are consistent with the expectations for free-free emission.

reasonable interpretation of the results and a connection to what has been found by other authors. Indeed, cross-talk between the CMB and foreground emission – the so-called ‘cosmic covariance’ of Chiang et al. (2009) – can likely both flatten and steepen the determined spectrum in some cases, and Dobler & Finkbeiner (2008b) have proposed that this effect makes it mandatory to subtract a CMB estimate before attempting any template fitting. In fact, we have observed such spectral changes in a set of simulations in which the foregrounds are described with idealised spectra (i.e.  $\beta_s = 3.1$ ,  $\beta_{ff} = 2.15$ ,  $\beta_d = 1.7$ , and a contribution from an  $H\alpha$  correlated WIM spinning dust component is either included or otherwise (see Chapter 5 for more details). However, the typical behaviour is such that the latter component is not detected erroneously when it is absent, and is detected when present. We do not then find a likely explanation of the inconsistency of our results with Dobler & Finkbeiner (2008b) given our unbiased recovery of the input coefficients.

## 4.5 Evaluation of the *WMAP* sky maps after foreground cleaning

Computing the coupling coefficients between the *WMAP* data and the foreground templates is useful both to study the spectral properties of the foregrounds and to allow the data to be cleaned for subsequent cosmological analysis. In our approach, we apply the  $p$  method to large sky areas, thus effectively assuming that each foreground has a single spectral index over the region of interest. In reality, this is incorrect, and we therefore expect that the cleaned data will contain residual foreground contamination, as a consequence both of this assumption, and the related one that the templates do indeed provide an adequate representation of the foreground morphology at microwave wavelengths.

The study of such residuals is then itself important in order to gain new physical insight, and to evaluate their impact on the statistical properties of the cleaned CMB maps. We have followed this approach in the 3-year analysis (see previous chapter).

Here, we wish to experiment further with the implementation of the cleaning approach applied to the data. To that end, we should recall that the coefficients are derived by maximising an approximation to the neg-entropy over the sky fraction of interest. Such a non-linear function could be dominated by one or two regions, that are either bright or have complex morphologies, thus the derived coefficients would be sub-optimal elsewhere. Indeed, one might speculate that such regions would occur close to the Galactic plane. Conversely, with a mask that only allows high latitude regions to be analysed where the spectral variations may be smaller, the coefficients may allow a relatively efficient cleaning of the data and small amounts of residuals. An interesting question then becomes by how much low latitude residuals would increase if the high latitude coefficients are applied. We therefore determine the template coefficients as a function of mask, apply the corresponding corrections to the multi-frequency data, and then inspect the corrected maps. In Section 4.6 we then investigate the related impact on an iterative cleaning approach. We show only the results based on the  $p$ -function analysis, since there is little visual difference relative to the  $g$ -function.

Figures 4.3 and 4.4 show the results for the Q-, V- and W-band data. We show two maps at each frequency - each map corresponds to the data as cleaned following the template analysis with the specified mask, but both the masked and full-sky are plotted. Showing the latter can help to reveal spectral mismatches between the emission in the masked region and the high-latitude area used for the analysis. The exception is the full-sky (FS) analysis where only the point-sources are masked in the analysis. It is clear that there is a complex pattern of behaviour as a function of mask, and some dependence on the template used to trace the synchrotron emission. However, with either the Haslam or K-Ka templates, it is apparent that in general, the thinner the applied mask, the higher the level of residual contamination. It should also be noted that in most cases the  $Kp0$  results seem anomalous and larger residuals are observed. When the Haslam template is used as the synchrotron tracer (Figure 4.3) the Q-band residuals towards the plane are the largest in amplitude, with the V- and W-band comparable. However, the residuals corresponding to fits performed on either the full-sky or using the processing mask are largely negative for the Q-band - there is an over-correction for foregrounds towards the Galactic plane. Interestingly, the low latitude residuals at V- and W-band for the full-sky analysis are relatively small and positive and confined to a very thin disk (a similar structure is seen at Q-band though embedded in the larger region of over-subtracted foregrounds). However, there appear to be more structures at higher latitudes than when more extensive masks are applied. Not surprisingly, the foreground coefficients are driven by the high amplitude Galactic plane signal, but are then sub-optimal for the high-latitude corrections. For the masks commonly applied for *WMAP* analysis, the cleaned maps at high latitude look very consistent, and the low-latitude residuals tend to be smallest for W-band. As noted above, the  $Kp0$  results are a curiosity and the level of residuals seems unexpectedly to be higher than for either the  $Kp2$ ,  $KQ85$  or  $KQ75$  masks, the latter case perhaps indicating some connection to the bright free-free regions that are masked therein. This argument may well be supported by the clear excess towards the Gum nebula. However, with the non-linear analysis

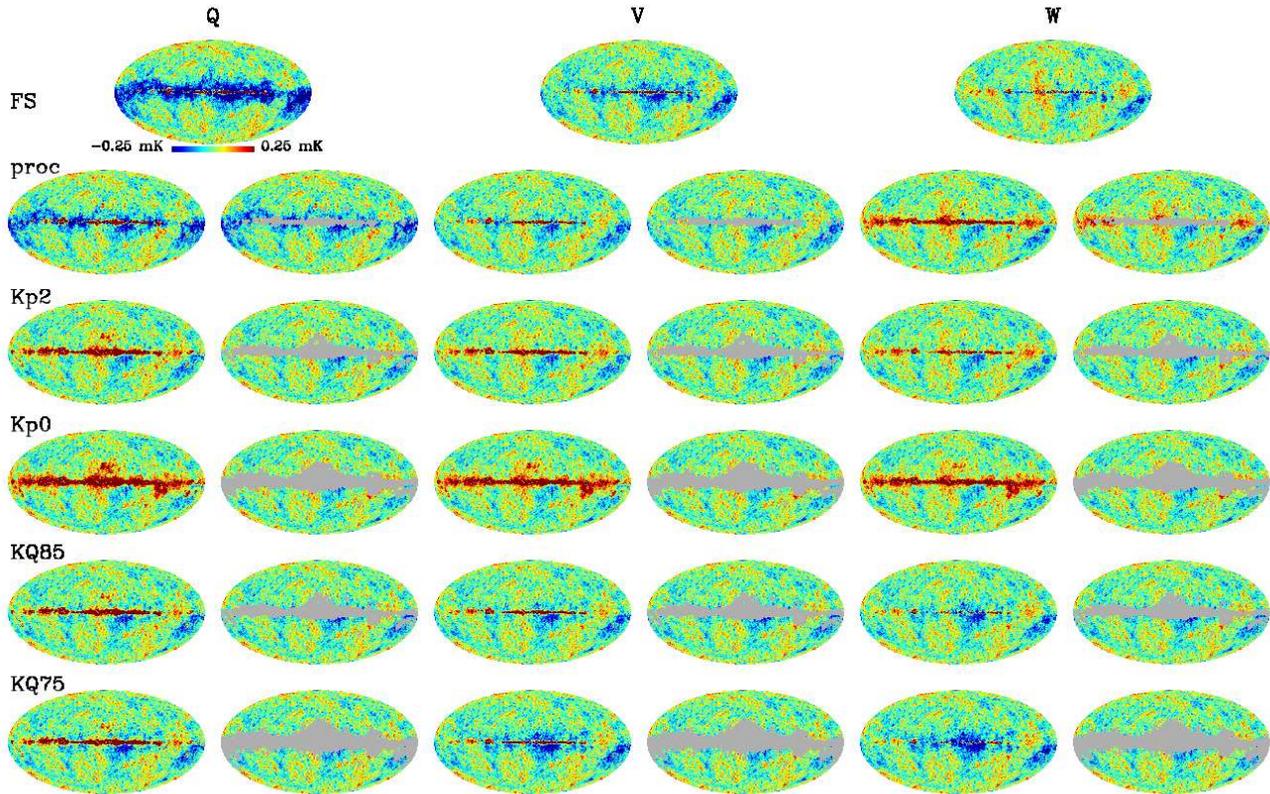


Figure 4.3: *WMAP* Q-, V- and W-band cleaned by subtracting foreground templates scaled by the coupling coefficients determined by  $\chi^2$  with the  $p$ -function. The maps are shown in a conventional mollweide projection in a Galactic frame-of-reference, with the north pole at the top of the image and the Galactic Center in the middle with longitude increasing to the left. The regions in grey correspond to the bright Galactic emission and point sources excised from the analysis by the masks. Each row corresponds to maps cleaned by coefficients derived from an analysis using one of 6 masks – FS (full sky, although the point source mask is still employed), proc (the processing mask), Kp2, Kp0, KQ85 and KQ75. The columns are divided into three pairs, each pair corresponding to one of the three frequencies considered as indicated. The left-hand plot in the pair is the full-sky map cleaned using the coefficients derived when applying the indicated mask, the right-hand plot then shows the cleaned map with the appropriate mask applied. The structure visible on the full-sky that falls within the analysis mask indicates the effect of spectral mismatch between the low- and high-latitude sky. Clearly, the first row can only show one map since the two cases coincide for a full-sky analysis. Here, the adopted synchrotron template is the Haslam map.

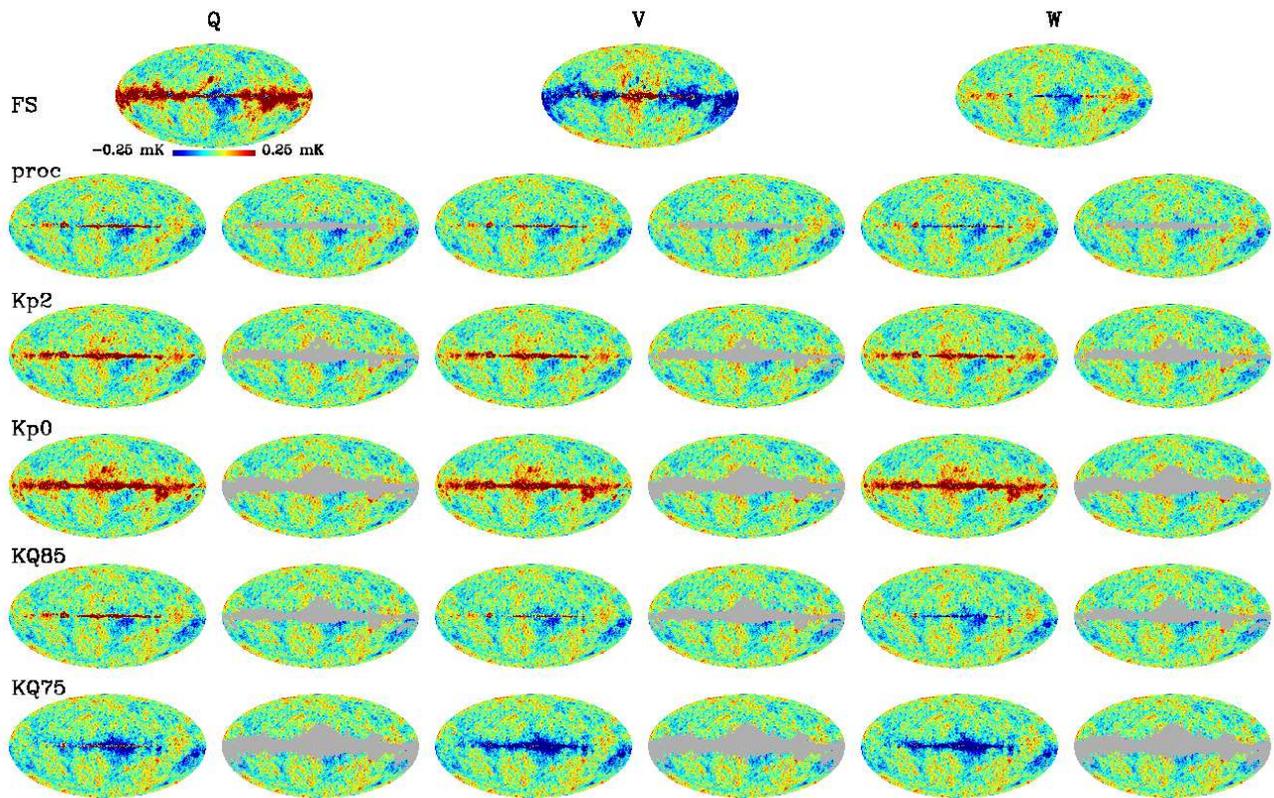


Figure 4.4: As the previous figure, but using the K-Ka template for tracing the low frequency foreground emission.

performed by it is difficult to be too precise about the origin of this behaviour. For the K-Ka template, the low-latitude residuals are generally positive except for the full-sky analysis where there is a complex mix of positive and negative structures at Q- and V-band, and for the *KQ75* mask where the residuals are predominantly negative. The results seem to suggest an improved cleaning of the data, particularly at intermediate latitudes. It is again reassuring to note that, in all cases, the dominant frequency-independent CMB structures are well-pronounced at high latitudes.

## 4.6 Iterative application of on pre-cleaned data

In Chapter 3, we introduced the concept of an iterative application of to the data. Specifically, we applied the algorithm to sky maps that had already been cleaned using templates with coupling coefficients themselves derived using a analysis. Here, we continue the study begun in the previous section by applying such an iterative step as a function of mask to data sets pre-cleaned with coefficients that are also a function of mask. However, we allow the mask utilised to determine the template coefficients to be different to that applied in the iterative phase to the cleaned data, and then investigate the resulting matrix of results. We utilise either the five frequency maps cleaned using the Haslam template, or the Q- V- and W- band data cleaned with either of the putative synchrotron templates (Haslam or K-Ka).

This approach to the problem is interesting in order to study in more detail the limits of application of the algorithm. Specifically we can address two related questions:

1. Can an initially poor cleaning of the data be compensated for by the iterative internal analysis of the cleaned data set?
2. Is it possible to derive a full-sky CMB map cleaned of any low-latitude residuals resulting from the application of corrections based on high latitude foreground coefficients? If not, can we determine the minimal mask, that allows reliable extraction of the underlying CMB signal?

In fact, we speculate that the combination of different masks at different stages of the iterative analysis can effectively introduce some sensitivity to the relative mix and spectral variations of the foreground components on the sky.

In our implementation, for a blind analysis of  $N$  input sky maps we expect  $N$  returned independent components. One of these will be identified as the CMB sky, the remainder correspond to other physical components on the sky. In Figures 4.5, 4.6 and 4.7, we show the spatial distribution of this residual foreground derived as a function of the mask used initially to compute the template coefficients for cleaning the input sky maps, and of the mask used to subsequently perform the iterative analysis. Each row of the plot corresponds to results derived from maps cleaned with those coefficients determined for a given mask, whilst each column corresponds to results derived when using the stated mask for the iterative analysis. Consequently, the diagonal maps are those returned when the same mask is adopted for the two steps of the analysis.

When the standard masks (i.e. *Kp2*, *Kp0*, *KQ85*, *KQ75*) are used in the two steps of the analysis, we find that only one of the additional non-CMB components returned by is actually consistent

with a foreground residual, the remainder being related to noise and residual dipole terms. A box in the plot highlights these maps which we retain as the most interesting results. Conversely, when either the processing mask is used or a full-sky analysis is performed in one of the two step of the analysis, the cleaned data remain strongly contaminated from the Galactic foregrounds, and there is no clear single map representing foreground residuals. Nevertheless, for completeness, we show representative maps that contain at least part of the strong foreground signal. In some cases, e.g. when the processing mask is used for iterative analysis after data pre-cleaning with the standard templates, there is evidence of consistent residuals with those shown in the red box in Figure 4.5. However, it is not possible to draw strong conclusions from these cases.

If we compare the new 5-year results with the 3-year results from our previous analysis, we can confirm the previous conclusions, which are actually reinforced thanks to the wider set of masks adopted here.

- There is strong evidence of a residual which is concentrated near the Galactic Centre and along the edge of the mask.
- There is a residual associated with the North Polar Spur seen in the Haslam data, clearly visible when all five *WMAP* frequencies are included in the analysis and the Haslam map is used as the synchrotron template. This oversubtraction of the emission implies a spectral difference between this region and high latitude regions. However, the amplitude falls off when a larger cut is used, which may indicate some variation of this index along the spur's extent.
- The spatial distribution of the contamination is larger when we analyze the K-, Ka-, Q-, V- and W-band data using the Haslam map as the synchrotron template than when only the Q-, V- and W-band maps are input data. Moreover, the amplitude of the residual is also lower in the latter case, implying a spectrum that falls with frequency. Whether the residual is a new physical component of emission or simply a region where the spectral index is flatter than the high latitude areas is unclear.
- The intensity of this component gets even weaker when the K-Ka map is adopted to model the synchrotron emission. This is a consequence of the fact that the K- and Ka-data do contain this emission component which also supports the concept that the use of internal templates such as the K-Ka map, even with a simple global scaling, provides a more accurate means to model the foreground contaminations than from data obtained at much lower or higher frequencies. However, the physical interpretation of such templates is difficult since they are certainly mixtures of several components.

Including the new cases of analysis, where the masks are mixed, other interesting points come out:

1. The processing steps with respect to masking do not commute - for example, the residuals from the iterative analysis using a *KQ85* mask applied to maps cleaned by *Kp2* coefficients are not identical to the case when the masks are reversed.

2. When the scaling factors are computed with the thinner masks (namely the  $Kp2$  and the  $KQ85$ ), the spatial distribution of the residual is similar regardless of the mask used for the iterative analysis. Therefore, the component is not a simple effect of the cut of the Galactic plane.
3. The algorithm provides a good component separation also when we adopt the coefficients relative to the largest masks: the algorithm is able to recover the residual whose amplitude depends on the mask applied in the second step of the analysis. What is not subtracted in the cleaning process, is actually recovered by the iterative analysis.
4. The residuals associated with the  $Kp0$  coefficients and the K-Ka template seem to show unexpectedly high amplitude compared to smaller cuts and the similar  $KQ85$  mask.
5. The  $KQ75$  coefficients computed with the K-Ka template seem to be the optimal ones in order to clean the data: the residual maps appear uniformly clean. The obvious conclusion would be that they give the most realistic description of the contaminated sky. However, this idea is not always supported by analogous considerations about the CMB component. In fact, the set of coefficients are reliable only when the  $KQ75$  mask is used also for the iteration. In the other cases, mixes all the components and part of the residual actually contaminates both the CMB map and a third component, which generally shows a noise pattern.

The last point highlights that a consideration of the residuals revealed by the iterative analysis can actually be misleading, and is not a sufficient figure of merit to quantify the performance of such an analysis by itself. To make such a judgement, we must consider the CMB component returned by the analysis and, in particular, its power spectrum. However, before we make this assessment in Section 4.7, we make a small digression.

#### 4.6.1 Is the foreground residual the ‘WMAP haze’?

Since we cannot simply associate the recovered residual component with any of the standard templates used to describe the diffuse Galactic emissions, we attempt to address the issue of its spatial distribution and physical origin. In fact, Dobler & Finkbeiner (2008a) have already attempted to give an answer to these questions, and identify the residual (the so-called *WMAP haze*) as a hard synchrotron component, whose origin is linked to dark matter annihilation in the centre of our Galaxy (see Chapter 1). They also provide a simple model for its spatial distribution, namely:

$$\mathbf{h} \propto \begin{cases} \frac{1}{r} - \frac{1}{r_0} & \text{for } r < r_0; \\ 0 & \text{for } r > r_0, \end{cases} \quad (4.2)$$

where  $r$  is the angular distance to the Galactic center and  $r_0$  is arbitrarily set equal to 45 degrees.

Examination of the maps returned by  $\mathbf{h}$  indicates that a spherically symmetric distribution is not the optimal one to describe the component. Nevertheless, a direct fit of this model to our maps shows a quite good agreement, with a correlation coefficient of 0.45. Therefore, we adopted it as a fourth template and repeated the iterative analysis using the Haslam map as the soft-synchrotron template.

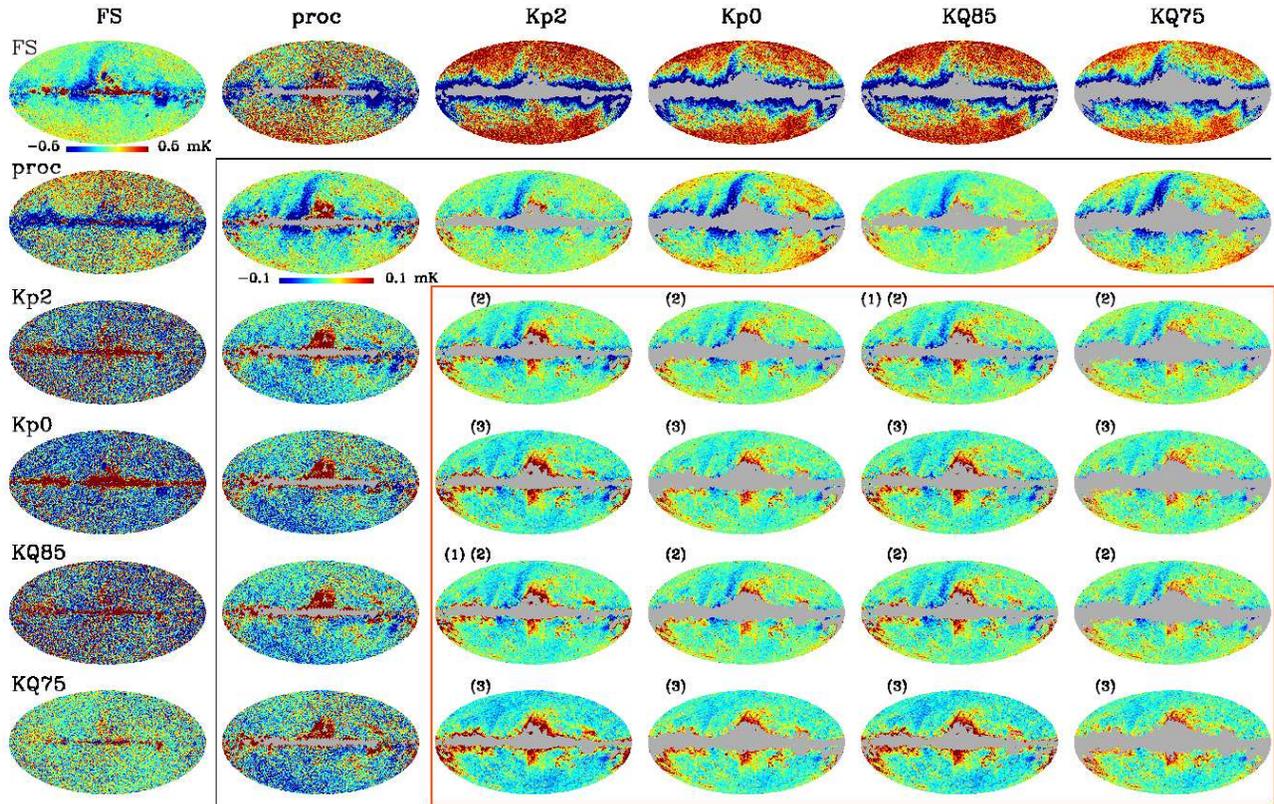


Figure 4.5: Maps of the residual component recovered by  $\chi^2$  with the iterative blind analysis of the cleaned K-, Ka-, Q-, V-, and W-band data. Results on a given row are derived from input maps cleaned using coefficients determined for the stated mask; results on a given column are derived when the stated mask is used in the iterative step of the analysis. The labels have the same meanings as Figure 4.3. The numbers over the maps refer to the list of comments in the text. Here, the Haslam map is used as a synchrotron template. The red line highlights the cases where single well-defined residual component is detected. The black line separates regions where different temperature scales have been used for the plots.

A physical interpretation of the coupling coefficients returned by  $\chi^2$  is not simple: the values are negative for all the frequencies except the K- and Ka-bands, making a spectral index computation very difficult. Nevertheless, we used the values to clean the data and considered the amount of contamination still present in the residual map from the corresponding iterative analysis. Although these coefficients are representative of the best  $\chi^2$  solution when the haze template is used, a negative value adopted to clean the data can effectively correspond to the introduction of a spurious foreground residual in the data. Therefore, we also studied the case where the haze template cleaning was only performed for the K- and Ka-bands; for the remaining frequencies we retained the old three-template coefficients. Figure 4.8 shows the residual maps obtained in these two cases when the  $Kp2$  mask is applied to the data together with the previous result derived from the three templates fit of the data.

When the haze template is employed, bright emission around the Northern extension of the Galactic Centre is still present, whilst the Southern hemisphere emission is essentially removed. There is little obvious difference between the two cases. Some refinement of the haze model is still required.

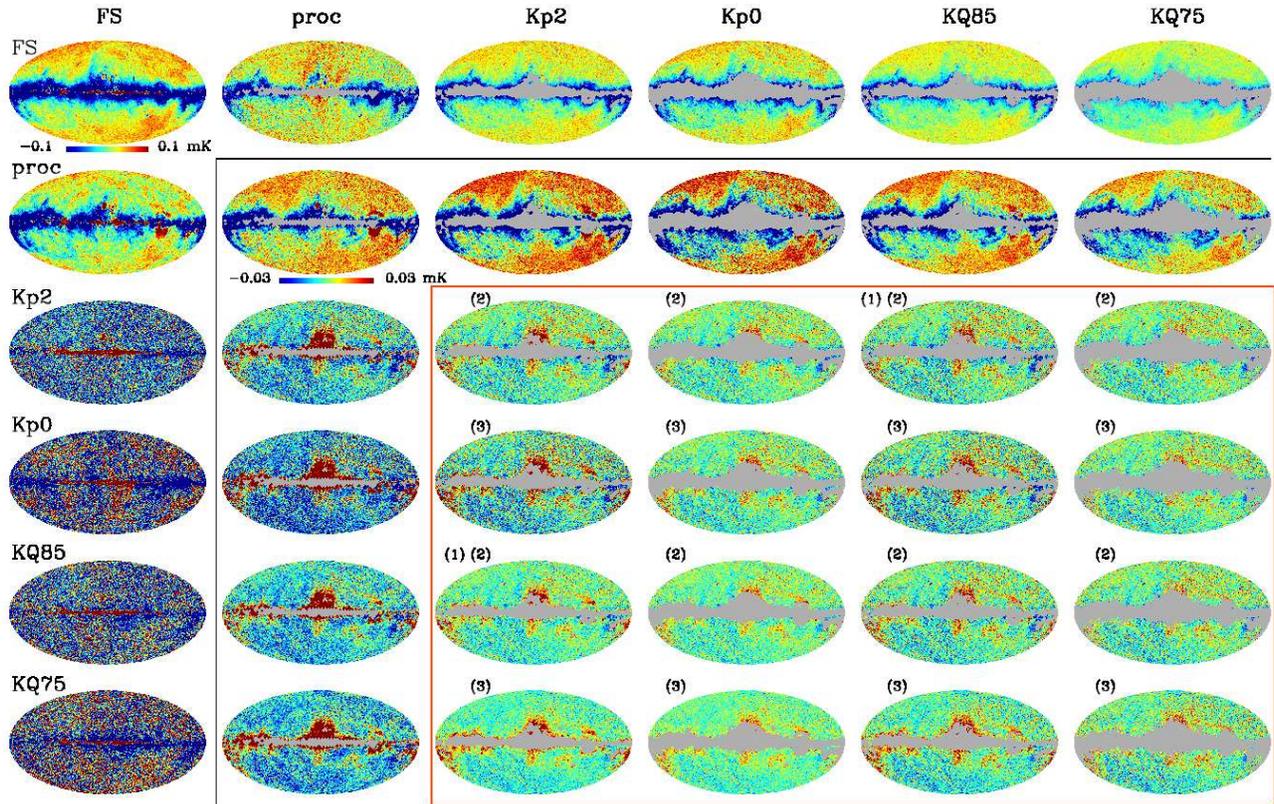


Figure 4.6: Maps of the residual component recovered by  $\text{proc}$  with the iterative blind analysis of the cleaned data Q, V and W band. Rows and columns have the same meaning as Figure 4.3 as well as the red box and the numbers, and again the Haslam map is used as the synchrotron template.

## 4.7 Power-spectrum evaluation of iteratively cleaned maps

Note that, in what follows we will use the *WMAP* team's 5-year power spectrum derived from the V- and W-band data using the MASTER algorithm over all angular scales<sup>1</sup> as a canonical reference. This should not imply that the result is indisputable, nevertheless several independent analyses have yielded results in good agreement. Moreover, we will make comparisons on a qualitative level regarding the broad features of the spectrum, rather than make detailed inferences about cosmological parameters.

In Figure 4.9, we show the CMB maps obtained when the five *WMAP* bands are cleaned using the coefficients determined from an analysis when the Haslam map is employed as the synchrotron template: we chose this case as the best example to comment on, since it connects to further analyses described later on. We divide a description of the results into two sections – the first dealing with more conventional analysis when only partial skies are analysed, the second considering cases when the signal for the full sky is utilised in at least part of the analysis.

<sup>1</sup>Private communication: G. Hinshaw (2009)

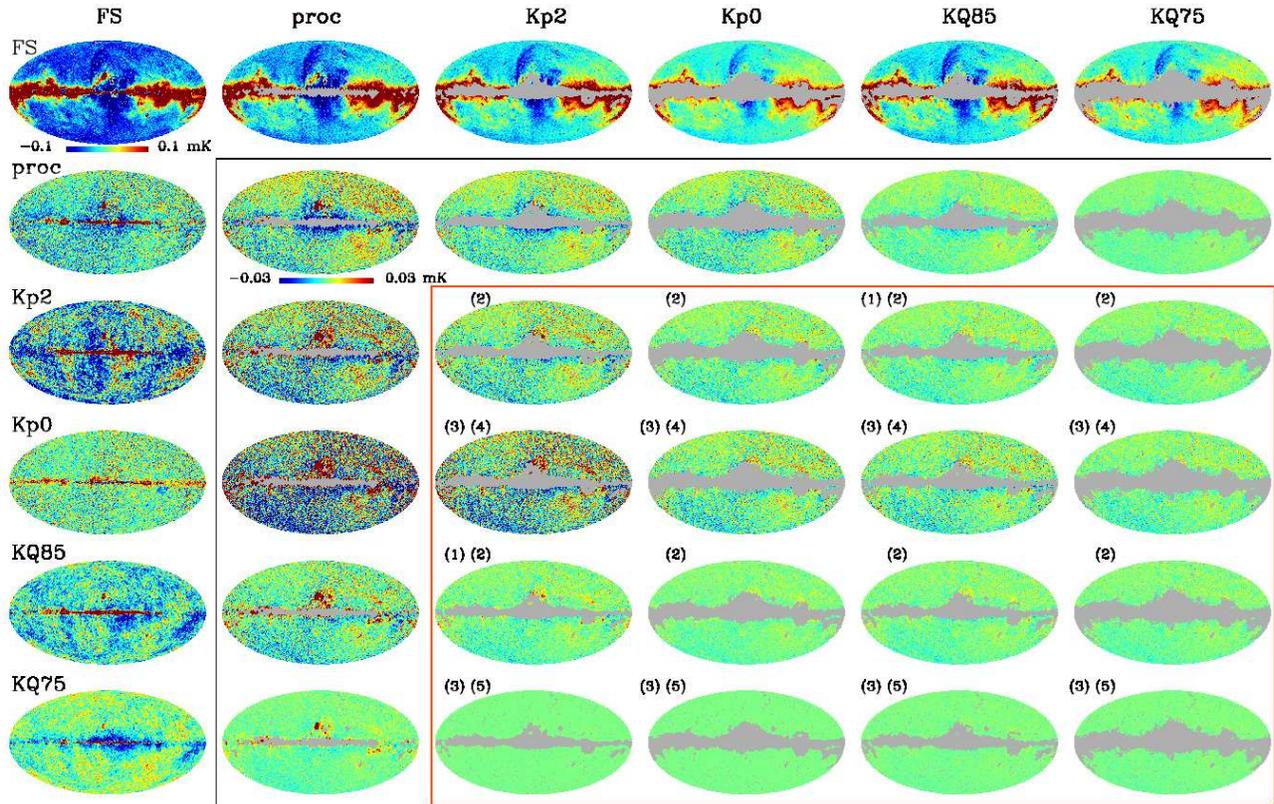


Figure 4.7: As Figure 4.6 but using the K-Ka map as the synchrotron template.

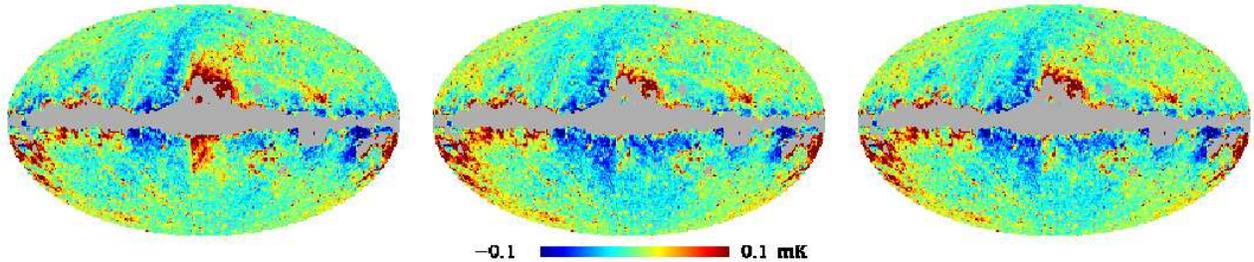


Figure 4.8: Maps of the residual component recovered by  $\chi^2$  with the iterative blind analysis of the cleaned data. These are produced using the coupling coefficients obtained from the templates fit where the haze template is not used (left), when it is added as a fourth template (center) and when the haze coefficients are only used to clean the K- and Ka-band maps. The Haslam map is adopted as the synchrotron model.

#### 4.7.1 Partial-sky analysis

Generally, with all the possible combinations of coefficients and masks, we obtain CMB maps whose power spectrum is highly consistent with the best estimation of *WMAP*. In Figure 4.10 we have plotted a subset of all possible power spectra computed by our iterative methodology. The selection was carried out on the basis of a reduced  $\chi^2$ -value computed relative to the *WMAP* spectrum, and we show some best (top panels) and worst (bottom panels) cases. Note that all plotted power spectra are corrected

for the contributions of the instrumental noise (determined using simulations) and the unresolved point sources following Wright et al. (2009).

The values of the  $\chi^2$ -statistic are generally driven by the low ( $\ell < 6$ ) and high multipoles ( $\ell > 250$ ). In order to exclude the possibility that such excesses seen in the power spectra could be associated with an underestimation of the noise, we implemented a cross-power spectrum estimator. Such a method requires two input CMB sky estimates, which we derived using subsets of the individual DA sky maps. For Q- and V-bands, the split is unambiguous since there are only two DAs per frequency. For the W band, we averaged W1 with W2 and W3 with W4. Finally, since for the K-band we have only one differential assembly, two independent inputs were created from averaged maps of the first, second and third year of observations and the fourth and fifth year, respectively. The inverse has been done for the Ka band where the same problem exists. The effect of these selections is to approximately balance the noise properties of the input maps for the power spectrum evaluation. For each of the two sets of maps, the individual frequencies were then combined with the coefficients derived from the standard analysis, and the cross-power spectrum evaluated. In all cases, this has confirmed the auto-power spectrum estimation after noise correction. Therefore, the excess likely reflects contamination from foreground residuals at low latitude, rising from an inability of the algorithm to disentangle the components. This contamination is probably due to the different extension of the masks used to clean the maps and to internally analyse them.

However, given this interpretation, some anomalies are difficult to explain. We would expect the contamination to be larger when the *Kp2* and *KQ85* masks are used for the internal analysis, yet the *KQ85* mask returns a CMB map with a lower contamination than for *Kp0*. Moreover, following this logic, we would expect the processing mask provided by the science team of *WMAP* to exhibit the most pronounced excess at high  $\ell$ , but on the contrary it is very small. This suggests that the component separation is mostly driven by specific features of what appears as the residual, rather than simply by its angular extension. This interpretation is also appropriate to explain the results obtained from the maps cleaned using the *Kp0* mask coefficients and then analysed with the processing cut. The CMB component returned by has a power spectrum which is consistently higher than the other cases, with a larger noise contribution at high multipoles.

### 4.7.2 Full-sky analysis

As a further experiment, we wanted to test the limits of the applicability of the algorithm by performing a component separation study on *full-sky* maps. In practice, by full-sky we mean the remaining sky coverage after exclusion of the point sources. This experiment is interesting in order to see if is able to function satisfactorily, even when the different sources are strongly mixed with each other, as they clearly are in the Galactic plane.

In the first stage of the analysis, we find that the component separation based on template fits is poor. This is only to be expected, since the spectral behaviour of the foregrounds traced by the templates has well-established differences at low- and high-latitude, and since the bright emission from the Galactic plane is likely to drive the fit coefficients. We find that the coefficients are very large for the dust

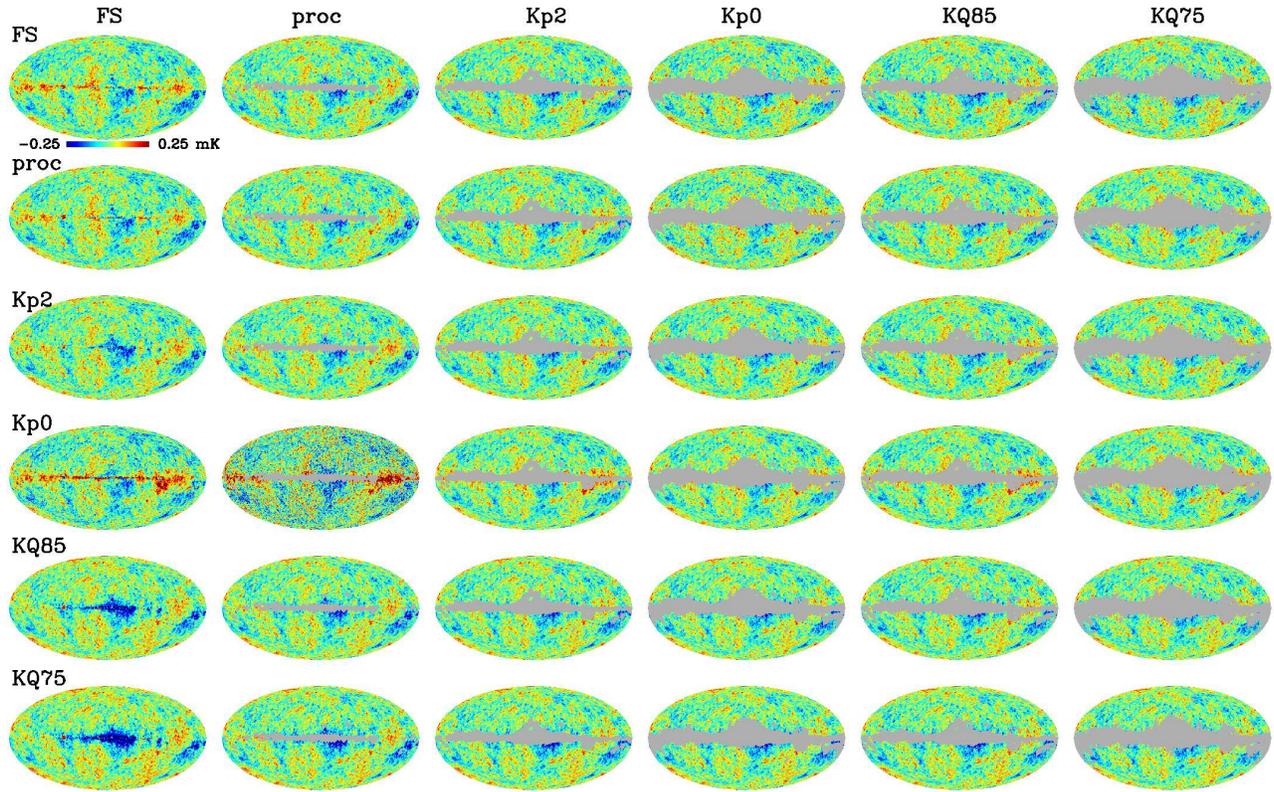


Figure 4.9: CMB components recovered by  $\chi^2$  minimization when the cleaned *WMAP* data are used as input maps. The data are cleaned according to the scaling factors obtained using the Haslam map as synchrotron template, the  $p$  function and *all* the masks. The iterative analysis itself has been performed with all the different cuts of the sky and the five maps. Each row is referred to the mask adopted to compute the coefficients, further used to clean the data. Inside the row, the single map is obtained from the internal analysis and using a specific mask, whose name is indicated on the top of each column.

and free-free emissions, but small for the synchrotron radiation. The latter factor probably reflects the degree of mixing between the sources, which is particularly strong between the synchrotron and free-free emissions. However, the general trends in spectral behaviour are maintained, and we adopt the coefficients to clean the data as usual.

After the iterative step, the final recovered full-sky CMB map shows clearly extraneous features along the Galactic plane, as shown in the top left corner of Figure 4.9. The corresponding power spectrum (see Figure 4.11) shows, accordingly, a substantial excess on almost all angular scales. Note that the enhancement on small scales ( $\ell > 250$ ) is in the regime limited by the  $1^\circ$  resolution of the analysis. Nevertheless, although we do not wish to over-emphasise its significance, it is likely indicative of foreground residuals.

The cleaned full-sky maps (top-row of Figure 4.9) were also iteratively analysed using the usual set of masks to exclude the Galactic plane. In these cases, we find that, on application of a mask, the excess of power at intermediate and high multipole values largely disappears, confirming that most of the previously observed excess is connected to the spurious structures along the Galactic plane, at

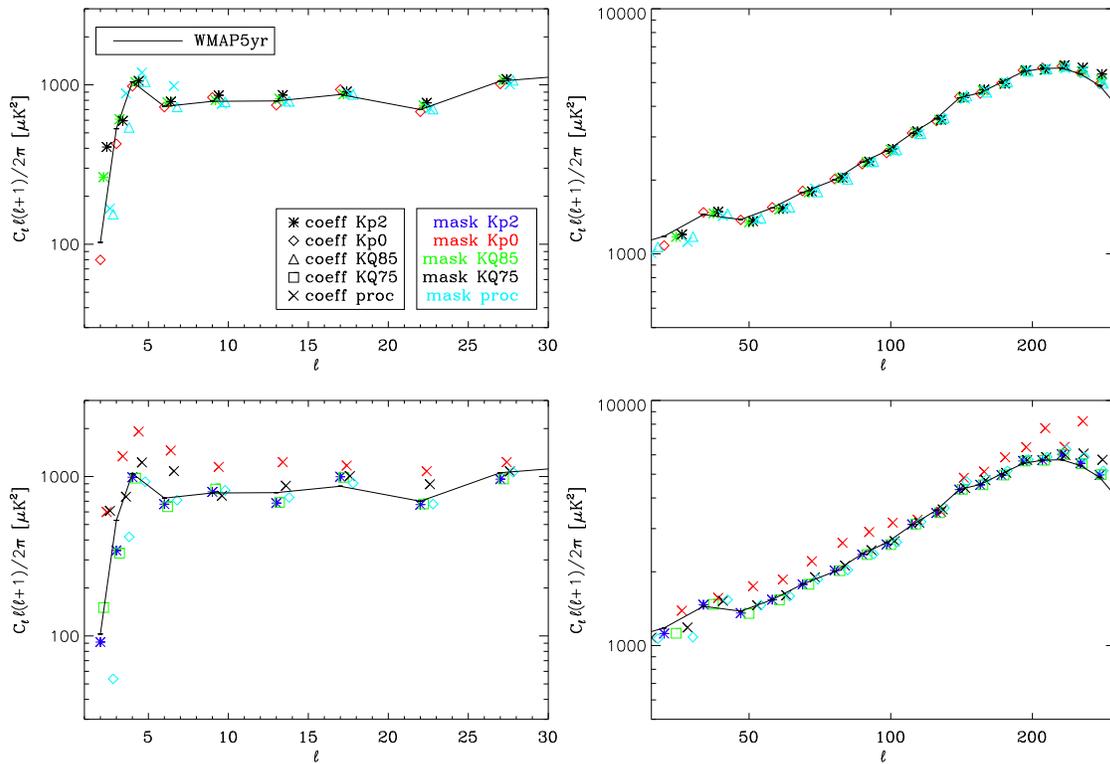


Figure 4.10: Subset of the power spectra of the CMB components recovered by an iterative application of when the cleaned *WMAP* data are used as input maps. The data are cleaned according to the scaling factors obtained using the Haslam map as synchrotron template, the  $p$  function and *all* the masks. The iterative analysis itself has been performed with all the different cuts of the sky and the Q, V and W maps. For an input set of maps cleaned using coefficients determined for a specific mask, best- (top row) and worst-case (bottom row) spectra are chosen from the corresponding iterative analyses for all possible analysis masks using a  $\chi^2$  statistic relative to the best-fit *WMAP* spectrum. For clarity, the left column shows the results at low- $\ell$  on a linear scale, the right column shows the intermediate and higher- $\ell$  values on a logarithmic scale.

latitudes smaller than  $\sim 5^\circ$ . Indeed, the processing mask itself is already sufficiently large to produce this result.

Finally, we have internally analysed the full-sky maps cleaned utilising the coefficients computed using a mask (leftmost column of Figure 4.9). As expected, the data appear strongly contaminated in the Galactic plane region, since the scaling factors describe the properties of the high-latitude sky outside of a specific cut. An intriguing result is constituted by the difference in the nature of this residual contamination as a function of mask: for the high-latitude cuts, there is generally a negative region in the Galactic plane, except when using the *Kp0* mask which retains positive Galactic plane structure. This again suggests anomalous behaviour for those results derived using this mask, following that described in Section 4.4 and further discussed in Section 4.4.1, dedicated to the spectral index estimation. The consequence of the existence of this contamination is again an excess at high values of

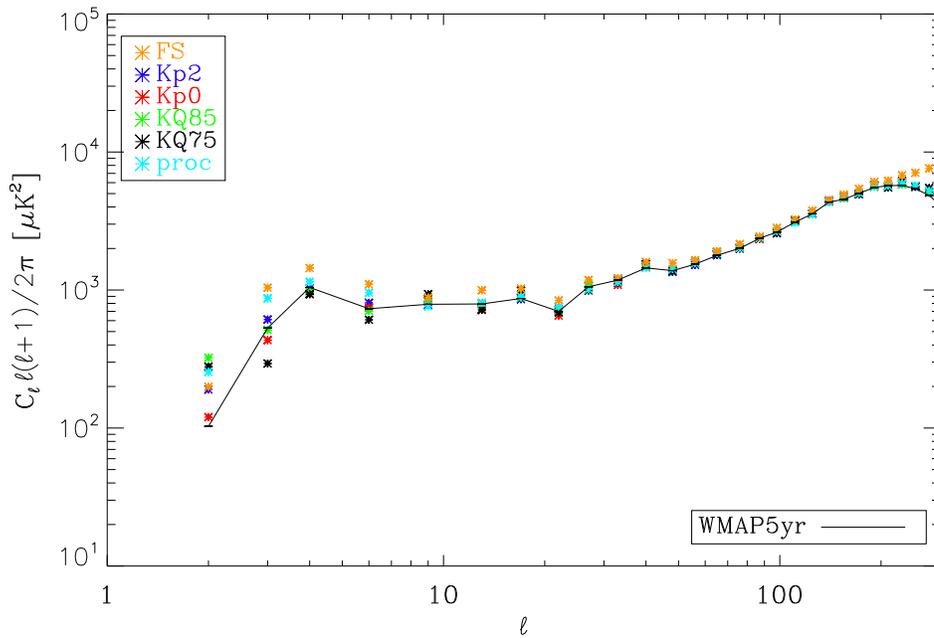


Figure 4.11: Power Spectra of the CMB components recovered by the algorithm applied to the data cleaned according with the coefficients obtained with a full-sky analysis. The Haslam map is adopted as synchrotron template and the coefficients were computed with the  $p$  function. The iterative analysis itself has been performed on all the five maps.

$\ell$  in the CMB power spectra, which extends to intermediate scales most notably when using the  $Kp0$  and  $KQ75$  masks. None of the spectra can be considered compelling for cosmological purposes.

## 4.8 Application of the algorithm on *WMAP* data and templates simultaneously

From what has been presented in the previous section, we conclude that the experimental full-sky analysis does not satisfy our expectations. This happens for two reasons: first of all, we lose information on the foreground components because the template fit is poorly behaved; secondly, the CMB reconstruction shows significant contamination, preventing its use in cosmological studies. However, we did not abandon the idea of producing an adequate full-sky CMB estimation, and considered a data set compromising the ensemble of multi-frequency maps dominated by the CMB component, simultaneously analysed with foreground templates. We considered different combinations of the input data: the five *WMAP* maps were analysed together with the dust and free-free maps, and afterward also including the Haslam map as synchrotron template. The same was done with only the Q- V- and W-band maps: in this case, when the synchrotron template was involved, we used both the Haslam and K-Ka maps. Finally, we considered the five *WMAP* frequency channels alone, in order to quantify the advantages of including foregrounds models.

In Figure 4.12, we show the CMB maps obtained with this new implementation of the algorithm.

Unsurprisingly, it appears that the higher the number of input maps, the greater the precision of the CMB map reconstruction. In particular, with all five *WMAP* maps and the three foreground templates, we obtained the best CMB map with respect to those recovered in all the other combinations of input data. Spurious structures along the Galactic Plane are clearly visible in all the maps, while in this case they get much weaker and are perhaps even confined to the inner part of the Galaxy. Quantitatively speaking, we checked the impact of these structures on the power spectrum, shown in Figure 4.13. The CMB map returned with the largest number of input data is the most consistent with the best estimate from the *WMAP* 5-year data, even though at high  $\ell$  values there is still an excess of power, which is clearly the consequence of these residuals.

However, we focused our attention on the map obtained from the combination of the largest data set, computing as usual the power spectrum as a function of applied mask. This analysis, shown in Figure 4.14 seems to confirm that the excess is related to structures observed along the plane since it disappears with more extensive cuts. Moreover, in Figure 4.15 we compared its power spectrum to those of the CMB maps derived with the internal full-sky analysis of the cleaned data. The aim of this exercise was to verify that the new implementation of  $\ell$  was an improvement on previous attempts.

The excess of power for  $\ell > 200$  is smaller than in the other cases of comparison, even though the difference is not marked and we are at the limit of the multipoles range imposed by the data angular resolution of  $1^\circ$ . On the other hand, however, at intermediate values of the multipoles, it is more consistent with the *WMAP* best estimation, while the other power spectra are slightly higher in amplitude.

Ultimately, we want to compare the  $\ell$  method with the results of other methods currently used to achieve the same aim. We considered the ILC map produced by the *WMAP* science team (Gold et al., 2009), a similar foreground-reduced map in which the frequency dependent weights were determined in harmonic space by Kim et al. (2008) (hereafter HILC) and a further alternative obtained by Delabrouille et al. (2009) using needlets as the basis of the analysis (hereafter NILC). Since these techniques make use of the *WMAP* data alone without any augmentation by internal templates, we revert to the fully blind analysis also performed for the *WMAP* three-year data in Maino et al. (2007), and consider the case when all five frequency bands are used. Again, the power spectrum has been chosen as the figure of merit to compare the performances of these methods. These are shown in Figure 4.16. We have used published information to compute the noise correction to the spectra based on simulations, except in the NILC case where a correction has been directly provided, though here no correction for unresolved point sources has been made. The power spectra of the HILC and the NILC maps are the most consistent with the *WMAP* best estimation, whereas the ICA map seems to be the most contaminated with residuals, since it shows the largest excess of power on small angular scales. However, in some sense, the comparison is unfair towards  $\ell$  since the other methods allow some regional dependence of the weights used to form the optimal CMB estimate, either by solving for coefficients in different regions of the sky separately, or finding a best spatially varying set of weights.

In conclusion, the ICA approach affords a fair estimate of the CMB signal, although improvements are still possible. However, at this stage, we simply took advantage of this result to address another

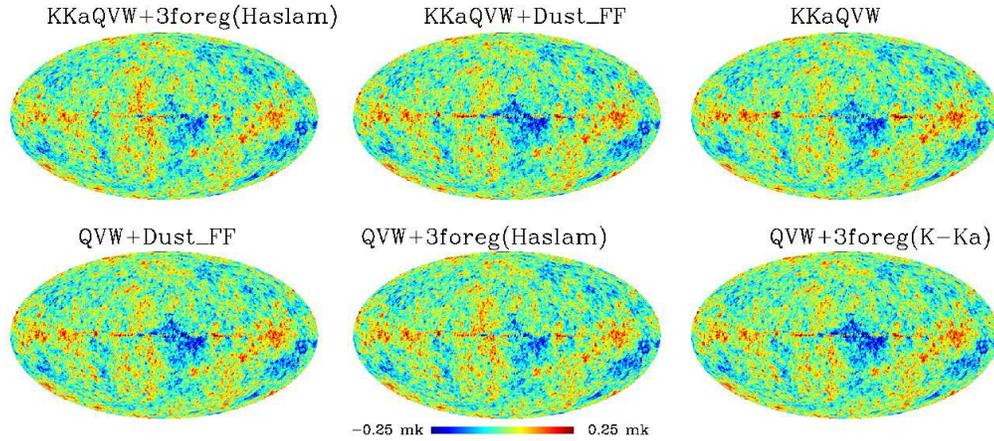


Figure 4.12: CMB maps returned by  $\text{WMAP}$  when applied on different combinations of data. Several cases have been considered: i) the *WMAP* maps at the K-, Ka-, Q-, V- and W-bands together with three foregrounds templates, where the Haslam map is used to describe the synchrotron emission; ii) the *WMAP* maps at the K-, Ka-, Q-, V- and W-bands together with the dust and free-free templates; iii) the *WMAP* maps at the K-, Ka-, Q-, V- and W-bands; iv) the *WMAP* maps at the Q-, V- and W-bands with the dust and free-free templates; v) the *WMAP* maps at the Q-, V- and W-bands together with three foregrounds templates, where the Haslam map is used to describe the synchrotron emission; v) the *WMAP* maps at the Q-, V- and W-bands together with three foregrounds templates, where the K-Ka map is used to describe the synchrotron emission. The cleanest map is the one obtained from the first case considered, but generally spurious structures are visible along the Galactic Plane in all the maps.

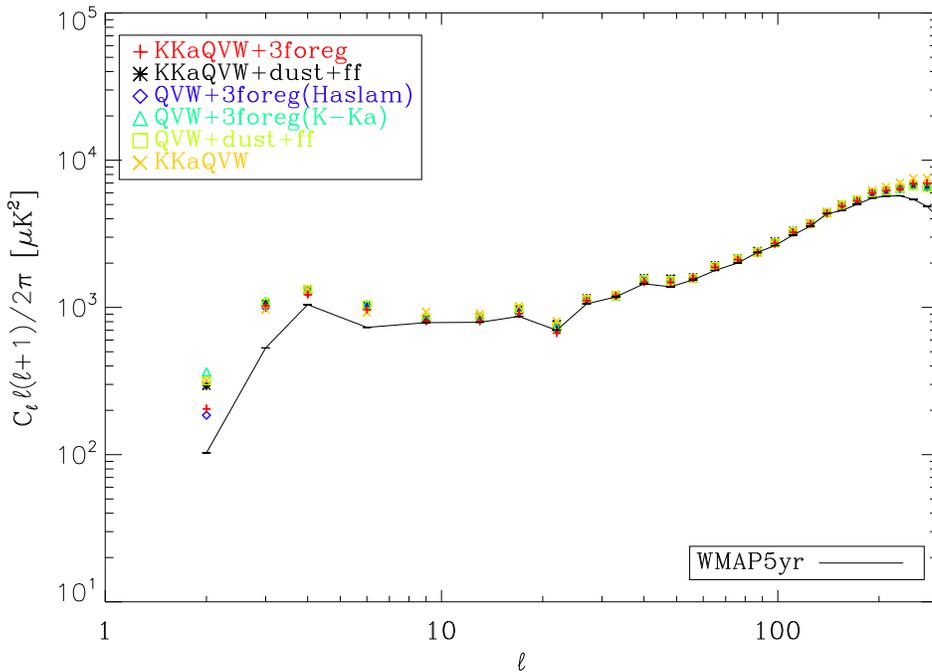


Figure 4.13: Full-sky power spectra of the CMB maps shown in Figure 4.12.

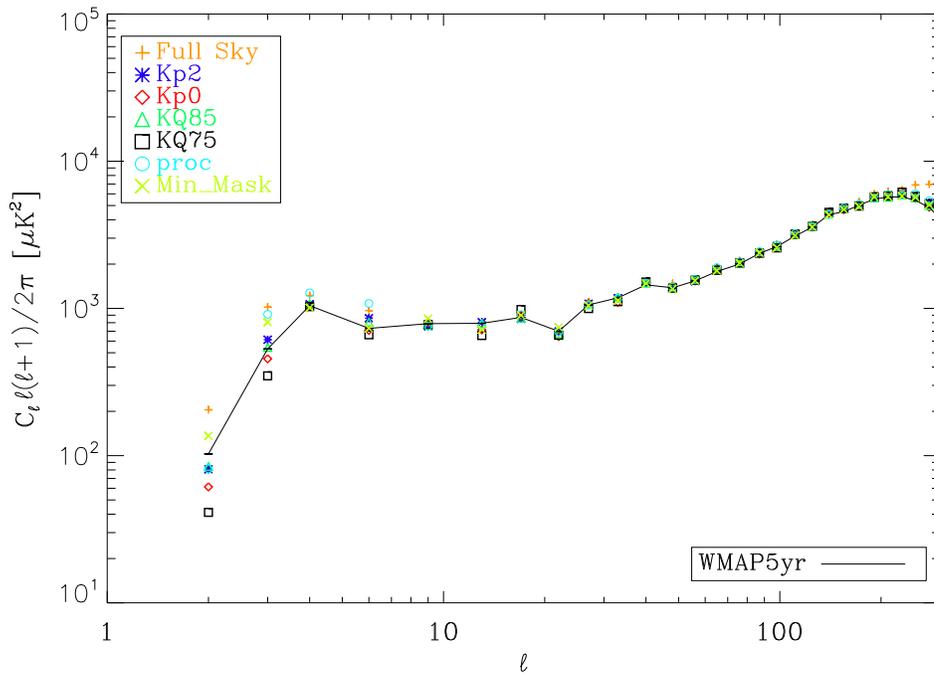


Figure 4.14: Power spectra of the CMB map (KKaQVW+3foreg in Figure 4.12) returned by a full-sky analysis when applied simultaneously to the five *WMAP* maps together with three foregrounds templates. We used the Haslam map as synchrotron template and the  $p$  function. Power-spectra are then evaluated after applying a variety of masks.

question.

#### 4.8.1 What is the minimal mask?

Given the imperfections of our full-sky CMB map reconstruction, we are forced to apply a cut in order to use it to extract cosmological information. We therefore took the opportunity to define the minimal mask required by to recovered a non-contaminated CMB map.

In order to determine the mask, we first implemented the *WMAP* thresholding method used to generate the  $Kp2$  and  $Kp0$  masks as described by Bennett et al. (2003), but here applied to the map. However, we found that, given the relatively low amplitudes of the residuals, pixels for a given threshold set soon included likely genuine CMB structures, which then contaminated the mask. Consequently, we considered simple parallel cuts of different latitude extension. Finally, we determined a minimal mask, which is the union of the processing mask and a parallel cut of the regions with  $|b| < 6^\circ$ . As usual, we also excluded the point sources, according to the mask provided by the *WMAP* science team. The result is shown in Figure 4.17. Figures 4.14 and 4.16 then show the power spectra derived from the *WMAP* data either including or excluding the three foreground templates respectively for this mask. In both cases, the agreement with the *WMAP* spectrum is impressive over all  $\ell$ , with the remaining differences being on large angular scales.

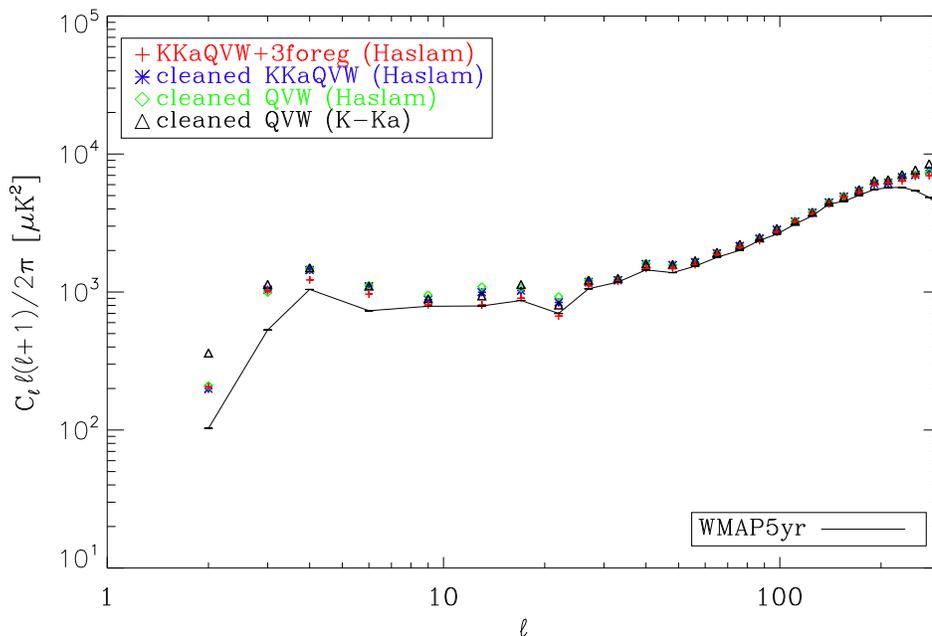


Figure 4.15: Comparison of the power spectrum of the CMB recovered from the simultaneous full-sky analysis of the *WMAP* data together with the templates and those obtained with the simultaneous internal analysis of the cleaned data. We show all the different combinations of the input data internally analysed. The results are obtained using the  $p$  function.

## 4.9 Discussion

In this chapter, we have undertaken a foreground analysis of the *WMAP* 5-year data using the algorithm, as previously applied to the *WMAP* 3-year data (see Chapter 3). Various improvements in the implementation have allowed us to address several open questions from our previous work, and allowed some experimentation with the technique.

We used the code to perform a foreground fit of the data on a frequency-by-frequency basis, where the Galactic components are described by all-sky templates obtained at wavelengths where the corresponding emission mechanisms dominate. Specifically, we adopted the Finkbeiner (2003)  $H\alpha$ -map as a template for the free-free emission, the Finkbeiner et al. (1999) FDS8 model for thermal dust emission, and the 408 MHz radio continuum all-sky map of Haslam et al. (1982) as utilized in the first year *WMAP* analysis for the synchrotron emission. In the latter case, we have also considered the differenced of the *WMAP* K- and Ka-band data, as preferred in their 3-year analysis.

The first step of the analysis is the computation of the coupling coefficients, which give an estimation of the contamination of the foreground emission level in the data. We know already that these coefficients depend on the extension of the mask used to exclude the most contaminated regions of the Galactic plane. We further investigated this dependence, thanks to the new *KQ85* and *KQ75* masks provided by the *WMAP* science team. We confirmed the result which reinforces the idea of a spatial

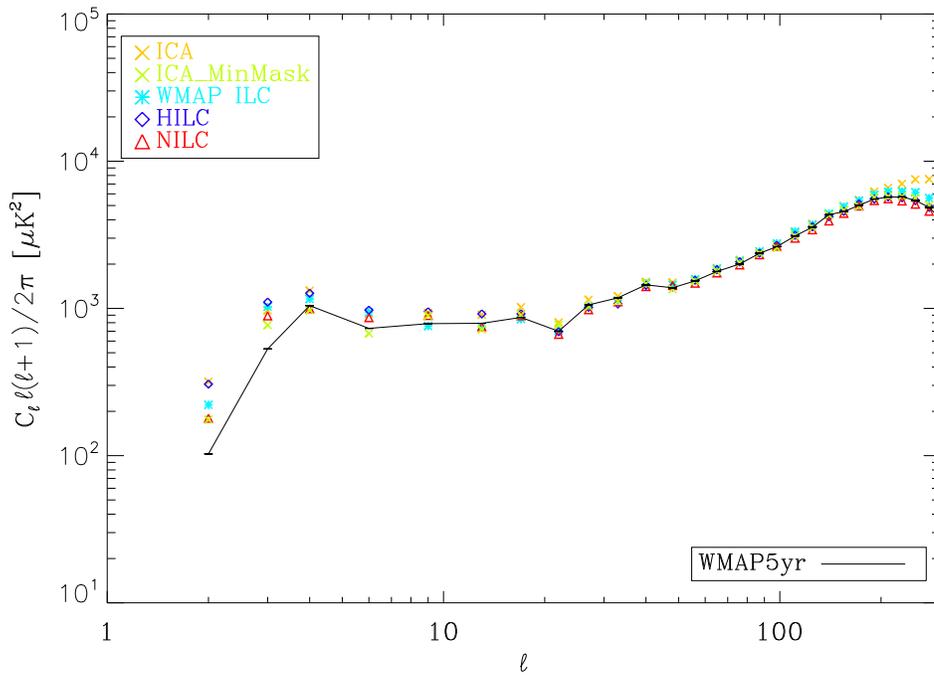


Figure 4.16: Power spectrum of our best full-sky map obtained with [ICA](#), compared with those of the *WMAP* ILC map, the HILC map produced by Kim et al. (2008) and the NILC map of Delabrouille et al. (2009).

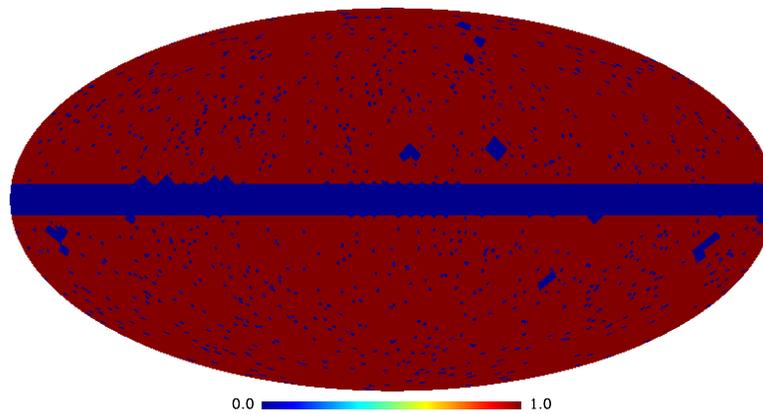


Figure 4.17: Minimal mask derived from the [ICA](#) simultaneous analysis of the five full-sky *WMAP* data with the three foreground templates. It is basically the combination of the processing mask and a parallel cut of the regions of the Galactic plane with  $|b| < 6^\circ$ . As usual we also excluded the point sources, according with the mask provided by the *WMAP* science team.

variation of the foreground emissions. Moreover, it suggests the component separation is driven by some key regions, rather than the extension of the mask itself. On the other hand, we reestablish some anomalies obtained when the *Kp0* mask is adopted that remain unexplained.

We have considered the spectral behaviour of the derived scaling factors when the Haslam et al.

(1982) data is used as the synchrotron template, since an interpretation of the analogous results derived using the K-Ka template is compromised by the fact that it is itself a mixture of foreground components. We evaluated the spectral index for the synchrotron emission, the anomalous dust-correlated component, and the free-free emission. In the first two cases, we found steeper, though statistically consistent, spectral behaviour as compared to previous analysis, e.g. Davies et al. (2006).

We then focused our attention on the estimation of the free-free spectral index. A non-trivial dependence of the spectral behaviour on the extension of the mask confirms what was already found in the previous chapter, with the *WMAP* 3-year data. However, an understanding of the behaviour is probably associated with two factors – the spatial variation of the physical properties of the component (e.g. electron temperature which directly relates to the scale-factor associated with the  $H\alpha$  template) which is mostly linked to specific structures close the Galactic plane, and the connection of the statistical properties of the foreground with the response of the algorithm as a function of the non-linear function used. The spectral behaviour is generally flat and even increasing with frequency, if the  $p$  function is used together with the *KQ85* and *KQ75* masks. We have interpreted these spectral features as a consequence of a self-similarity of the Gaussian/non-Gaussian mixtures at each frequency. When the  $g$  function is adopted, a more uniform behaviour, consistent with theoretical expectations, is obtained. Finally, in both the cases, the *Kp0* mask gives a steeper spectral index than the expected value of 2.14. We note that none of our results seem to be compatible with the existence of a bump in the spectrum, as claimed by Dobler & Finkbeiner (2008b) and subsequently explained by the existence of spinning dust in the WIM that is traced morphologically by the  $H\alpha$  template.

A more extensive approach to the cleaning of the data has been introduced in connection with the iterative application of . A set of coefficients obtained with a specific mask is used to remove the foreground contaminations in the maps, which are then internally analysed with *all* the Galactic cuts available. This new approach allowed us to check the capability of the iterative step to recover from a poor initial cleaning of the data and still yield a credible CMB estimate. A useful figure of merit for the quality of the results is then the corresponding CMB power spectrum as compared to the *WMAP* estimation. In general, the iterative approach is very robust, although some not completely satisfactory results provide a evidence of limitations of the algorithm. Specifically, it seems that the algorithm provides poorer estimates of the CMB if applied on smaller regions of the sky.

Associated with the iterative analysis, we were able to determine the presence of a residual foreground whose spatial distribution is concentrated along the Galactic plane, with pronounced emission near the Galactic center. The extension of this residual is independent of the mask used, meaning that it is not just the effect of a poor component separation. However, it decreases in amplitude if the K and Ka band are not included in the input data set, and more so if the K-Ka map is used to compute the synchrotron coefficients. This indicates that the component has a falling spectrum with frequency, but whether it is a new physical component or simply a manifestation of a region with different spectral behaviour from the average is difficult to determine. In any case, it does confirm the utility of the K-Ka data as a better template to describe the low-frequency foreground mixture. In fact, this emission was already observed in the SMICA analysis of Patanchon et al. (2005) and is clearly consistent with the *WMAP* haze of Finkbeiner (2004a). The putative model of its spatial

distribution proposed by Dobler & Finkbeiner (2008a) is found to be in reasonable agreement with the data, although some refinements are required.

Finally, we attempted a full-sky analysis of the same data set. Even though the code then challenged to work in regions of the sky where the cross-talk among the components is very high, we find that is still able to achieve good results, particularly based on the iterative analysis of the cleaned data as originally implemented in Chapter 3. We focused our attention on the CMB reconstruction, comparing the results with those produced by a simultaneous analysis of the multi-frequency *WMAP* data with the templates. In this case, the result is slightly better. On the other hand, a direct comparison of the power spectra of the ICA CMB map with variants of the ILC approach proposed recently (HILC, NILC) suggests the need to include more spatial dependence in our analysis. Indeed, the excess of power observed on small angular scales is likely the signature of some residual structures along the Galactic plane. Consequently, to partially compensate for such structures, we defined a minimal mask to be used in a practical analysis of the derived CMB map. The resulting power spectra for our preferred data sets then agree remarkably with the best estimate of the CMB spectrum provided by the *WMAP* team.



## Chapter 5

# A refined study of the free-free frequency spectrum

### 5.1 Introduction

The analysis presented in Chapters 3 and 4 has shown that, given the foreground scaling coefficients, it is possible to derive the frequency spectra of the single foreground components.

The coefficients can generally be computed by means of different kinds of analysis, as described in Chapter 2. Therefore, it is interesting to compare the results obtained with different methods. By this point of view, the work presented by Dobler & Finkbeiner (2008a,b) and Dobler et al. (2009) and their claims about the free-free frequency spectrum are particularly interesting. They assert that the  $H\alpha$  map, generally used to trace the free-free emission, should also do so for a spinning dust emission component, and that such a model is consistent with the values of the  $H\alpha$  coefficients which they derived. Specifically, they found a *bump* in the spectral behaviour of these coefficients, which can be explained as a contribution from the spinning dust emission in the Warm Ionised Medium (WIM).

The spectrum derived from our analysis did not show any consistency with the model proposed by Dobler et al. (2009), although, depending on the mask used, it demonstrated some peculiar behaviour.

The work presented in this chapter aims to investigate in more detail the discrepancies between the results of the two analysis, taking into account their differences and the impact of them on the results themselves. Practically, we repeated our analysis conforming our working conditions to those used by Dobler et al. (2009).

### 5.2 Fitting procedure used by Dobler *et al.*

The foreground study carried out by Dobler & Finkbeiner (2008a) is a simple cross-correlation analysis (see Chapter 2), applied to the *WMAP* data previously cleaned of the CMB emission. The data are described by the equation:

$$Pa = w, \tag{5.1}$$

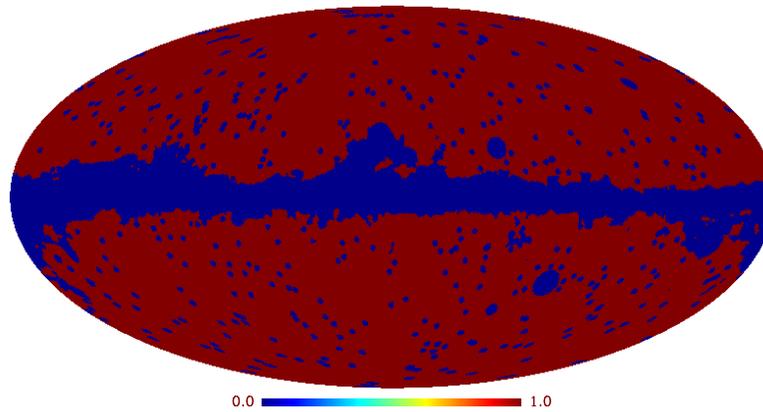


Figure 5.1: Mask of the Galactic plane produced by Dobler et al. (2009). They masked out the regions of the sky where the dust extinction at  $H\alpha$ , as estimated by the map of Schlegel et al. (1998), is  $A(H\alpha) \equiv 2.65E(B - V) > 1$  mag. Additionally, they masked out all point sources listed by the *WMAP* science team as well as the LMC, SMC, M31, Orion-Barnard’s Loop, NGC 5090, and the HII region around  $\zeta$  Oph.

where  $\mathbf{w}$  is the CMB-subtracted *WMAP* data, and  $P$  is a ‘template matrix’ whose columns consist of foreground templates, for the coefficient vector  $\mathbf{a}$ , whose entries represent the weights of the individual foregrounds. The adopted templates are the Haslam map at 408 MHz (Haslam et al., 1982), the FDS 94 GHz map by Finkbeiner et al. (1999) and the  $H\alpha$  map produced by Finkbeiner (2003) for the synchrotron, dust and free-free emission respectively. They assumed no knowledge of the foreground components spectra. The latter are derived by fitting the weights themselves.

The correlation between the cosmological signal and the foreground emissions makes the cleaning of the data mandatory, since it affects any foreground analysis and is a major source of systematic errors. However, subtracting the CMB contribution, estimated by means of the ILC, introduces a bias in the data, due to the foreground residual which affects the CMB estimation itself (see Section 2.3). Therefore, the choice of the CMB estimator can dramatically affect the inferred foreground spectra in a given fit, due to the different amount of residual which depends on the way the estimator is generated. This issue is particularly relevant for the characterisation of the free-free spectrum, given the claims about the presence of a bump.

To assess such an impact, Dobler et al. (2009) experimented with different CMB estimators, which are basically different realizations of ILC maps.

- First of all, they used the published ILC map derived by the *WMAP* team for the five-year data.
- The set of coefficients obtained for the Region 0 (according to the subdivision of the sky proposed by *WMAP* shown in Figure 2.1, basically the region off the Galactic plane) have been adopted to generate the second estimator.
- The third one is the ILC which minimizes the variance over their unmasked pixels. In their mask, which is also used throughout the later analysis, they cut out the regions of the sky where the dust

extinction at  $H\alpha$ , as estimated by the map of Schlegel *et al.* (1998), is  $A(H\alpha) \equiv 2.65E(B - V) > 1$  mag. Additionally, they masked out all point sources listed by the *WMAP* science team as well as the LMC, SMC, M31, Orion-Barnard's Loop, NGC 5090, and the HII region around  $\zeta$  Oph (see Figure 5.1). This mask covers 22.2% of the sky: in comparison the nominal mask for the temperature analysis of *WMAP* (the *KQ85* mask) excludes 22% of the pixels, although slightly differently distributed mainly along the Galactic plane .

- Furthermore, they adopted the so called ‘high frequency estimator’ that removes the dominant foregrounds (thermal dust and free-free) from the 94 GHz *WMAP* data. It is described by  $CMB = T_{94GHz} - FDS - AH\alpha$ , where the constant  $A$  is determined from the approximate free-free amplitude at 23 GHz and traced to 94 GHz assuming a power law of the form  $T \sim \nu^{-2.15}$ .
- A fifth estimator is created from the internal linear analysis of the *WMAP* data previously cleaned of the thermal dust contribution, which is assumed to be described by  $T_{dust} = (\nu/94GHz)^{1.7} \times FDS$ .
- Finally, they generated a map using the Tegmark *et al.* (2003) method, which makes use of a linear weighting of the data in which the weights depend on the multipole  $\ell$  of the spherical harmonic expansion of each of the five *WMAP* bands.

The solution to the Equation 5.1 is given by the coefficients

$$\mathbf{a} = (P/\sigma)^+ (\mathbf{w}/\sigma) \quad (5.2)$$

which minimize the quantity:

$$\left\| \frac{P}{\sigma} \mathbf{a} - \frac{\mathbf{w}}{\sigma} \right\|^2 = \frac{\|P\mathbf{a} - \mathbf{w}\|^2}{\sigma^2} \equiv \chi^2. \quad (5.3)$$

These are the coupling coefficients between the *WMAP* data and the foreground templates, by fitting which it is possible to derive the frequency spectrum of the associated emission. The chosen foreground model to fit the  $H\alpha$  coefficients is given by the contribution of free-free plus CMB and WIM spinning dust emission. For the last one, they assumed the model of Draine & Lazarian (1998b) (hereafter DL98), with the maximum of emission shifted to 40 GHz:

$$I_{mod} = F_0 \left( \frac{\nu}{23GHz} \right)^{-0.15} + D_0 \left( DL98, \nu_{peak} = 40GHz \right) + C_0 \left( \frac{\nu}{23GHz} \right)^2 a(\nu) \quad (5.4)$$

where  $a(\nu)$  is the factor which converts the  $\Delta T$  thermodynamic temperature to antenna temperature at a given frequency  $\nu$ , and  $F_0$ ,  $D_0$  and  $C_0$  are the free parameters which quantify the amplitudes of the three components. The last one is actually a component with a CMB-like spectrum (i.e. flat in thermodynamic temperature), which is included to take into account the residuals in the ILC maps used to clean the data.

### 5.3 Monte Carlo simulations

Dobler et al. (2009) have pointed out that the spectral cross correlation coefficients do depend on the CMB estimator used. However, they also stressed the fact that since they fitted out a CMB spectrum in the interpretation of the correlation spectra, the final results for the  $H\alpha$ -correlated spinning dust do not depend on the choice of the CMB estimator and, therefore, on the subtraction of the CMB itself. Nevertheless, given the discrepancy with our results (see Chapters 3 and 4), it is natural to investigate whether a fake detection of a bump in the free-free spectrum is possible, mainly as a consequence of the subtraction of the CMB emission from the data.

To answer this question, we used simulations, which are also an useful tool for addressing other interesting points:

1. whether a fake detection of a bump in the free-free spectrum is induced by a cross-talk among the components: this is the main reason why we did not study only the free-free spectrum, but we carried out the same analysis also for the other components.
2. what the error bars are for a template fit when a CMB map is pre-subtracted. From the *WMAP* noise alone, the error bars should be very small because of the high *WMAP* signal-to-noise. However, the CMB cross-correlation bias is not small, and it imprints a CMB spectrum on the derived frequency spectrum: this could explain why the error bars derived in the previous analysis (see Chapter 4) are so much larger at 94 GHz compared to 23 GHz, and generally larger than the error bars derived by Dobler et al. (2009).
3. whether not detecting a bump, but instead a rise in the spectrum, is possible when the WIM contribution is included in the foreground model: one of the results of our analysis proposed in Chapter 4 indeed, is a free-free spectrum which arises with frequency depending on the mask used to exclude the Galactic plane.

In practice, we run two sets of simulations, considering two different foreground models:

- A. The first one is a model where the free-free emission is not correlated with spinning dust. This is created using the usual three templates for synchrotron, free-free and thermal dust emission, scaled at the different frequencies with idealised spectra, i.e.  $\beta_s = 3.1$ ,  $\beta_{ff} = 2.15$ ,  $\beta_d = 1.7$ . The amplitude is chosen in order to make the resultant coupling coefficients comparable to those derived from the real data analysis. However for the dust spectrum, we added an actual spinning dust component from Cold Neutral Medium (CNM), whose model is derived using the SPDUST code developed by Ali-Haïmoud et al. (2009) and shown in Figure 5.2.
- B. The second one includes also the free-free correlated spinning dust emission from WIM, again as derived by the SPDUST code (see Figure 5.2).

As usual, the foreground model for each frequency is then added to the CMB random Gaussian realization, according to the *WMAP* best estimation of the angular power spectrum, and the noise

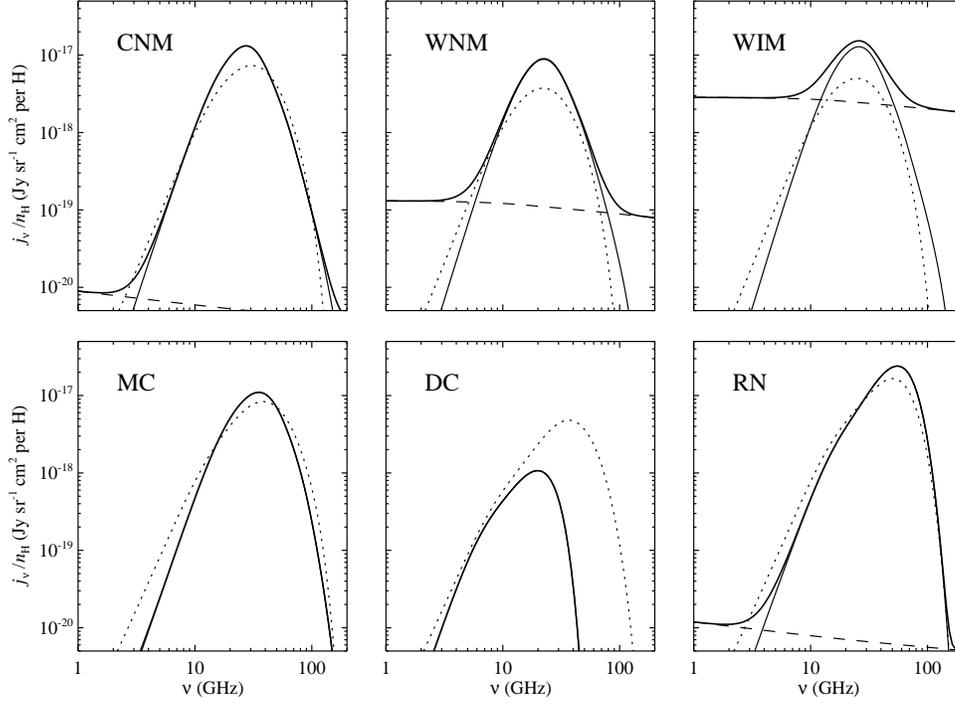


Figure 5.2: Spinning dust spectra for several environment conditions: Cold Neutral Medium (CNM), Warm Neutral Medium (WNM), Warm Ionized Medium (WIM), Molecular Cloud (MC), Dark Cloud (DC) and Reflection Nebula (RN). The environments are defined by Draine & Lazarian (1998b), in their Table 1. The thin solid line is the result of Ali-Haïmoud et al. (2009) using the SPDUST code, while dotted line is DL98 prediction, and the dashed line is the free-free emission (the adopted free-free gaunt factor is derived from Sutherland (1998)). The plot was taken from Ali-Haïmoud et al. (2009). We considered the spectra defined for the WIM and CNM. The apparent systematic increase of power around the peak frequency for the SPDUST code result is mainly due to the used grain size distribution, which has an increased number of small grains compared to that used by Draine & Lazarian (1998b).

contribution, as described in Appendix A. The maps are subsequently used to run the  $\chi^2$  templates fits in the standard way, as done in Chapters 3 and 4. Furthermore, they are internally analysed to produce an ICA and ILC estimation of the CMB which is subtracted from the simulated data themselves. Finally, also the CMB cleaned simulated data are fitted with templates with  $\chi^2$  and the  $\chi^2$  approach.

Although the real data analysis has been carried out considering different cuts of the Galactic plane, for the simulations we used only the  $Kp2$  mask, due to the fact that the pipeline is quite time consuming. For the same reason, we adopted only the  $p$  function among those generally used with  $\chi^2$ . We tested that these choices do not affect our results and conclusions.

As already mentioned, the subtracted CMB emission is estimated either by ICA or the ILC method. For the latter case, we reproduced some of the maps used by Dobler et al. (2009), which we refer to as:

- ‘noICA’: the CMB emission is traced by the ICA map computed outside the  $Kp2$  mask and it is

subtracted before the templates fit.

- ‘noILC’: the CMB emission is subtracted and traced by the ILC map computed over the whole sky, using the 12 regions defined by the *WMAP* science team.
- ‘noILCKp2’: the CMB emission is subtracted and described by the ILC map computed over the region of the sky outside the *Kp2* mask.
- ‘noILCnoThDust’: the CMB emission is subtracted and traced by the ILC map computed over the whole sky, using the 12 regions defined by the *WMAP* science team. The thermal dust emission is not included in the data.
- ‘noILCKp2noThDust’: the CMB emission is subtracted and described by the ILC map computed over the region of the sky outside the *Kp2* mask. The thermal dust emission is not included in the data.

### 5.3.1 Simulations results

An interesting cross-check on whether the modeled sky actually matches the *WMAP* observations is whether the  $\alpha$  and ILC coefficients are similar to those derived from the analysis of the observed data or not. To assess this point, we studied the distributions of the weights obtained from the internal analysis of the raw simulated data, and compared them to those derived from the analysis of the real ones. These are shown in Figure 5.3 and 5.4 respectively for the model which does not include the WIM emission and the one that includes it.

It is easy to note that the  $\alpha$  results are more spread than the ILC ones. Furthermore, in the latter case, the mean is generally not consistent with the values expected from the real data analysis: for a frequency larger than 40 GHz (Q band), the values are outside the distributions. This is not the case for  $\beta$  although this result is mainly due to the fact that the distributions are very spread: the real weights are still consistent with the mean of the distributions or, at least, they lie on the tail. This discrepancy in the width of the distributions is probably the effect of the different statistic employed by the two methods as well as the fact that the ILC imposes a constraint on the CMB spectrum. On the other side, the inconsistency of the mean of the distributions with the weights derived from the real data analysis, suggests the idea that the foreground model used for simulations does not realistically reproduce the observations. Thus, our poor knowledge about foregrounds partially compromises the usage of simulations to predict real data analysis results.

We also examined the distribution of the derived coupling coefficients of all the components with respect to the mock data, and the cross-talk among them. As stated before, although the attention was focused on the free-free component, the distribution of the synchrotron and dust coefficients have been analysed as well, in order to have a complete picture of the simulation response, and to check that the results obtained for the free-free emission were not just a consequence of a bad reconstruction of the other components. For a non-standard analysis, we corrected the values for the bias due to the CMB subtraction. This value has been derived by fitting all the coefficients to the models assumed, as follows:

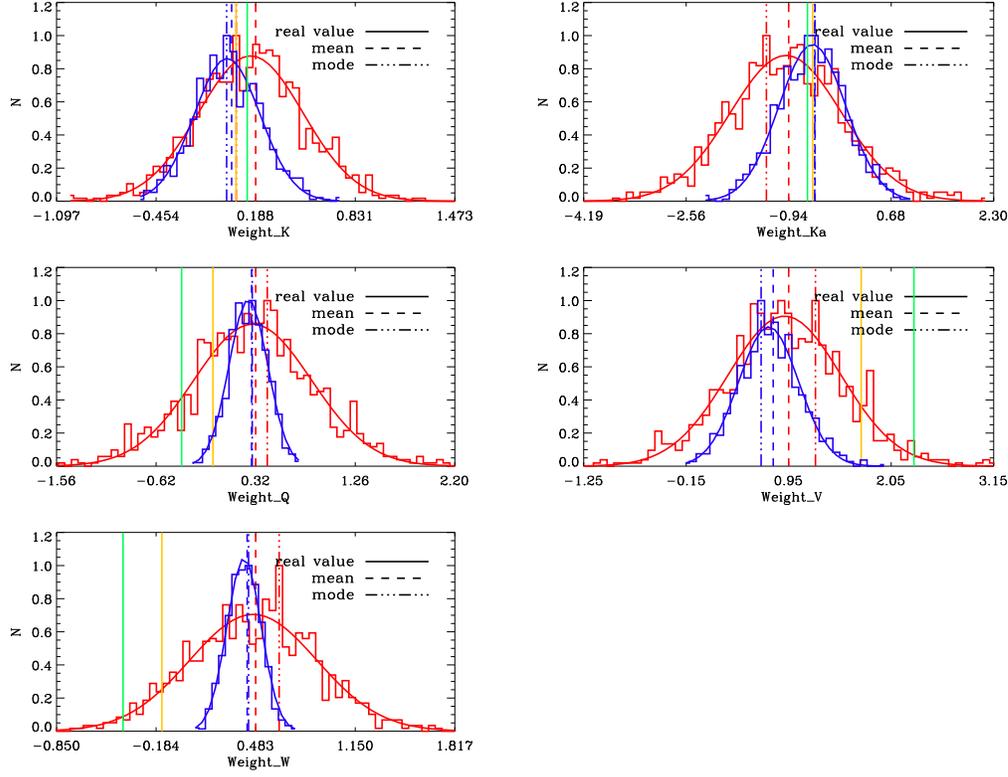


Figure 5.3: **Simulations without WIM emission.** Statistical distribution of the weights derived to compute the CMB map using (red) and the ILC code (blue). The first one is computed outside the  $Kp2$  mask and the latter has been computed using the partition of the sky proposed by the *WMAP* science team. The green (ICA) and yellow (ILC) lines show the values of the weights recovered by the real data analysis.

$$I_{mod,synchrotron} = A_{sync} \left( \frac{\nu}{23GHz} \right)^{-3.1} + C_0 a(\nu) \left( \frac{\nu}{23GHz} \right)^2, \quad (5.5)$$

$$I_{mod,dust} = A_{ThDust} \left( \frac{\nu}{94GHz} \right)^{1.7} + A_{CNM}(CNM) + C_0 a(\nu) \left( \frac{\nu}{23GHz} \right)^2 \quad (5.6)$$

where  $A_{sync}$ ,  $A_{ThDust}$  and  $A_{CNM}$  are the amplitudes of the synchrotron, the thermal and spinning dust emissions respectively, and  $C_0$  is the offset. Equation 5.4 is the equivalent one for the free-free emission.

The coefficients of the synchrotron and dust emission derived from the standard analysis (without subtraction of the CMB) have a symmetric and not biased distribution. In the cases where the CMB is subtracted, instead, there is a bias which can be significant, if compared with the error on the mean. This is probably the consequence of a cross-talk between the two emissions, which is further enhanced by the bias introduced by the ICA and ILC maps. Such a deviation is not present in the distributions of the free-free coefficients, meaning that they are not involved in the cross-talk.

Finally, comparing the coupling coefficients obtained from simulations and those derived from the real data analysis (see Section 5.4), we detected a disagreement between them. When the CMB is pre-

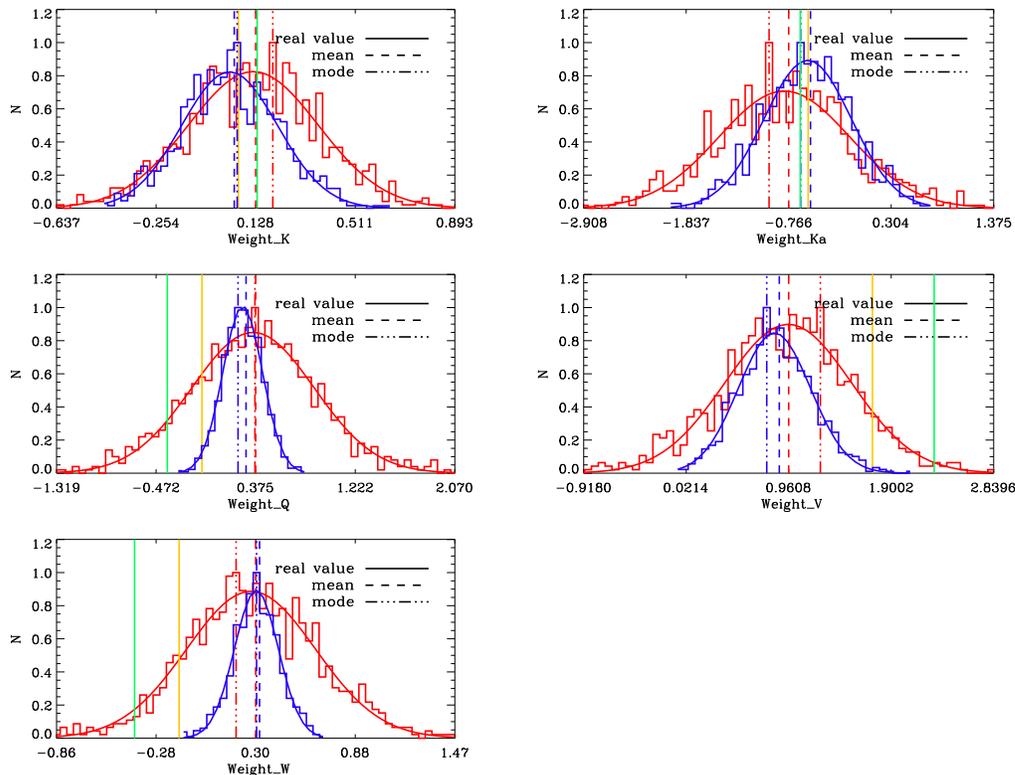


Figure 5.4: **Simulations with WIM emission.** Statistical distribution of the weights derived to compute the CMB map using  $\chi^2$  (red) and the ILC code (blue). The first one is computed outside the  $Kp2$  mask and the latter has been computed using the partition of the sky proposed by the *WMAP* science team. The green (ICA) and yellow (ILC) lines show the values of the weights recovered by the real data analysis.

subtracted, the coefficients are generally not consistent with the real ones, in both the cases where we include the WIM emission and we do not include it, and with both the  $\chi^2$  and  $\chi^2$  methods. This is a further indication of the fact that the adopted foreground model is not perfectly consistent with the real observations. The presence of the bias in the CMB tracers makes the discrepancy even larger.

A critical issue is to disentangle the different contributions encoded in the free-free coefficients following the prescription proposed by Dobler et al. (2009) for the  $\chi^2$  and ILC corrected data. For both the sets of simulations, we fitted the coefficients with both the models, looking for the genuine free-free emission as well as for the correlated spinning dust component. In practice, in the case where the model A is adopted, we checked that there were no fake detections: the mean of the coefficients for the spinning dust component should be consistent with zero. On the other side, when the B model is used, fitting the coefficients under the hypothesis that they can be described as purely free-free, is interesting in order to see how the fits behave. In other words, in the first case we can test whether we can obtain a bump when there is none, and, in the case where there is a bump, whether we do not have evidence of this.

Figures 5.5, 5.6, 5.7 show the spectrum behaviour derived from the scaling coefficients for the three components when using the real input models. It is clear that for all the cases of analysis, the spectral

behaviour for both the components is recovered according to the expectations. Specifically the bump in the free-free spectrum is visible, if the WIM emission is actually present in the data. This is true for either the case where the CMB contribution is subtracted from the data or not. Therefore, the subtraction of the CMB contribution does not seem to induce any artifact on the free-free spectrum, excluding the hypothesis that the bump is due to the contamination of the CMB map by residual foregrounds. Indeed, the simulations where the WIM contribution is not included in the foreground model, return spectra perfectly consistent with the theoretical expectations, also when the CMB is subtracted from the data. This is why we retained only the fit with the model used to simulate the data.

Looking at the plots of the frequency spectra, it is also easy to note that, when the CMB emission is not subtracted from the data, the uncertainty of the coefficients is much larger than in the other cases, due to the cross-talk between the CMB and the templates. Thus, taking into account the amplitude of the error bars, several scenarios are possible for the spectral behaviour of the free-free coupling coefficients.

Furthermore there are not dramatic differences between the results obtained with  $\chi^2$  analysis. However, there are cases of analysis where  $\chi^2$  seems to be unstable: when the ILC map (created according to the *WMAP* prescriptions) is assumed as CMB tracer, the algorithm sometimes fails in the reconstruction of the components. The same happens when the ILC map is computed using data without contribution from the thermal dust emission.

### 5.3.2 Fit parameters

Tables 5.1, 5.2, 5.3 and 5.4 show the values of the mean, the mode and the errors of the parameters which define the foreground model used to fit the simulated data. We show the results for the three or four foreground components, depending on the model adopted, and different cases of analysis.

The statistical properties of the fit parameters reflect those of the coefficients used to derive them. In the case of the free-free, the distributions are symmetric and the mean is perfectly in agreement with the input values. Some odd results however, are obtained when the ILC map is used to trace the CMB emission, both in the case where it is computed with and without the thermal dust emission. For what concerns the other two components, instead, the distributions of the fit parameters are fine if derived from the standard analysis, but biased, even though symmetric, in the other cases. All these results are consistent with what already pointed out about the cross-talk between the two components.

The offset parameter, included when the CMB is pre-subtracted from the data, is on average consistent with zero. However, the contribution for the single simulations can be significant.

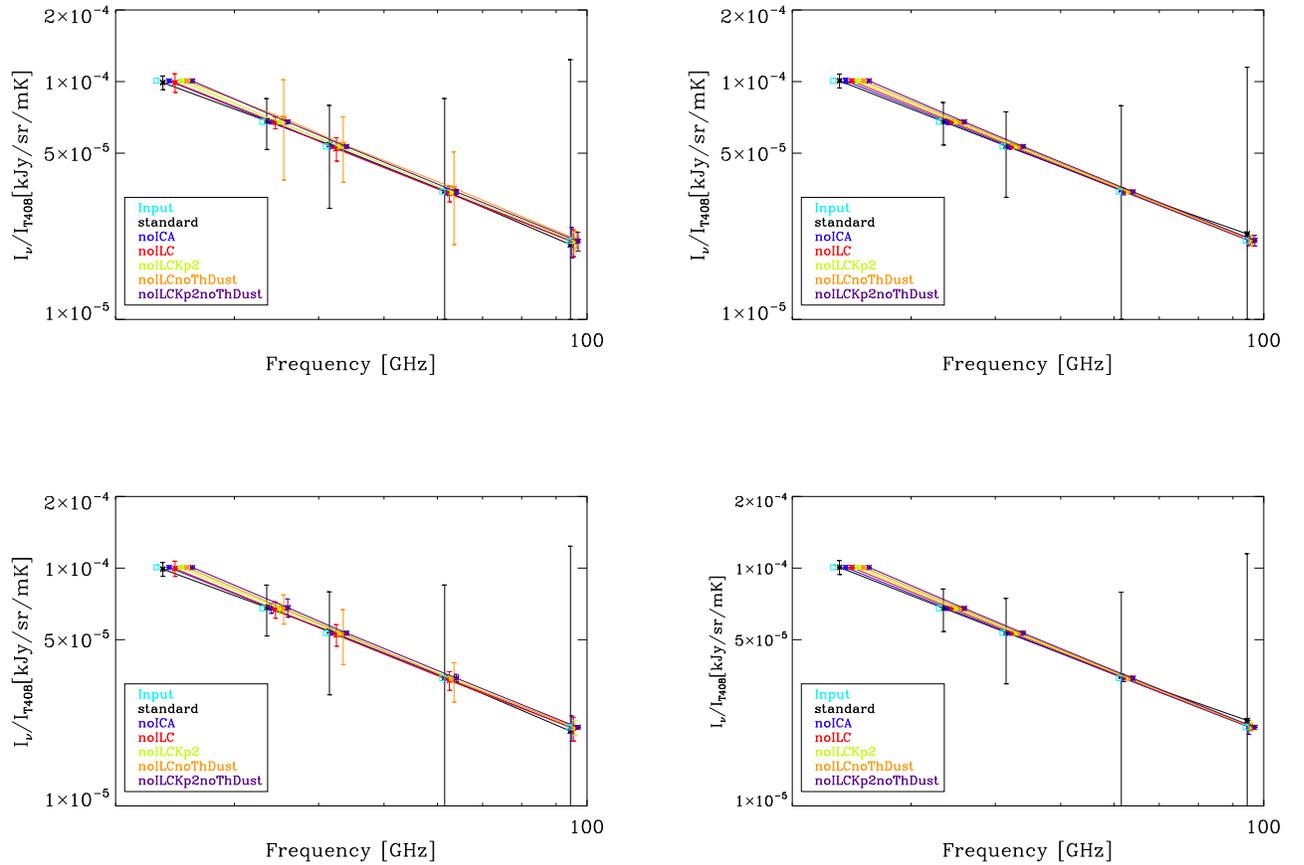


Figure 5.5: **Synchrotron emission.** Mean spectral behaviour of the synchrotron emission computed from 1000 simulations. The data are simulated as composed by a superposition of the CMB, synchrotron emission, dust emission (which includes the anomalous component as due to the CNM emission), and free-free emission. For the latter component, we considered both the cases where the WIM emission is added to the foreground model (top panels) and where it is not (bottom panels). The analysis to derive the coefficients is performed using the code (left plot) and the  $\chi^2$  analysis (right plot). We have considered only the Kp2 mask.

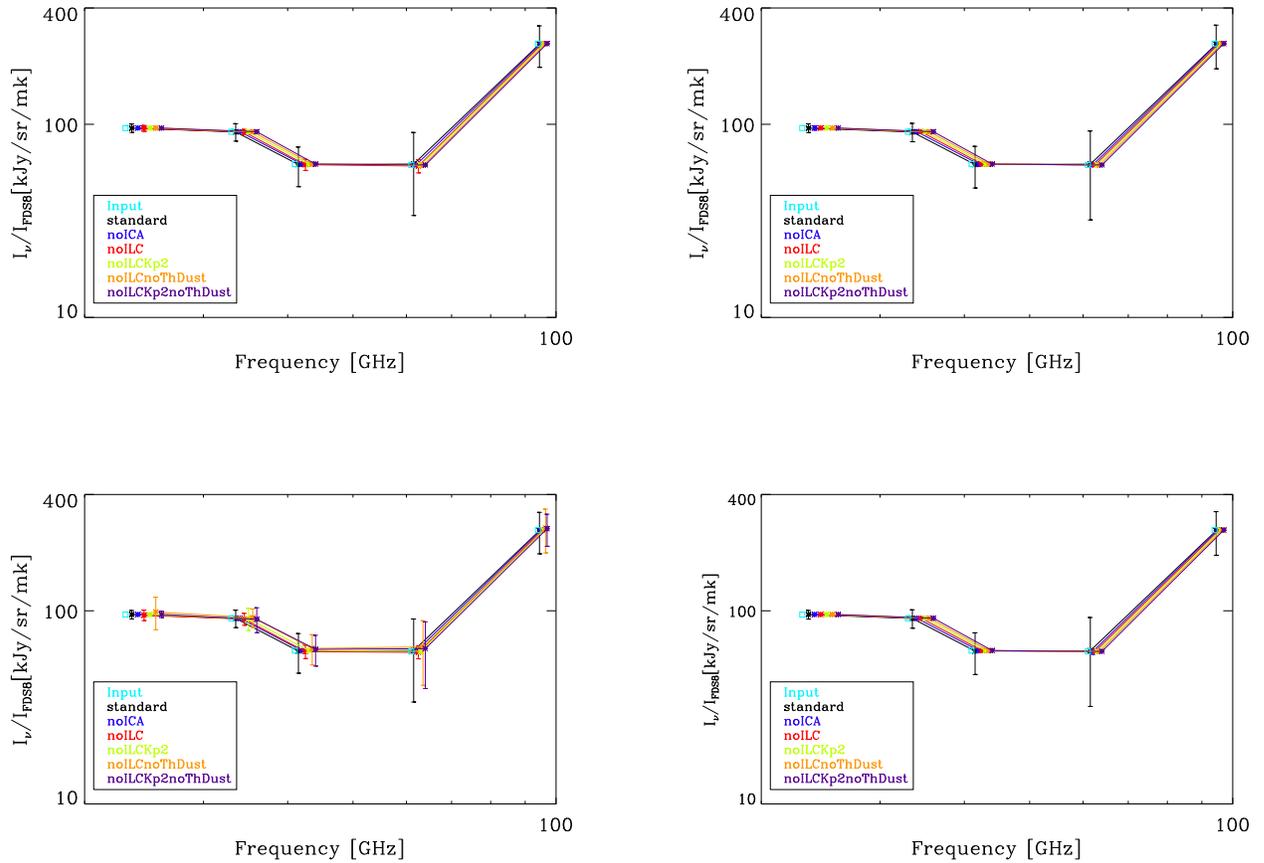


Figure 5.6: **Dust emission.** Mean spectral behaviour of the dust emission computed from 1000 simulations. The data are simulated as composed by a superposition of the CMB, synchrotron emission, dust emission (which includes the anomalous component as due to the CNM emission), and free-free emission. For the latter component, we considered both the cases where the WIM emission is added to the foreground model (top panels) and where it is not (bottom panels). The analysis to derive the coefficients is performed using the `code` (left plot) and the  $\chi^2$  analysis (right plot). We have considered only the Kp2 mask.

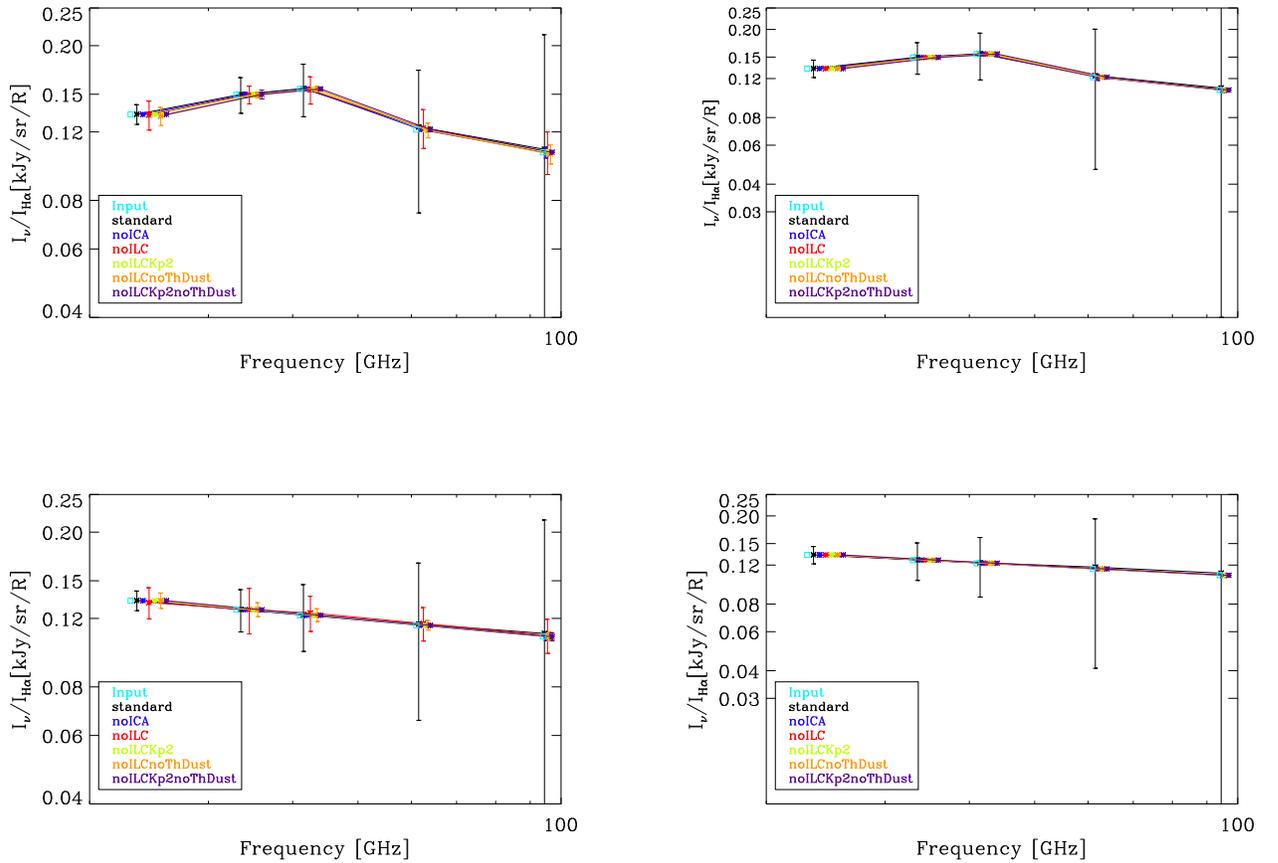


Figure 5.7: **Free-free emission.** Mean spectral behaviour of the free-free emission computed from 1000 simulations. The data are simulated as composed by a superposition of the CMB, synchrotron emission, dust emission (which includes the anomalous component as due to the CNM emission), and free-free emission. For the latter component, we considered both the cases where the WIM emission is added to the foreground model (top panels) and where it is not (bottom panels). The analysis to derive the coefficients is performed using the code (left plot) and the  $\chi^2$  analysis (right plot). We have considered only the Kp2 mask.

	$\chi^2$ analysis							
	mean	$\sigma$	$\sigma_{MPFIT}$	mode	mean	$\sigma$	$\sigma_{MPFIT}$	mode
<b>Synchrotron</b>								
$A_{sync}$ [expected value 6.200]								
<i>standard</i>	6.222	0.687	0.008	5.958	6.202	0.576	0.008	5.935
<i>ICA</i>	6.190	0.021	0.032	6.190	6.193	0.007	0.032	6.192
<i>ILC</i>	6.186	0.118	0.028	6.189	6.193	0.006	0.028	6.190
<i>ILCKp2</i>	6.190	0.019	0.026	6.187	6.193	0.006	0.026	6.190
<i>ILCnoThDust</i>	6.183	0.027	0.049	6.176	6.192	0.008	0.049	6.195
<i>ILCKp2noThDust</i>	6.189	0.016	0.052	6.193	6.192	0.009	0.052	6.195
offset								
<i>ICA</i>	0.031	0.449	0.004	0.134	0.030	0.446	0.004	-0.059
<i>ILC</i>	0.135	0.318	0.006	0.110	0.184	0.354	0.006	0.198
<i>ILCKp2</i>	0.038	0.214	0.006	0.076	0.038	0.214	0.006	0.010
<i>ILCnoThDust</i>	-0.065	0.313	0.003	-0.062	-0.079	0.355	0.003	-0.083
<i>ILCKp2noThDust</i>	-0.010	0.213	0.004	-0.033	-0.010	0.214	0.004	0.042
<b>Thermal Dust + CNM</b>								
$A_{CNM}$ [expected value 5.793]								
<i>standard</i>	5.788	0.309	0.003	5.850	5.787	0.320	0.003	5.627
<i>ICA</i>	5.795	0.010	0.012	5.797	5.796	0.006	0.012	5.797
<i>ILC</i>	5.794	0.015	0.010	5.793	5.796	0.006	0.010	5.794
<i>ILCKp2</i>	5.795	0.009	0.010	5.791	5.796	0.006	0.010	5.797
<i>ILCnoThDust</i>	5.796	0.008	0.024	5.789	5.796	0.005	0.024	5.794
<i>ILCKp2noThDust</i>	5.795	0.007	0.026	5.790	5.796	0.005	0.026	5.793
$A_{ThDust}$ [expected value 0.963]								
<i>standard</i>	0.951	0.465	0.004	1.125	0.945	0.494	0.004	0.701
<i>ICA</i>	0.960	0.008	0.011	0.963	0.960	0.005	0.011	0.961
<i>ILC</i>	0.960	0.011	0.005	0.962	0.960	0.005	0.005	0.960
<i>ILCKp2</i>	0.960	0.008	0.005	0.962	0.960	0.005	0.005	0.960
<i>ILCnoThDust</i>	0.961	0.008	0.012	0.954	0.960	0.005	0.012	0.960
<i>ILCKp2noThDust</i>	0.960	0.009	0.013	0.961	0.960	0.005	0.013	0.961
offset								
<i>ICA</i>	-0.013	0.298	0.006	0.111	-0.012	0.297	0.006	0.103
<i>ILC</i>	-0.025	0.105	0.005	-0.027	-0.033	0.043	0.005	-0.054
<i>ILCKp2</i>	-0.007	0.034	0.004	-0.005	-0.006	0.030	0.004	-0.011
<i>ILCnoThDust</i>	-0.014	0.151	0.015	-0.005	0.019	0.041	0.015	0.029
<i>ILCKp2noThDust</i>	0.003	0.032	0.016	0.003	0.004	0.030	0.016	-0.002

Table 5.1: **Simulations with WIM emission.** Values of the mean, the mode and the errors of the fit parameters obtained from 1000 simulations for the three foreground components and different cases of analysis. See next table for further details.

	$\chi^2$ analysis							
	mean	$\sigma$	$\sigma_{MPFIT}$	mode	mean	$\sigma$	$\sigma_{MPFIT}$	mode
<b>Free-free + WIM</b>								
$F_0$ [expected value 8.200]								
<i>standard</i>	8.039	0.418	0.006	8.044	8.043	0.645	0.006	7.983
<i>ICA</i>	8.031	0.013	0.026	8.031	8.033	0.008	0.026	8.035
<i>ILC</i>	8.031	0.015	0.021	8.030	8.034	0.008	0.021	8.036
<i>ILCKp2</i>	8.032	0.012	0.020	8.024	8.034	0.008	0.020	8.034
<i>ILCnoThDust</i>	8.033	0.015	0.041	8.031	8.034	0.009	0.041	8.033
<i>ILCKp2noThDust</i>	8.032	0.015	0.044	8.026	8.034	0.009	0.044	8.033
$D_0$ [expected value 0.150]								
<i>standard</i>	0.151	0.068	0.001	0.131	0.152	0.105	0.001	0.144
<i>ICA</i>	0.150	0.002	0.003	0.149	0.150	0.002	0.003	0.150
<i>ILC</i>	0.150	0.003	0.003	0.150	0.150	0.001	0.003	0.150
<i>ILCKp2</i>	0.150	0.002	0.003	0.150	0.150	0.001	0.003	0.150
<i>ILCnoThDust</i>	0.150	0.002	0.006	0.148	0.150	0.001	0.006	0.150
<i>ILCKp2noThDust</i>	0.150	0.002	0.006	0.150	0.150	0.001	0.006	0.150
offset								
<i>ICA</i>	0.012	0.435	0.004	-0.018	0.010	0.433	0.004	-0.026
<i>ILC</i>	0.006	0.133	0.005	0.043	-0.007	0.089	0.005	0.008
<i>ILCKp2</i>	-0.006	0.082	0.005	0.003	-0.007	0.082	0.005	-0.001
<i>ILCnoThDust</i>	0.003	0.098	0.003	-0.029	-0.005	0.090	0.003	-0.027
<i>ILCKp2noThDust</i>	-0.013	0.083	0.004	-0.016	-0.013	0.083	0.004	-0.067

Table 5.2: **Simulations with WIM emission.** Values of the mean, the mode and the errors of the fit parameters obtained from 1000 simulations for the three foreground components and different cases of analysis. The free parameters of the fitted model are the amplitude of the synchrotron emission, the amplitude of the anomalous component and the thermal dust emission, and finally the amplitudes of the free-free emission and the WIM contribution: their expected values are shown in brackets. An offset has been included among the parameters, when the CMB contribution is subtracted from the simulated data: it takes into account the residuals of the CMB map itself. The spectral indexes are fixed at -3.1, 1.7 and -2.15 for the synchrotron, the thermal dust and the free-free emission respectively. The WIM and CNM spectrum are derived from the SPDUST code. The fitted coefficients are derived with  $\chi^2$  and the  $\chi^2$  analysis. To perform the fit we used the IDL MPFIT function which returns the value of the error for each parameter of the model used to fit the data. For comparison, we show the error bars derived as the mean of the MPFIT estimations.

	$\chi^2$ analysis							
	mean	$\sigma$	$\sigma_{MPFIT}$	mode	mean	$\sigma$	$\sigma_{MPFIT}$	mode
<b>Synchrotron</b>								
$A_{sync}$ [expected value 6.200]								
<i>standard</i>	6.222	0.680	0.008	5.879	6.201	0.570	0.008	5.939
<i>ICA</i>	6.191	0.022	0.041	6.194	6.193	0.012	0.041	6.193
<i>ILC</i>	6.186	0.072	0.032	6.192	6.193	0.008	0.031	6.191
<i>ILCKp2</i>	6.190	0.016	0.035	6.192	6.193	0.010	0.035	6.189
<i>ILCnoThDust</i>	6.189	0.027	0.048	6.195	6.193	0.014	0.048	6.192
<i>ILCKp2noThDust</i>	6.191	0.024	0.048	6.188	6.193	0.014	0.048	6.190
offset								
<i>ICA</i>	0.071	0.426	0.010	-0.068	0.072	0.423	0.010	0.033
<i>ILC</i>	0.210	0.241	0.009	0.221	0.271	0.287	0.009	0.226
<i>ILCKp2</i>	0.062	0.204	0.009	0.024	0.061	0.204	0.009	0.040
<i>ILCnoThDust</i>	0.083	0.304	0.010	0.077	0.101	0.337	0.010	0.087
<i>ILCKp2noThDust</i>	0.023	0.211	0.010	-0.006	0.022	0.212	0.010	-0.001
<b>Thermal Dust + CNM</b>								
$A_{CNM}$ [expected value 5.793]								
<i>standard</i>	5.788	0.303	0.005	5.851	5.787	0.313	0.005	5.633
<i>ICA</i>	5.794	0.013	0.010	5.793	5.795	0.008	0.010	5.796
<i>ILC</i>	5.795	0.023	0.009	5.790	5.795	0.007	0.009	5.797
<i>ILCKp2</i>	5.794	0.011	0.009	5.789	5.795	0.007	0.009	5.794
<i>ILCnoThDust</i>	5.794	0.014	0.011	5.792	5.795	0.009	0.011	5.794
<i>ILCKp2noThDust</i>	5.794	0.014	0.011	5.794	5.795	0.009	0.011	5.794
$A_{ThDust}$ [expected value 0.963]								
<i>standard</i>	0.950	0.489	0.003	1.062	0.944	0.520	0.003	0.806
<i>ICA</i>	0.960	0.008	0.014	0.959	0.960	0.005	0.014	0.959
<i>ILC</i>	0.960	0.011	0.011	0.960	0.960	0.005	0.011	0.960
<i>ILCKp2</i>	0.960	0.008	0.012	0.959	0.960	0.005	0.012	0.960
<i>ILCnoThDust</i>	0.960	0.008	0.016	0.962	0.960	0.005	0.016	0.959
<i>ILCKp2noThDust</i>	0.960	0.008	0.016	0.955	0.960	0.005	0.016	0.960
offset								
<i>ICA</i>	-0.015	0.279	0.007	0.042	-0.015	0.279	0.007	0.040
<i>ILC</i>	-0.046	0.093	0.005	-0.063	-0.047	0.039	0.005	-0.069
<i>ILCKp2</i>	-0.011	0.029	0.006	-0.008	-0.010	0.029	0.006	-0.014
<i>ILCnoThDust</i>	-0.008	0.046	0.008	-0.026	-0.009	0.043	0.008	-0.008
<i>ILCKp2noThDust</i>	-0.002	0.032	0.008	-0.006	-0.001	0.031	0.008	0.001

Table 5.3: **Simulations without WIM emission.** Values of the mean, the mode and the errors of the fit parameters obtained from 1000 simulations for the three foreground components and different cases of analysis. See next table.

	$\chi^2$ analysis							
	mean	$\sigma$	$\sigma_{MPFIT}$	mode	mean	$\sigma$	$\sigma_{MPFIT}$	mode
	<b>Free-free</b>							
	$F_0$ [expected value 8.200]							
<i>standard</i>	8.211	0.679	0.005	8.377	8.216	1.050	0.005	8.134
<i>ICA</i>	8.198	0.023	0.009	8.193	8.200	0.015	0.027	8.200
<i>ILC</i>	8.194	0.038	0.022	8.196	8.200	0.012	0.021	8.195
<i>ILCKp2</i>	8.198	0.020	0.024	8.188	8.200	0.014	0.024	8.197
<i>ILCnoThDust</i>	8.198	0.026	0.030	8.180	8.200	0.018	0.009	8.198
<i>ILCKp2noThDust</i>	8.199	0.027	0.029	8.180	8.200	0.018	0.029	8.197
	offset							
<i>ICA</i>	0.008	0.418	0.009	0.051	0.009	0.419	0.009	0.031
<i>ILC</i>	-0.011	0.129	0.008	-0.033	-0.013	0.086	0.008	-0.020
<i>ILCKp2</i>	-0.016	0.080	0.008	-0.056	-0.016	0.079	0.008	-0.003
<i>ILCnoThDust</i>	-0.023	0.087	0.009	-0.051	-0.025	0.086	0.009	-0.044
<i>ILCKp2noThDust</i>	-0.033	0.081	0.009	-0.041	-0.033	0.080	0.009	-0.016

Table 5.4: **Simulations without WIM emission.** Values of the mean, the mode and the errors of the fit parameters obtained from 1000 simulations for the three foreground components and different cases of analysis. The free parameters of the fitted model are the amplitude of the synchrotron emission, the amplitude of the anomalous component and the thermal dust emission, and finally the amplitudes of the free-free emission: their expected values are shown in brackets. An offset has been included among the parameters, when the CMB contribution is subtracted from the simulated data: it takes into account the residuals of the CMB map itself. The spectral indexes are fixed at -3.1, 1.7 and -2.15 for the synchrotron, the thermal dust and the free-free emission respectively. The CNM spectrum is derived from the SPDUST code. The fitted coefficients are derived with  $\chi^2$  analysis. To perform the fit, we used the IDL MPFIT function which returns the value of the error for each parameter of the model used to fit the data. For comparison, we show the error bars derived as the mean of the MPFIT estimations.

### 5.3.3 Noise impact simulations

Given the cross-talk between the CMB and the foregrounds, it is interesting to quantify the role of the noise in such a coupling. We did that by running a new set of simulations, starting from those described above. We chose a CMB realisation for which the recovered coefficients seemed quite bad compared to the expected input values, and we kept it fixed, only varying the noise (see the fixed coefficients in Tables 5.5 and 5.6 compared to the original ones). We considered only the case where we did not include the WIM emission and we run a standard analysis, i.e. we did not subtract the CMB from the simulated data.

The difference between the mean value of the coefficients of the simulations with the original input ones represents the cross-talk for this CMB realisation. Furthermore, the dispersion of the values around the mean should give an idea of the impact of the noise (see Tables 5.5 and 5.6). This is two orders of magnitude lower than the one recovered by a standard analysis, meaning that the noise contributes very little to the variation of the coefficients (see previous tables). Moreover, we conclude that the uncertainty in the estimation of the coefficients, over 1000 simulations with different CMB realizations, is mainly driven by the CMB cross-talk with the foreground templates.

	input coeff.	fixed coeff.	mean	$\sigma$	$\sigma_{mean}$	mode
<b>Synchrotron</b>						
$K$	6.2000	6.1006	6.0976	0.0064	0.0002	6.0955
$Ka$	2.0247	1.9442	1.9353	0.0063	0.0002	1.9343
$Q$	1.0330	0.9428	0.9473	0.0057	0.0002	0.9478
$V$	0.3015	0.2292	0.2206	0.0066	0.0002	0.2212
$W$	0.0789	0.0080	0.0089	0.0070	0.0002	0.0050
<b>Thermal Dust + CNM</b>						
$K$	5.8808	5.6120	5.6052	0.0037	0.0001	5.6051
$Ka$	2.7336	2.4639	2.4638	0.0035	0.0001	2.4629
$Q$	1.2040	0.9377	0.9376	0.0032	0.0001	0.9357
$V$	0.5387	0.2849	0.2858	0.0039	0.0001	0.2854
$W$	0.9635	0.7447	0.7409	0.0040	0.0001	0.7395
<b>Free-free</b>						
$K$	8.2000	10.1160	10.1256	0.0087	0.0003	10.1264
$Ka$	3.7733	5.6689	5.6717	0.0084	0.0003	5.6717
$Q$	2.3662	4.2438	4.2379	0.0076	0.0003	4.2378
$V$	1.0071	2.7687	2.7839	0.0092	0.0003	2.7850
$W$	0.3975	1.9435	1.9601	0.0093	0.0003	1.9554

Table 5.5: Statistical properties of the coefficients distribution derived from the simulations which have been run to test the noise impact. The sky is generated without WIM emission and fixing the CMB emission to be the same for all the realizations. The noise is the only contribution which changes from one realization to the other. The coefficients are computed using the  $\text{Kp2}$  mask and following a standard analysis with the  $\text{Kp2}$  mask. The expected values are shown in the first column of the table.

	$\chi^2$ analysis						
	input coeff.	fixed coeff.	mean	$\sigma$	$\sigma_{anal}$	$\sigma_{mean}$	mode
<b>Synchrotron</b>							
<i>K</i>	6.2000	6.2210	6.2209	0.0037	0.0082	0.0001	6.2197
<i>Ka</i>	2.0247	2.0506	2.0500	0.0036	0.0086	0.0001	2.0501
<i>Q</i>	1.0330	1.0594	1.0592	0.0033	0.0069	0.0001	1.0582
<i>V</i>	0.3015	0.3281	0.3270	0.0040	0.0135	0.0001	0.3249
<i>W</i>	0.0789	0.1045	0.1021	0.0042	0.0194	0.0001	0.1030
<b>Thermal Dust + CNM</b>							
<i>K</i>	5.8808	5.7031	5.7001	0.0025	0.0028	0.0001	5.6991
<i>Ka</i>	2.7336	2.5534	2.5535	0.0025	0.0030	0.0001	2.5532
<i>Q</i>	1.2040	1.0262	1.0258	0.0023	0.0024	0.0001	1.0258
<i>V</i>	0.5387	0.3724	0.3693	0.0026	0.0046	0.0001	0.3689
<i>W</i>	0.9635	0.8126	0.8141	0.0029	0.0067	0.0001	0.8140
<b>Free-free</b>							
<i>K</i>	8.2000	10.9841	10.9940	0.0053	0.0056	0.0002	10.9944
<i>Ka</i>	3.7733	6.5284	6.5261	0.0053	0.0059	0.0002	6.5243
<i>Q</i>	2.3662	5.0890	5.0794	0.0050	0.0047	0.0002	5.0785
<i>V</i>	1.0071	3.5813	3.5819	0.0056	0.0092	0.0002	3.5778
<i>W</i>	0.3975	2.6492	2.6623	0.0063	0.0133	0.0002	2.6630

Table 5.6: Statistical properties of the coefficients distribution derived from the simulations which have been run to test the noise impact. The sky is generated without WIM emission and fixing the CMB emission to be the same for all the realizations. The noise is the only contribution which changes from one realization to the other. The coefficients are computed using  $\chi^2$  and following a standard analysis with the Kp2 mask. The expected values are shown in the first column of the table.

## 5.4 Real data analysis

Once the analysis with simulations has been tested, we repeated it with real observations, the data set being the same as in the previous chapters, though cleaned of the CMB emission. This is evaluated by internally analysing the *WMAP* 5-year data with  $\chi^2$ .

Both the non-linear function  $p$  and  $g$  have been used: when one of these has been employed for the internal analysis, it has been kept for the templates fit as well.

Furthermore, as before, a simple  $\chi^2$  analysis has been used as a reference for the  $\chi^2$  results. Such a comparison seems to emphasize a significant instability of  $\chi^2$  in performing the components separation with respect to the simple  $\chi^2$  analysis. This instability is confirmed by the fact that the results depend on the non-linear function used to perform the component separation. The algorithm is more stable using  $g$ , rather than  $p$ . This suggests that the instability is generated by the internal analysis of the data and propagated to the templates fit, subtracting the CMB map from the data. Indeed, the work done by Maino et al. (2007) showed that the internal analysis of CMB observations with  $\chi^2$  is more stable when the  $g$  function is used.

In order to reduce the instability of the algorithm, we have derived the CMB estimation using a full-sky analysis, and applied a mask of the Galactic plane only for the templates fit. It allows a better

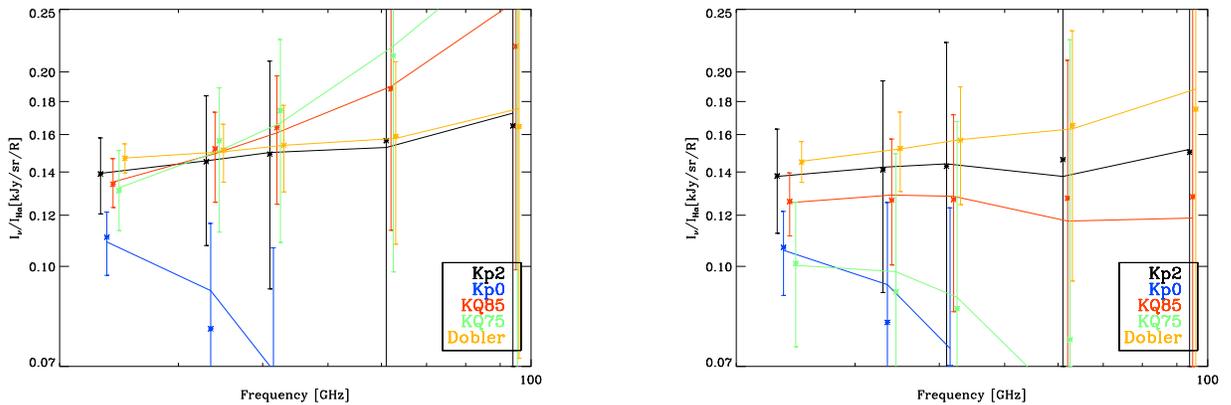


Figure 5.8: We show the free-free coupling coefficients between the *WMAP* data and the  $H\alpha$  template, as computed by using the non-linear functions  $p$  (left) and  $g$  (right). Different colors are used for the different masks applied to the data. We show the results obtained with all the masks provided by the *WMAP* science team and the one adopted by Dobler et al. (2009). The asterisk shows the values of the coefficients, while the line shows the best fit derived using the model proposed by Dobler et al. (2009) (see Equation 5.4).

performance of , thanks to the fact that in these conditions it has a larger statistic sample (i.e. a larger number of pixels is available to perform the component separation).

In Figure 5.8, we report, as a reference, the frequency spectrum of the  $H\alpha$  coupling coefficients derived in the previous chapter. As already mentioned, the spectral behaviour is not necessarily in favor of the presence of a bump: the spectrum is generally flat, as expected, or rising with frequency, strongly depending on the mask used in the analysis and on the non-linear function. However, the large error bars make the spectrum still consistent with the theoretical expected one as well as with other completely different scenarios. This means that the presence of an  $H\alpha$  correlated additional component can not be ruled out: the rising spectrum might still be a different signature of the contribution from spinning dust grains from the WIM, as claimed by Dobler et al. (2009).

When the *WMAP* data are cleaned of the CMB contribution, instead, some hints of a bump appear in the spectra: it is visible in Figure 5.9, where we show the free-free frequency spectrum derived using different masks of the plane.

We tested that the shape of the spectrum does not depend on the template used to trace the free-free emission: we performed the analysis using the  $H\alpha$  templates assembled by Finkbeiner et al. (1999), but also the map produced by Dickinson et al. (2003). Furthermore, we checked that the bump is not induced by the correction for the dust absorption applied to the  $H\alpha$  used in the analysis. Here, we did not use any correction, whose introduction would only decrease the amplitude of the coefficients. These results are all in agreement with what found by Dobler et al. (2009).

They also studied the dependence of the spectrum on the extension of the cut of the Galactic plane and they did not find any significant variation of their results, apart from a change in amplitude of the coupling coefficients. Similarly, we see such a variation in amplitude. However, the effect is visible also in terms of irregularities in the global spectrum. These are actually the signature of a bad performance

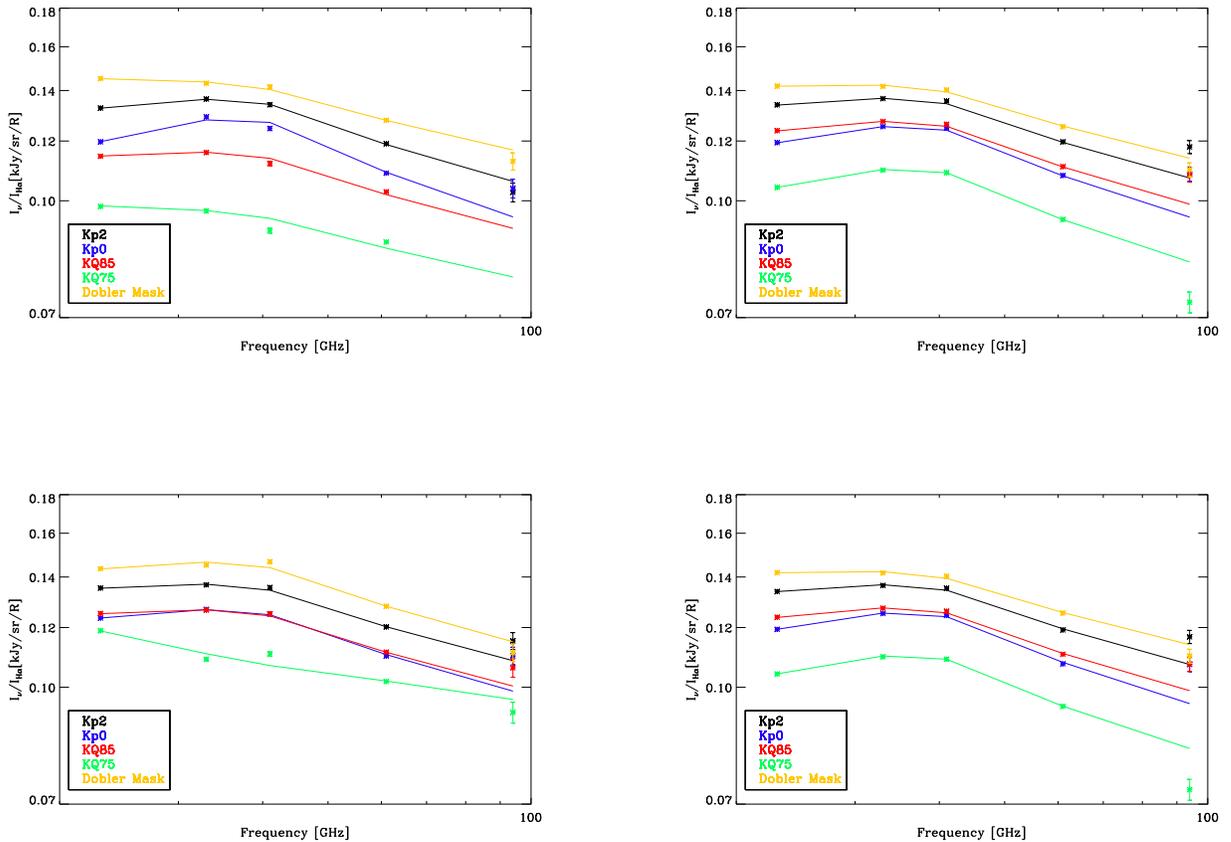


Figure 5.9: We show the free-free coupling coefficients between the *WMAP* data and the  $H\alpha$  template, as computed by  $\chi^2$  (left) and the  $\chi^2$  analysis (right). The CMB emission is subtracted from the data and estimated with a blind analysis of the full-sky five maps: we excluded only the point sources. Different colors are related to different masks applied to the data for the templates fit. We show the results obtained with all the masks provided by the *WMAP* science team and the one adopted by Dobler et al. (2009). The analysis has been performed with both the  $p$  (upper panels) and  $g$  (lower panels) functions. The asterisk shows the values of the coefficients, while the line shows the best fit derived using the model proposed by Dobler et al. (2009) (see Equation 5.4).

of  $\chi^2$  in separating the components: what is left after the CMB subtraction from the data, is not well disentangled in the different components. Specifically, it is hard to distinguish the dust components from everything else. The situation is different if the analysis is performed by minimizing the  $\chi^2$ : the spectra are more regular and the results are generally more consistent with those derived by Dobler et al. (2009).

Finally, the free-free coupling coefficients have been fitted to the model proposed by Dobler et al. (2009) and defined in Equation 5.4: the values of the parameter which describe the model are shown in Table 5.7, while the resultant curves are plotted over the coefficients in Figure 5.9.

Consistently with the picture derived looking at the frequency spectra,  $\chi^2$  sometimes returns negative values for the  $D_0$  parameter. It happens for the *KQ85* and *KQ75* maps when the analysis

<b>Free-Free -noICA</b>					
– function p					
	$Kp2$	$Kp0$	$KQ85$	$KQ75$	<i>Dobler Mask</i>
$F_0$	$0.1307 \pm 0.0002$	$0.1171 \pm 0.0002$	$0.1132 \pm 0.0002$	$0.0976 \pm 0.0002$	$0.1440 \pm 0.0002$
$C_0$	$0.0005 \pm 0.0001$	$0.0009 \pm 0.0001$	$0.0026 \pm 0.0001$	$0.0030 \pm 0.0001$	$-0.0005 \pm 0.0001$
$D_0$	$0.1560 \pm 0.0043$	$0.2147 \pm 0.0043$	$0.1043 \pm 0.0043$	$0.0482 \pm 0.0043$	$0.0907 \pm 0.0043$
– function g					
	$Kp2$	$Kp0$	$KQ85$	$KQ75$	<i>Dobler Mask</i>
$F_0$	$0.1337 \pm 0.0002$	$0.1218 \pm 0.0002$	$0.1237 \pm 0.0002$	$0.1191 \pm 0.0002$	$0.1416 \pm 0.0002$
$C_0$	$0.0006 \pm 0.0001$	$0.0006 \pm 0.0001$	$0.0004 \pm 0.0001$	$-0.0020 \pm 0.0001$	$0.0002 \pm 0.0001$
$D_0$	$0.1300 \pm 0.0043$	$0.1443 \pm 0.0043$	$0.1193 \pm 0.0043$	$-0.0259 \pm 0.0043$	$0.1548 \pm 0.0043$
$\chi^2$ analysis – function p					
	$Kp2$	$Kp0$	$KQ85$	$KQ75$	<i>Dobler Mask</i>
$F_0$	$0.1321 \pm 0.0002$	$0.1171 \pm 0.0002$	$0.1219 \pm 0.0002$	$0.1021 \pm 0.0002$	$0.1404 \pm 0.0002$
$C_0$	$0.0026 \pm 0.0001$	$-0.0131 \pm 0.0001$	$0.0049 \pm 0.0001$	$0.0008 \pm 0.0001$	$0.0097 \pm 0.0001$
$D_0$	$0.1451 \pm 0.0043$	$0.1812 \pm 0.0043$	$0.1495 \pm 0.0043$	$0.1665 \pm 0.0043$	$0.1163 \pm 0.0043$
$\chi^2$ analysis – function g					
	$Kp2$	$Kp0$	$KQ85$	$KQ75$	<i>Dobler Mask</i>
$F_0$	$0.1321 \pm 0.0002$	$0.1171 \pm 0.0002$	$0.1219 \pm 0.0002$	$0.1021 \pm 0.0002$	$0.1404 \pm 0.0002$
$C_0$	$-0.0008 \pm 0.0001$	$-0.0113 \pm 0.0001$	$-0.0006 \pm 0.0001$	$0.0291 \pm 0.0001$	$0.0103 \pm 0.0001$
$D_0$	$0.1452 \pm 0.0043$	$0.1813 \pm 0.0043$	$0.1494 \pm 0.0043$	$0.1665 \pm 0.0043$	$0.1164 \pm 0.0043$

Table 5.7: Values of the fit parameters obtained fitting the coefficients with the model proposed by Dobler et al. (2009). The fitted coefficients are computed using  $\chi^2$  and adopting different masks of the sky. We considered all the masks provided by the *WMAP* science team and the mask produced by Dobler et al. (2009). The data used to derive the fitted coupling coefficients are previously cleaned of the CMB emission which is traced by the ICA map. This one has been computed using full-sky data maps. The values are expressed in terms of intensity, to more easily compare the results with those derived by Dobler et al. (2009).

implemented with the function  $p$  and for the mask  $KQ75$ , with  $g$ . Again, this is the effect of the irregularities of the spectra mentioned before. On the contrary, the  $\chi^2$  analysis returns always regular spectra characterised by a positive value of the  $D_0$  parameter. Anyway, both the methods give a value of the amplitude of the WIM component not consistent with zero.

#### 5.4.1 Dependence on the CMB estimation

Simulations showed that a critical issue is the bias of the CMB estimation due to foreground residuals. To check the impact of it, in the real data analysis, we experimented with different estimations of the CMB additionally to the one derived with  $\chi^2$ . Namely, we used the same ILC maps as Dobler et al. (2009) and already described in Section 5.2. Furthermore, to better compare our results with theirs, we tried to reproduce the conditions of their analysis, by substituting the  $H\alpha$  map with their free-free template and by adopting their mask of the Galactic plane. The only difference in the analysis is the fact that we did not fit the data with the Haze template.

We performed the analysis using  $\chi^2$  as well as the simple  $\chi^2$ , also because it is the method

closer to the one used by Dobler et al. (2009). The results are compared in Figure 5.10.

Again, the  $\chi^2$  analysis is extremely stable regardless of the CMB tracer used: the shape of the bump is quite evident and the results basically match those derived by Dobler et al. (2009). It is true also when the CMB emission is described by the map derived from the analysis. On the contrary, when the algorithm is used to compute the coefficients, the presence of the bump is less clear as a consequence of irregular features in the spectrum. It depends strongly on the non linear function used:  $g$  shows now the largest instability especially with specific ILC maps adopted to subtract the CMB contribution.

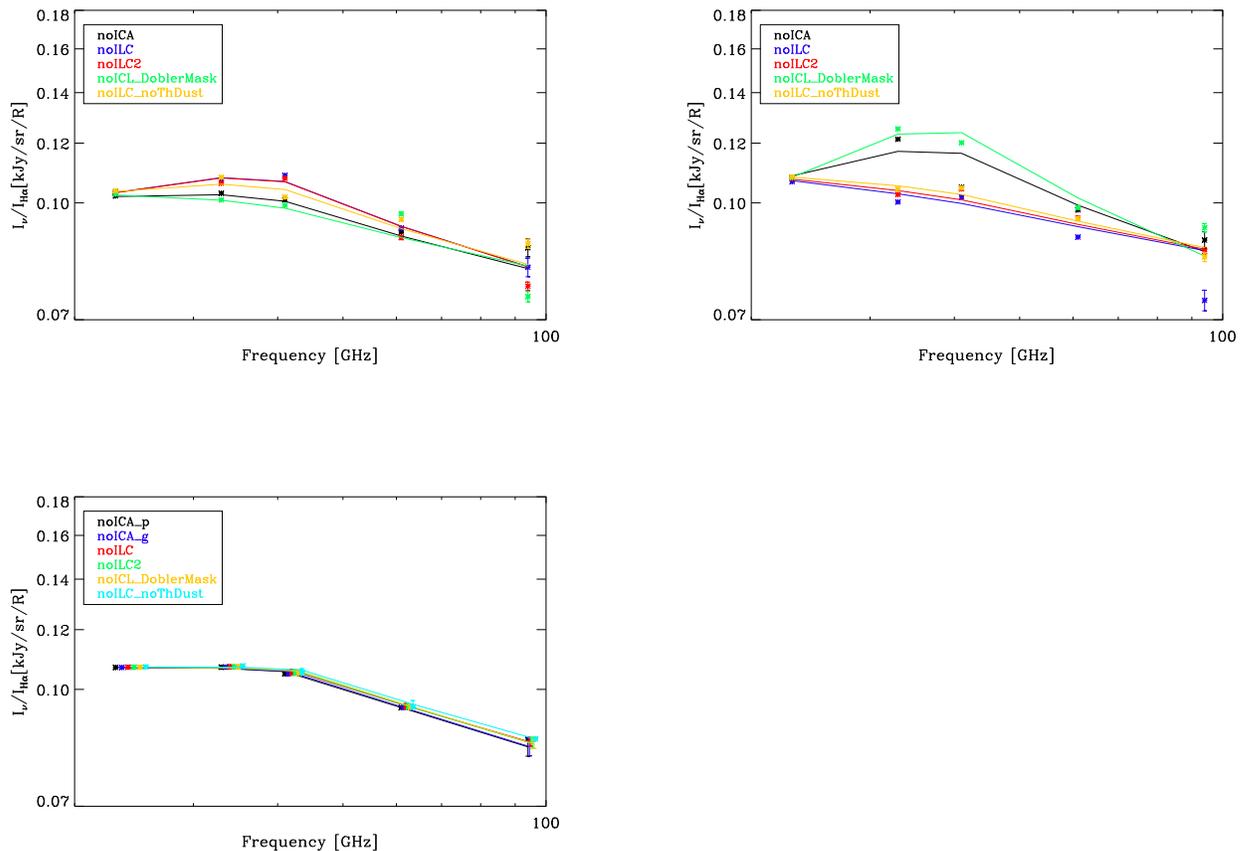


Figure 5.10: We show the free-free coupling coefficients between the *WMAP* data and the  $H\alpha$  template, as computed by using the  $p$  function (top), the  $g$  function (center) and the  $\chi^2$  analysis (bottom). The CMB emission is subtracted from the data and estimated using both the algorithm and the ILC method: in the latter case, we consider the ILC maps used by Dobler et al. (2009). Their mask is also used in the analysis, as well as for computing the ICA CMB map and one of the ILC map ( $ILC_{DoblerMask}$ ). We used the asterisks for showing the values of the coupling coefficients, and the lines for the fit.

The fit of the coefficients, as shown in Table 5.8, confirms these statements. The values of the parameter  $D_0$  derived by the  $\chi^2$  are extremely stable. This is not the case for , although the parameter is never consistent with zero, meaning that the spectrum is always consistent with the

presence of a bump.

	$F_0$	$D_0$	$C_0$
	p		
<i>ICA</i>	$0.10076 \pm 0.00058$	$0.08364 \pm 0.01292$	$0.00734 \pm 0.00031$
<i>ILC</i>	$0.10121 \pm 0.00059$	$0.14915 \pm 0.01303$	$0.00107 \pm 0.00031$
<i>ILC<sub>Kp2</sub></i>	$0.10124 \pm 0.00059$	$0.14974 \pm 0.01301$	$0.00070 \pm 0.00031$
<i>ILC<sub>doblermask</sub></i>	$0.10205 \pm 0.00059$	$0.06166 \pm 0.01301$	$0.00288 \pm 0.00031$
<i>ILC<sub>noThDust</sub></i>	$0.10218 \pm 0.00059$	$0.11458 \pm 0.01317$	$0.00226 \pm 0.00031$
	g		
<i>ICA</i>	$0.10624 \pm 0.00058$	$0.19697 \pm 0.01292$	$0.00665 \pm 0.00031$
<i>ILC</i>	$0.10688 \pm 0.00059$	$0.02733 \pm 0.01303$	$0.00133 \pm 0.00031$
<i>ILC<sub>Kp2</sub></i>	$0.10712 \pm 0.00059$	$0.02923 \pm 0.01301$	$0.00118 \pm 0.00031$
<i>ILC<sub>doblermask</sub></i>	$0.10421 \pm 0.00059$	$0.30248 \pm 0.01301$	$0.00215 \pm 0.00031$
<i>ILC<sub>noThDust</sub></i>	$0.10770 \pm 0.00059$	$0.04313 \pm 0.01317$	$0.00298 \pm 0.00031$
	$\chi^2$ analysis		
<i>ICA<sub>p</sub></i>	$0.10590 \pm 0.00058$	$0.08700 \pm 0.01292$	$0.00707 \pm 0.00031$
<i>ICA<sub>g</sub></i>	$0.10590 \pm 0.00058$	$0.08699 \pm 0.01292$	$0.00765 \pm 0.00031$
<i>ILC</i>	$0.10591 \pm 0.00059$	$0.08660 \pm 0.01303$	$0.00137 \pm 0.00031$
<i>ILC<sub>Kp2</sub></i>	$0.10591 \pm 0.00059$	$0.08661 \pm 0.01301$	$0.00152 \pm 0.00031$
<i>ILC<sub>doblermask</sub></i>	$0.10591 \pm 0.00059$	$0.08662 \pm 0.01301$	$0.00287 \pm 0.00031$
<i>ILC<sub>noThDust</sub></i>	$0.10591 \pm 0.00059$	$0.08648 \pm 0.01317$	$0.00249 \pm 0.00031$

Table 5.8: Values of the fit parameters obtained fitting the coefficients with the model proposed by Dobler et al. (2009). The fitted coefficients are computed using as well as the  $\chi^2$  analysis, from data previously cleaned of the CMB emission. This is traced by different maps: we considered the ICA map together with the ILC maps adopted by Dobler et al. (2009). Finally, we adopted their mask. The values are expressed in terms of intensity, to more easily compare the results with those derived by Dobler et al. (2009).

These differences in the frequency spectra can be better understood looking at the features of the CMB maps used in the analysis. Specifically, we plotted the difference between the ILC maps adopted by Dobler et al. (2009) and the ICA map derived with the  $p$  function for the regions of the sky outside their mask. The maps show that the estimation of the CMB is generally more contaminated along the Galactic plane than the ILC ones. Conversely, the latter has different residuals depending on the way it is computed, mostly concentrated at high latitudes and around the Gum Nebula. These residuals introduce a contamination in the data cleaned of the CMB, enhancing the mix and cross-talk among the components. This is reflected in a poor components separation, if the ICA approach is used. In fact, the cubic and exponential form of the non-linear functions used as neg-entropy approximations, amplify the effect of the residuals. The consequent diverse sensitivity to the statistical properties of the sky, explains the discrepancy in the response of the methods implemented. Indeed, it is interesting to observe that the residuals do not impact the  $\chi^2$  analysis.

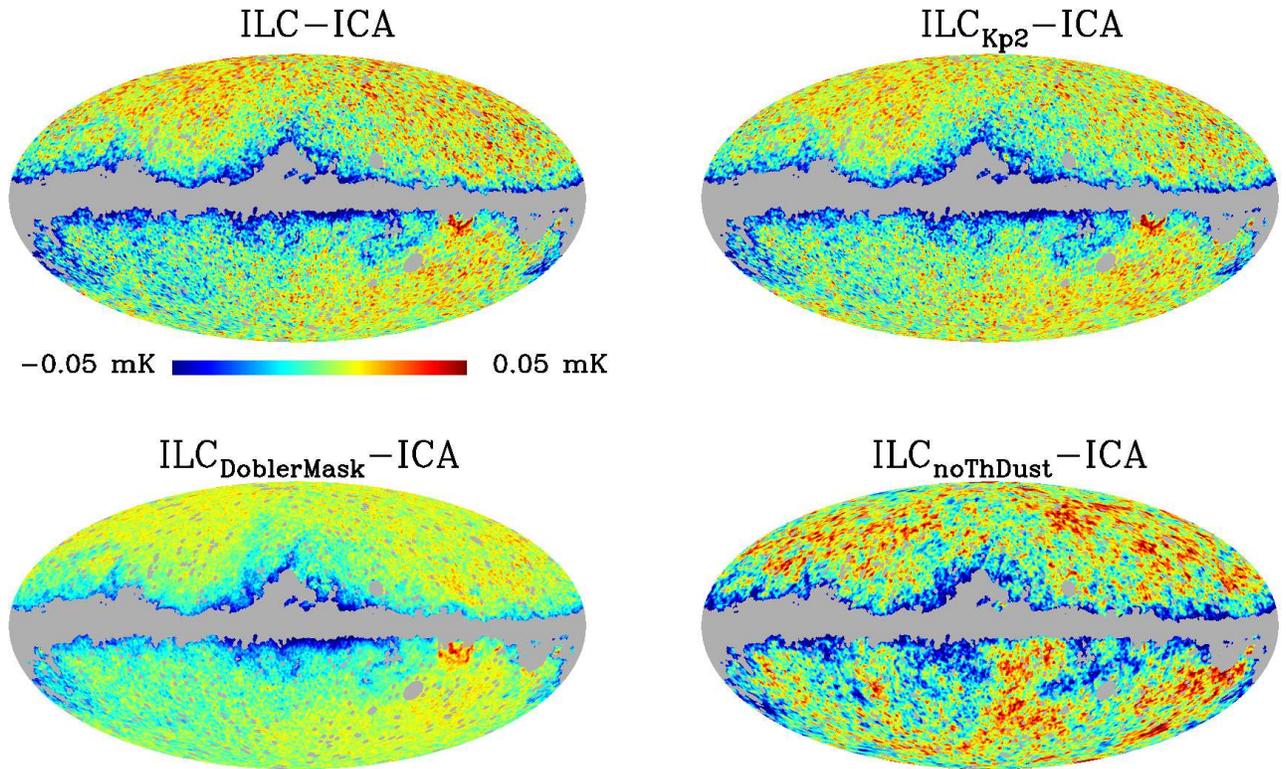


Figure 5.11: We show the difference between the ILC maps adopted by Dobler et al. (2009) and the ICA map derived with the  $p$  function for the regions of the sky outside their mask. The same mask is also applied to all the other maps.

## 5.5 Discussion

The work presented in this chapter aimed at investigating the presence of a bump in the free-free spectrum, by experimenting with the analysis of CMB observations, where the cosmological emission is removed. The impact of this subtraction has been studied as one of the possible reasons for the existence of the bump as claimed by Dobler & Finkbeiner (2008b) and Dobler et al. (2009).

After testing the analysis using simulated data, we carried out a study of the *WMAP* 5-year observations cleaned of the CMB emission.

The first step of the work has shown that, given a specific foreground model, both and the  $\chi^2$  analysis return spectra of the single components which are perfectly in agreement with the expected one. Specifically, a bump is detected in the free-free spectrum, only when the contribution from the spinning dust emission in the WIM is included in the foreground model.

Therefore, we have been able to exclude the CMB subtraction as a possible reason for the manifestation of the bump in the free-free spectrum. The same response concerns the cross-talk among the three considered foreground components. This affects only the synchrotron and dust emission, whose coefficients distributions are slightly biased with respect to the expected values. Besides, the free-free coefficients are generally well recovered.

Simulations have also shown that a rising free-free spectrum, as derived with a standard analysis of the *WMAP* data, seems not to be the favored solution, if the CMB is removed from the data. Instead, when the CMB is not pre-subtracted, the cross-talk between the cosmological signal and the foregrounds is the main reason why the spectral behaviour deviates from the theoretical one, being remarkably steep or rising with frequency. Consequently, the uncertainty of the coefficients is very large compared to the case where the CMB is excluded from the data, making the derived free-free spectrum consistent with different scenarios. The existence of an additional component correlated with the  $H\alpha$  emission, as claimed by Dobler et al. (2009), is therefore plausible: the rising spectrum could just be a different signature of it.

However, both the assumed foreground models in the simulations are not completely in agreement with the real observations. The comparison between the distribution of the linear weights, which define the ILC and ICA maps used as a tracer of the CMB, and the analogous weights derived from the real data analysis, highlighted a large inconsistency between the latter and the mean of the distributions. Such a limitation of the model does not invalidate the conclusions derived from simulations, but prevents us from understanding the results obtained from the same analysis of the *WMAP* data.

The application on the *WMAP* data cleaned of the CMB seems to favour the existence of a bump in the free-free spectrum. This is strongly confirmed by the  $\chi^2$  analysis, whose response is remarkably stable in this direction, independently of the tracer of the CMB. However, becomes more unstable once the CMB is excluded from the data, so that it is difficult to separate the single emissions. It is most likely the consequence of the larger complexity of the foregrounds in the real sky, together with the cross-talk among the components, which in the real case possibly involves also the free-free emission. Consequently, the ICA and ILC maps used to remove the cosmological signal from the data are probably more affected by residuals. Therefore, we interpreted these residuals and their different features depending on the method used to estimate the CMB, as the main responsible for the instability of . The neg-entropy statistic employed in the component separation makes the method particularly sensitive to residuals and to the consequent complex features of the data.

The degeneracy between the components - especially between the synchrotron and dust emissions - can only be broken having observations with sufficient frequency coverage, such as those provided by the Planck satellite. Moreover, full-sky and high resolution maps between 5 and 15 GHz could help to eliminate this ambiguity and, at the same time, could certainly favour the detection of the presence of the WIM emission, better defining its spectrum. Forthcoming experiments such as C-BASS (C-Band All Sky Survey) and QUIJOTE (Q-U-I JOint TENERife), will be very useful in this sense.



## Chapter 6

# Internal analysis of the *WMAP* five-year data on patches of the sky.

### 6.1 Introduction

The analysis of the *WMAP* five-year data presented by Maino et al. (2007), and discussed in Chapters 3 and 4 has confirmed  $\text{WMAP}$  as a useful and stable tool to disentangle the mixture of signals in the sky as measured by *WMAP*. Moreover,  $\text{WMAP}$  turned out to be a powerful method to recover a good estimation of the CMB emission, even in the extreme working conditions of a full-sky analysis. To achieve this result, it is convenient to combine the foreground analysis with the internal one.

Nonetheless, significant improvements can still be introduced in the analysis: they concern both the treatment of the data and the implementation of the component separation problem itself.

One of the possible modifications is connected to the most significant limitation of the approach, as it has been generally used. In the standard implementation of the code, both for the internal and the foreground analysis of a set of CMB data, the values of the scaling factors (see Equation 3.2) are unique for the entire sky, regardless of the existing variations of the spectral and spatial properties of the different emissions. Therefore a natural improvement to the component separation could be the possibility of taking into account such variations.

The main point of the work presented in this chapter is to experiment with the derivation of a CMB estimation as a result of the implementation of the internal analysis of the *WMAP* 5-years data on small patches of the sky.

Several subdivisions of the sky have been adopted, in order to answer to the following questions:

1. Is the partition of the sky actually useful to improve the performance of our analysis?
2. Which is the optimal number of regions to consider?
3. How should these regions be selected in order to get the best reconstruction of the CMB signal?

Given our previous considerations, we expect a partition of the sky to be helpful, particularly if it is a subdivision which reflects the actual properties of the sky. However, it might be that a better

performance is simply driven by the number of subdivisions of the sky, independently of the way they are defined.

This is the reason why we have examined both a simple schematic subdivision of the sky, as well as a less straightforward one. Specifically, we looked for an internal prescription for selecting regions guided by the results themselves (see Section 6.1.1).

Besides the regional approach to the problem, two simple improvements can be introduced to the ICA implementation, as explained in Section 2.4.4. The first one consists of introducing a constraint (or *prior*) to require the CMB spectrum to follow the Planck law (e.g. to have a flat spectrum in terms of thermodynamic temperature). Thus the component is estimated accordingly to the prior itself. However, it has been demonstrated by Vio & Andreani (2008) that this modification makes become identical to the ILC method. Therefore, switching on and off this constraint, is equivalent to use either the or the ILC code. For the work presented here, we adopted both of them, aiming to an orderly and explicit comparison of their performances.

The second modification to the original code consists of compressing the output of the algorithm, in order to reduce any redundancy, but preserving all the information enclosed in the data. As explained in Section 2.4.4, the standard data compression and Principle Component Analysis techniques are combined in order to find a linear combination of the input maps which preserves all the information concerning the sky signal, but which consists of independent data.

Both these new features of the algorithm (hereafter ) have been used simultaneously only in the process of defining a physically interesting partition of the sky (see next section). In the subsequent steps of the analysis, they have been dropped and only has been used and compared to the ILC analysis.

### 6.1.1 Partitioning of the sky

Defining the criterion to use in partitioning the sky is a critical issue. As stated before, we experimented with simple partitions, as well as more sophisticated ones, taking advantage of the known physical properties of the sky and their spatial variations. An optimal partition consists of a subdivision of the sky into regions across which the foreground properties can be considered constant. The issue, however, is that the foreground properties are also part of what is being sought in the analysis.

We first adopted a simple selection of regions (see Figure 6.1) where the Galactic and Ecliptic Plane have been used to separate two regions: they are defined as the ensemble of pixels with  $|b| < 15^\circ$  or  $|b| > 15^\circ$ . Thus, three regions maps have been created, in a straightforward way, distinguishing between the northern and southern Galactic or Ecliptic Hemispheres. Although simple, these partitions of the sky are interesting in order to test the existence of Hemispherical asymmetries as well as anomalies connected to the Ecliptic plane. Moreover the lack of complexity makes them very useful tools to test the applicability of the algorithm on patches of the sky, at least in a simple context.

Nevertheless, we are ultimately interested in defining a subset of regions which represents the spatial variation of the foregrounds emission on the entire sky as well as possible. By this point of view, precious hints can be derived looking at the features of the non-cosmological maps returned by .

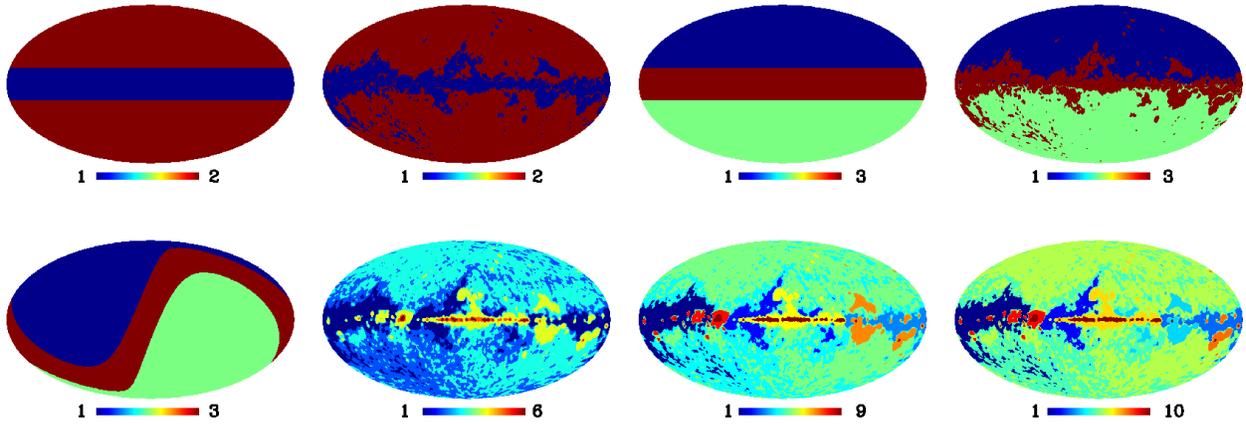


Figure 6.1: Full-sky maps color-coded to show the different sets of regions that were used to generate the and ILC CMB maps, both for the simulations and for the real data analysis.

Indeed, when the code is internally applied to a set of  $N$  measurements at different frequencies, it returns a map which is the best CMB reconstruction, but also other  $N - 1$  maps which are the best estimation of the independent foregrounds signals or the noise. Taking advantage of the compressing (see Section 2.4.4), it is possible to concentrate all the main components but the CMB, in only one map. Specifically, it is convenient to compress the input data to three output maps (e.g. CMB, foregrounds and noise). In order to limit the cross-correlation among the components and, therefore, to optimize the performance of the component separation, it is also useful to impose the prior for the CMB.

The map relative to the foreground components will then include all the contributions from any emission. As visible, for example, in Figure 6.2, its main features tell us the dominant components and regions of the sky with the same properties. Therefore, this map is suitable for tracing a subdivision of the sky, such that these regions can be treated separately.

Both the non-linear functions ( $p$  and  $g$ ) have been used in this analysis. The results derived using  $g$  are particularly interesting in terms of the reconstruction of the foregrounds. The smoothed map at a resolution of 3 degrees is shown in Figure 6.2.

Note that we have also experimented with the computation of the non-linear statistics maps generated from this foreground component map: in practice, we applied the  $p$  and  $g$  functions to the foregrounds map shown in Figure 6.2. Nonetheless, we retained the latter as the most significant result, since the other ones did not show revealing improvements.

The foregrounds map has been used to define a new set of regions: we separated groups of pixels selected in a way that each region has the same total integrated temperature. The integrated temperature criterion is useful since it naturally focuses on regions of the sky where the foreground variations are largest, namely in the Galactic plane. Note however, that the term “region” here is used improperly since we are actually defining a set of thresholds. Regions are instead sets of pixels with well defined coherent and integrated structure. Looking at the pixels comprising a given threshold, there are possibly many separate areas on the sky, i.e. regions. Using temperature thresholds is motivated by Miville-Deschênes et al. (2007): indeed, they have shown that, at least for dust emission, there is a correlation between

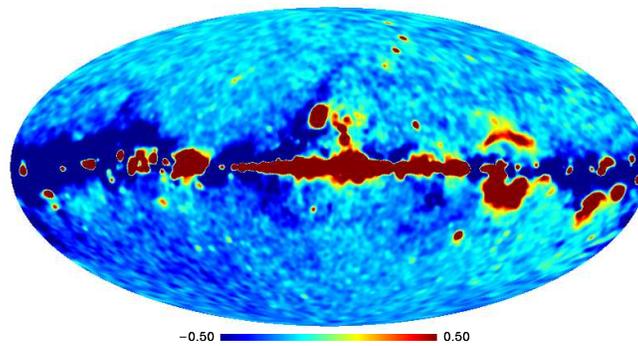


Figure 6.2: Plot of the map of the integrated foreground components recovered by  $\dots$ , imposing a compression to three output maps. The color coding represents the relative temperature (i.e. red is hotter and blue is colder).

regions defined this way and their spatial power spectrum, and that this connects physical properties to temperature. Consequently, it is reasonable to assume that these physical properties are also related to frequency spectral behaviour.

Following the threshold prescription, we ended up with a six regions map, shown in Figure 6.1 (second map in the second row, from the left). It is interesting to note that the high latitude sky is mostly one region, with a second confined largely to one quadrant.

Merging them in one region, this can be distinguished from the regions along the Galactic plane which again can be treated as a unique patch. In this way we have defined a new map of two regions: it is the second top map, from the left, in Figure 6.1. Again, as for the simple cut along the Galactic plane, it is possible to consider the case where the northern Hemisphere is distinguished from the southern one: this gives a three regions map (the fourth top map).

In the six regions map, it is not difficult to recognize some structures along the Galactic plane which can be associated with well known sources: the Gum Nebula, the Cygnus region, the Orion nebula as well as the North Polar Spur are easily visible in the map. Because of that, we decided to further subdivide the map by hand, separating the main structures in the Plane, in what we can now properly call regions. Moreover, following the same criterion of coherency and integration of the structures, we merged the fifth and sixth regions: this also allows us to avoid too small areas. We propose two solutions, with 9 and 10 regions, presented in Figure 6.1. Note that these last two maps are influenced by our intuition and thus, might not be optimal choices.

Finally, beside our own subdivisions of the sky, for the proposed analysis of the real *WMAP* data, we have also considered the 12 regions defined by the *WMAP* science team to perform the ILC analysis (Hinshaw et al., 2007), and the 400 regions used by Park et al. (2007) (see Figure 6.3). They are again an attempt of a physically driven selection of different areas of the sky, and therefore, they give us the opportunity of comparing and extending our results to a broader scenario: both of these partitions, in fact, contain a larger number of regions than ours, derived following different criteria. The first one is basically a subdivision of the regions close to the Galactic plane, being the largest area of the sky concentrated in a single region (e.g. Region 0). The partition proposed by Park et al. (2007), instead,

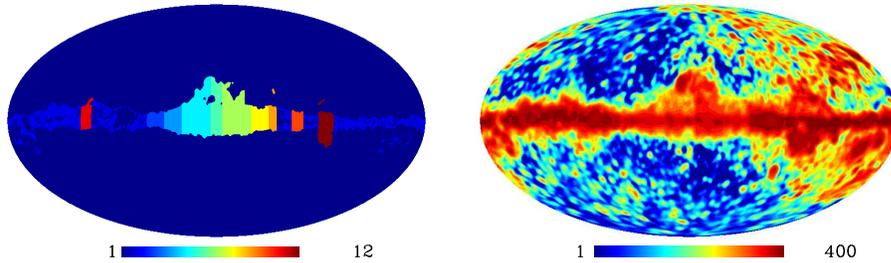


Figure 6.3: Full-sky maps color-coded to show the partitions proposed by Hinshaw et al. (2007) (left) and Park et al. (2007) (right), which we have adopted for the analysis of the *WMAP* observations.

is based on the distributions of the spectral indices on the sky measured at low- and high-frequency intervals: it is particularly interesting since the regions selection is explicitly done by taking into account the spectral index variations which are studied and derived by using the MEM maps (Hinshaw et al., 2007).

## 6.2 Method of analysis

The code has been used to analyse sets of real and simulated CMB data at different frequencies, in different patches of the sky. The same scheme of analysis has been repeated with the ILC code to compare the performance of the codes.

Having fixed the regions of the sky to be considered, the component separation is achieved as described in Chapter 2 (see Sections 2.4.3 and 2.3, respectively). The CMB solutions for the single regions are then merged in a full-sky map, by co-adding them: in order to suppress boundary effects, the edges of the regions are smoothed with an effective Gaussian beam of amplitude  $1.5^\circ$ .

## 6.3 Monte Carlo simulations

To test the accuracy of the methods with a regional approach and to compare the results derived with them, we first applied the regional analysis to simulations of the *WMAP* observations. Specifically, we generated 200 maps of the microwave sky at the 5 *WMAP* frequencies, each containing a realisation of the CMB signal, the Galactic foreground emission, and instrumental noise appropriate to the specific channel. As usual, the study has been undertaken at an effective resolution of  $1^\circ$ .

Each CMB component corresponds to a Gaussian realisation of the theoretical best fit CMB angular power spectrum as determined by the *WMAP* five-year analysis. The white noise contribution is simulated for each of the 10 differencing assemblies (DAs) in the usual manner, as described in the Appendix A. The CMB and noise realisations are then combined, and the maps smoothed. For those frequencies with multiple associated DAs, the band averaged sky maps are formed using simple averaging. For the foreground emission, we added to each simulated frequency map four simulations of the main Galactic components (synchrotron, thermal and anomalous dust, and free-free emissions)

according to the Planck Sky Model (PSM, see Miville-Deschênes et al. (2008) and Chapter 1 for more details). This is in fact the most complete model currently available, although a direct comparison with the *WMAP* data (see Section 1.5.4) shows that, at these frequencies, the PSM maps do not match perfectly with the observations. Besides, other models available in the literature seem to better trace the foreground emissions as measured by *WMAP*. Therefore, we also used the MEM solutions produced by the *WMAP* science team, which by construction are closer to the data. Furthermore, we experimented with the foreground maps recently published by Ghosh et al. (2010) (hereafter GFM). In both the cases, the model is derived from the *WMAP* data pre-cleaned of the CMB emission, whose estimation is computed using the ILC implemented in the real and needlet space respectively.

The five frequencies maps for each simulation have been used as an input to `ilcmap` and the ILC code, to be analysed simultaneously, on different patches of the sky. The CMB solution derived for each region has been used to produce a full-sky map, following the prescription described in Section 6.2.

Different figures of merit have been used to evaluate the results. We first studied the weights of the linear combination of the five input maps: this is an important consistency test of the simulations with data which has not been undertaken anywhere else. Afterwards, we focused on the statistical properties of the returned CMB maps.

Figures 6.4, 6.5 and 6.6, show the distribution of the weights obtained for each frequency band and for each region, in the case where we used the three regions, defined with respect to the Galactic plane, and where we adopted the PSM to trace the foregrounds. These are just an example of what has been computed for all the partitions of the sky: the derived results are generally valid for all of the partitions shown in Figure 6.1. Analogous plots are shown in Figures 6.7, 6.8, 6.9, 6.10, 6.11 and 6.12 for the MEM and GFM simulations. To compare the results, we retained the same partition of the sky.

- For each single method, the distributions are symmetric and well described by a Gaussian function. However, for the PSM simulations, the mean and the mode are generally not consistent with the expected values derived from the analysis of the real data: they lie well outside of the distributions. This suggests that the simulated data are actually not representative of the real observations, although the model assumed has been proposed as one of the most realistic. The situation is visibly different if either the MEM or the GFM maps are used in the simulations (see Figures 6.7, 6.8, 6.9, 6.10, 6.11, and 6.12): in these cases, indeed, the weights derived using real data, are generally within the distribution, although often on the tail. This suggests a better match between the foreground model adopted and the observations than for the PSM.
- The `ilcmap` weights are generally larger in absolute value, than the ILC ones. The fact that the two methods use different statistics for the estimation of the CMB plays certainly an important role. Besides, the fact that ILC imposes a constraint on the CMB frequency spectrum, probably helps to stabilise the solution.
- The regions closer to the Galactic plane are the most contaminated by foregrounds. Therefore, the component separation becomes more challenging, due to the strong mix of the physical emissions.

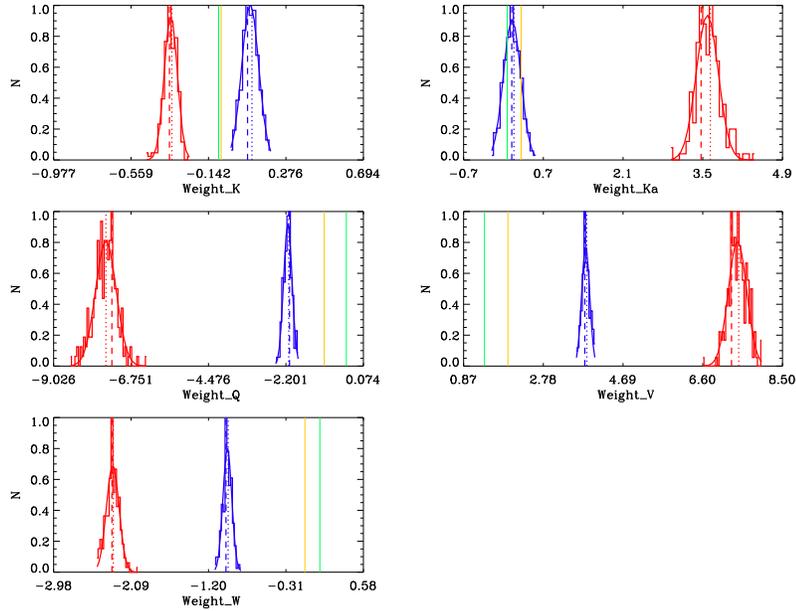


Figure 6.4: **PSM simulations.** Statistical distribution of the CMB weights derived using  $\text{ICA}$  (red) and the ILC code (blue), adopting the three regions defined with respect to the Ecliptic plane: the values used are relative to the northern Hemisphere. The yellow (ICA) and green (ILC) lines show the values of the weights recovered by the real data analysis. The dot lines indicate the mean of the distributions, while the dashed one show their maxima.

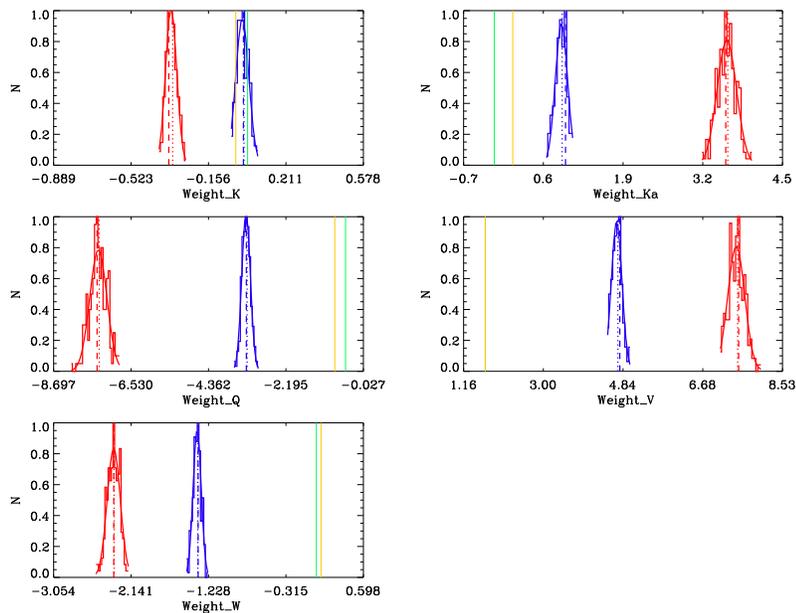


Figure 6.5: **PSM simulations.** Statistical distribution of the CMB weights derived using  $\text{ICA}$  (red) and the ILC code (blue), adopting the three regions defined with respect to the Ecliptic plane: the values are computed using the pixels within 15 degree from the Ecliptic plane. The yellow (ICA) and green (ILC) lines show the values of the weights recovered by the real data analysis. The dot lines indicate the mean of the distributions, while the dashed one show their maxima.

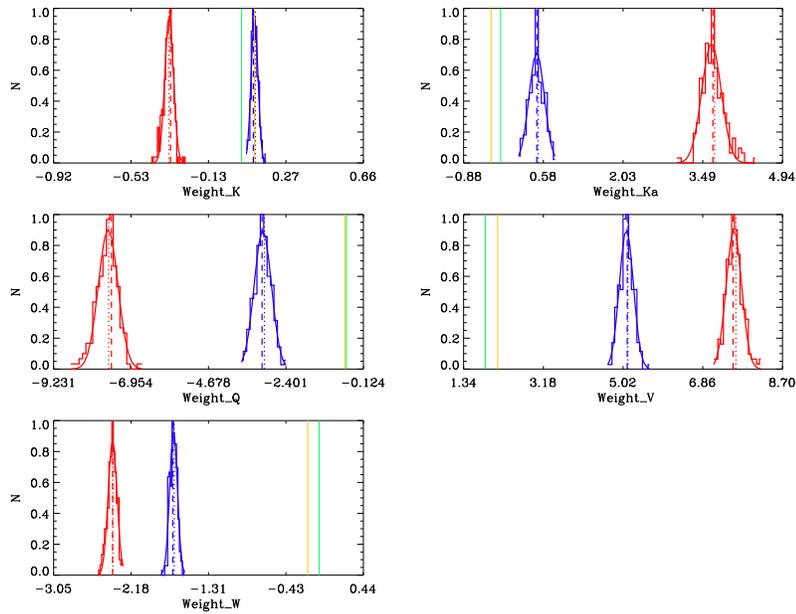


Figure 6.6: **PSM simulations.** Statistical distribution of the CMB weights derived using PSM (red) and the ILC code (blue), adopting the three regions defined with respect to the Ecliptic plane: the values used are relative to the southern Hemisphere. The yellow (ICA) and green (ILC) lines show the values of the weights recovered by the real data analysis. The dot lines indicate the mean of the distributions, while the dashed one show their maxima.

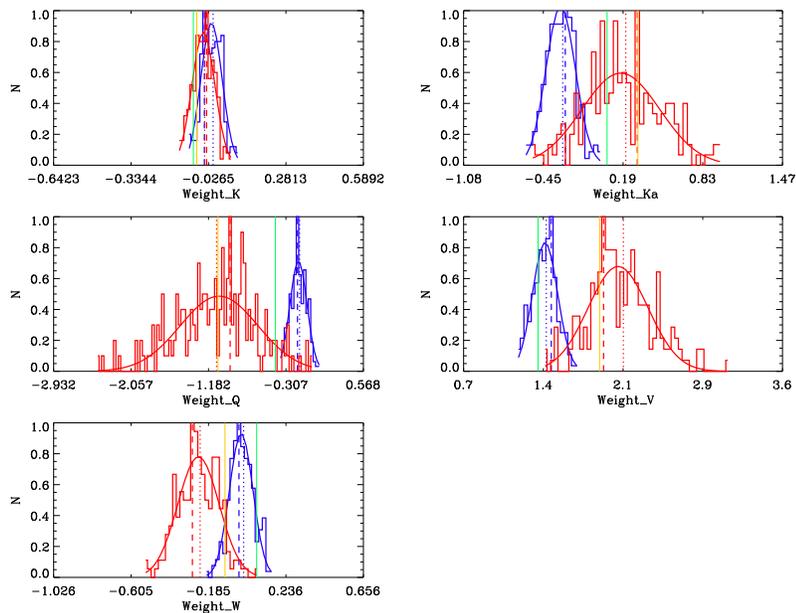


Figure 6.7: **MEM simulations.** Statistical distribution of the CMB weights derived using MEM (red) and the ILC code (blue), adopting the three regions defined with respect to the Ecliptic plane: the values used are relative to the northern Hemisphere. The yellow (ICA) and green (ILC) lines show the values of the weights recovered by the real data analysis. The dot lines indicate the mean of the distributions, while the dashed one show their maxima.

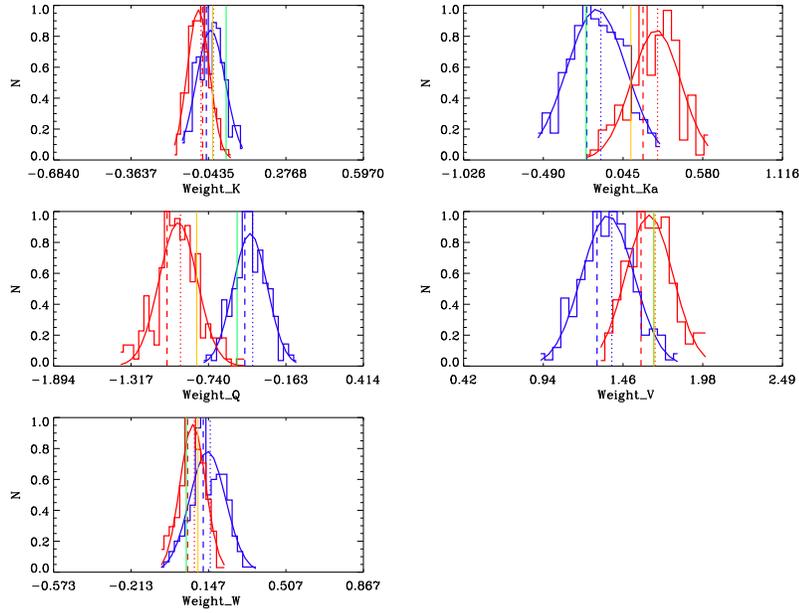


Figure 6.8: **MEM simulations.** Statistical distribution of the CMB weights derived using MEM (red) and the ILC code (blue), adopting the three regions defined with respect to the Ecliptic plane: the values are computed using the pixels within 15 degree from the Ecliptic plane. The yellow (ICA) and green (ILC) lines show the values of the weights recovered by the real data analysis. The dot lines indicate the mean of the distributions, while the dashed one show their maxima.

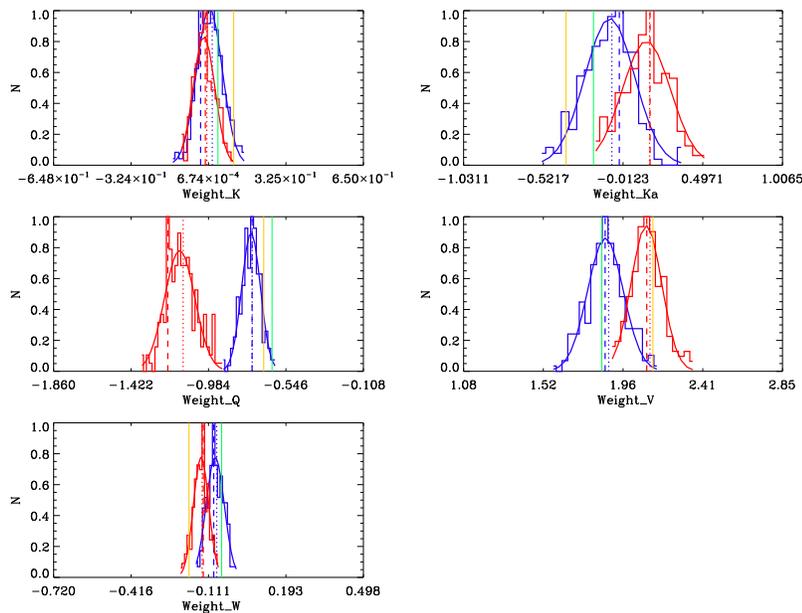


Figure 6.9: **MEM simulations.** Statistical distribution of the CMB weights derived using MEM (red) and the ILC code (blue), adopting the three regions defined with respect to the Ecliptic plane: the values used are relative to the southern Hemisphere. The yellow (ICA) and green (ILC) lines show the values of the weights recovered by the real data analysis. The dot lines indicate the mean of the distributions, while the dashed one show their maxima.

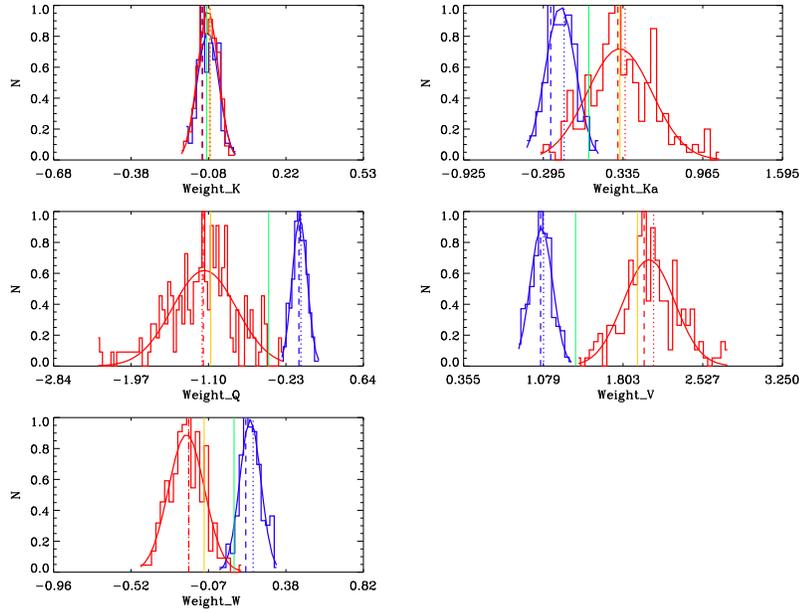


Figure 6.10: **GFM simulations.** Statistical distribution of the CMB weights derived using  $\text{GFM}$  (red) and the ILC code (blue), adopting the three regions defined with respect to the Ecliptic plane: the values used are relative to the northern Hemisphere. The yellow (ICA) and green (ILC) lines show the values of the weights recovered by the real data analysis. The dot lines indicate the mean of the distributions, while the dashed one show their maxima.

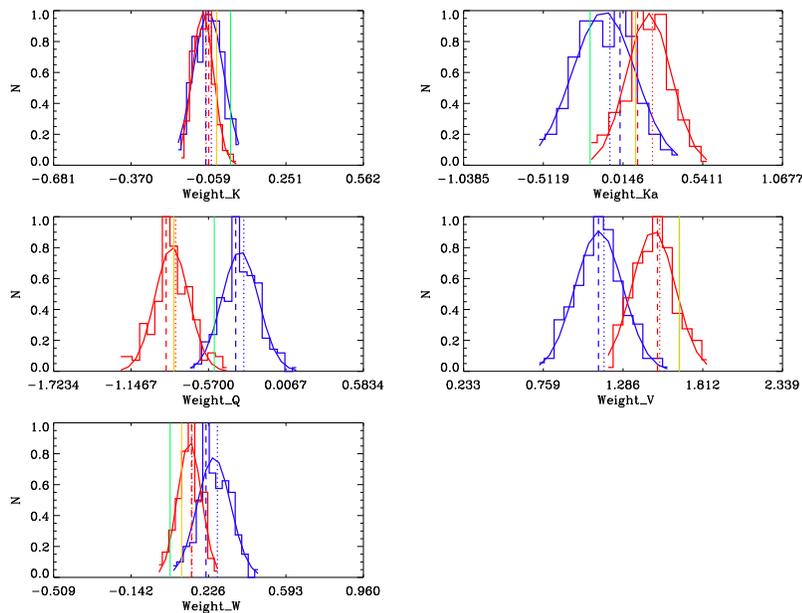


Figure 6.11: **GFM simulations.** Statistical distribution of the CMB weights derived using  $\text{GFM}$  (red) and the ILC code (blue), adopting the three regions defined with respect to the Ecliptic plane: the values are computed using the pixels within 15 degree from the Ecliptic plane. The yellow (ICA) and green (ILC) lines show the values of the weights recovered by the real data analysis. The dot lines indicate the mean of the distributions, while the dashed one show their maxima.

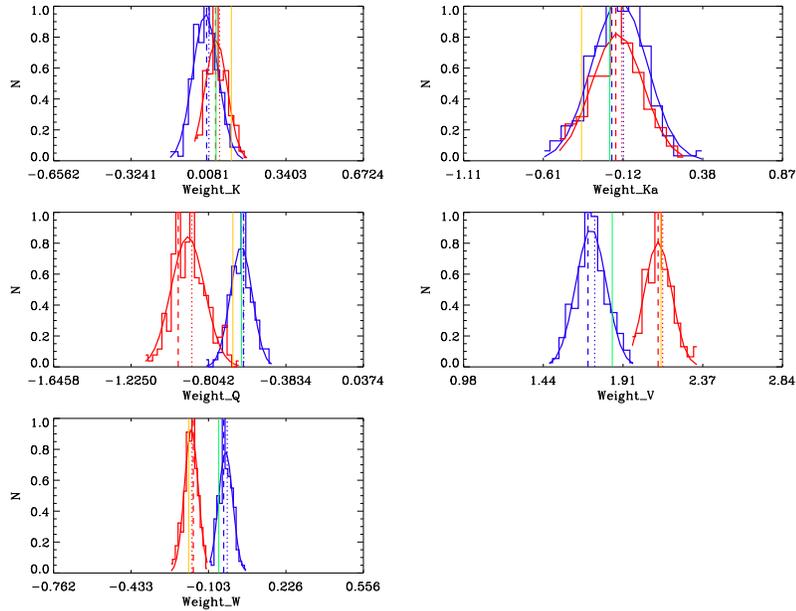


Figure 6.12: **GFM simulations.** Statistical distribution of the CMB weights derived using  $\text{GFM}$  (red) and the ILC code (blue), adopting the three regions defined with respect to the Ecliptic plane: the values used are relative to the southern Hemisphere. The yellow (ICA) and green (ILC) lines show the values of the weights recovered by the real data analysis. The dot lines indicate the mean of the distributions, while the dashed one show their maxima.

This is reflected by a slightly larger distribution of the weights returned by the two methods, with respect to the one derived from the analysis of regions at high latitudes of the sky.

Besides the weights, the output maps of the simulations have been used as a figure of merit to evaluate the results of the analysis. These have been compared to three other maps: the input CMB, the noisy maps, meaning the input maps with the addition of the noise contribution according to the weights of the specific simulation, and the output maps corrected for the bias due to the foregrounds residual (see Section 2.3). The bias has been estimated as the average of the input foreground templates co-added with realisation specific weights.

In particular, comparing the debiased CMB maps with the input ones is interesting in order to evaluate the effect of the applied correction. Note though, that the mean bias is not necessarily a good estimation of the foreground residual in the CMB map for each simulation, although it is the best estimation we can achieve if the foreground model is correct. Of course, this is the case only for the simulations, since we know the foregrounds by construction. It is difficult to assess whether such an estimation of the bias is suitable for the CMB maps derived from data. In fact, the discrepancies highlighted before between data and model, do not support this idea in the case of the PSM: it would be difficult to quantify the impact of this mismatch on a statistical assessment of the data. Again, the issue makes it clear why the *WMAP* science team has preferred the MEM solutions as better tracers of the foregrounds in the data, regardless of their difficult noise properties and the effect of the ILC subtraction in the initial stages of the analysis. Using the GFM maps goes in the same direction of a

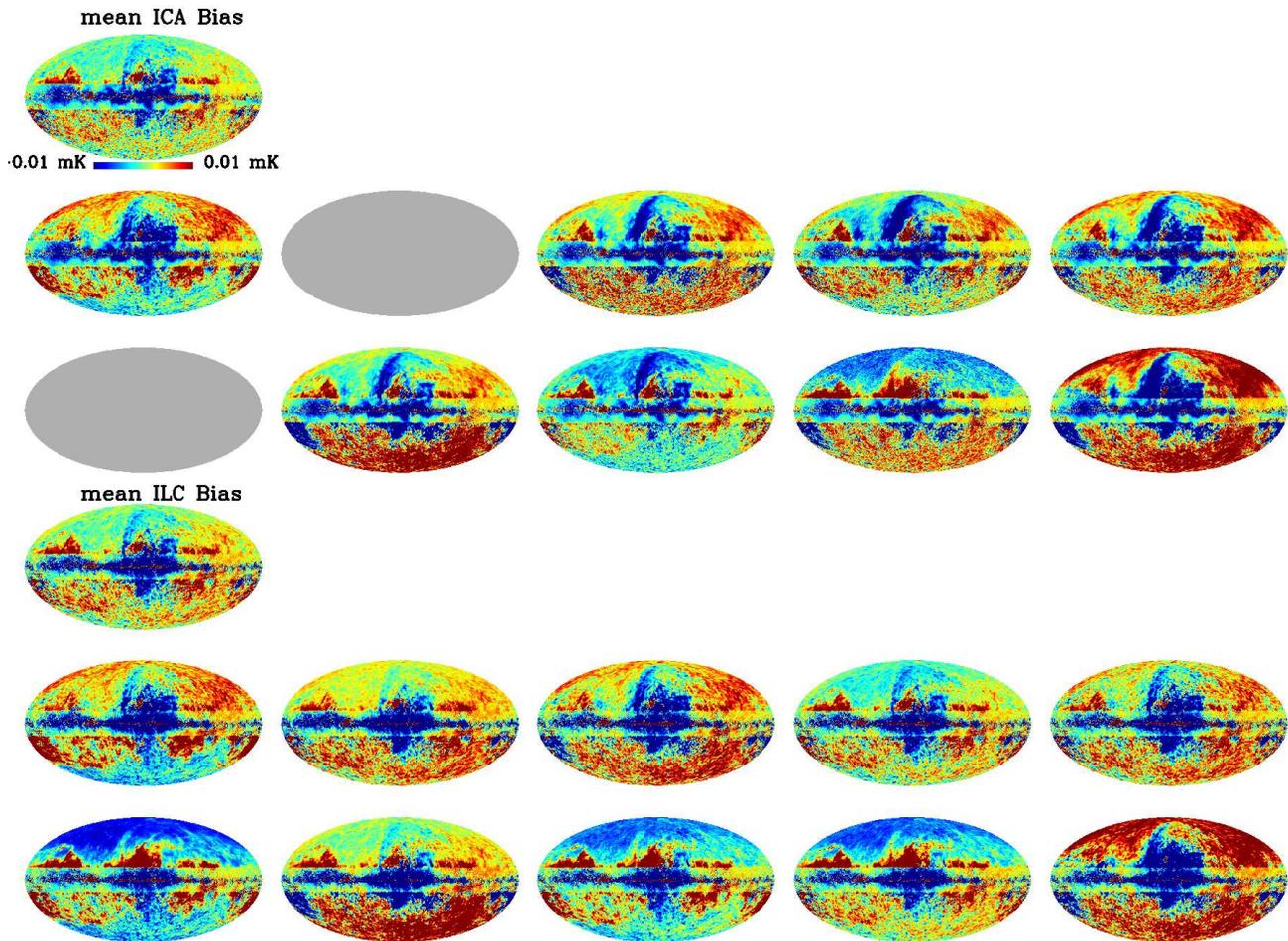


Figure 6.13: **Component separation analysis on three regions defined with a parallel cut along the Galactic plane.** We show the foreground residual map of some simulations randomly selected among 200 available. They are derived from the  $\text{ICA}$  and ILC analysis: the two rows of maps on the top of the pictures are those derived with  $\text{ICA}$ , while the rows on the bottom are those from the ILC analysis. These maps can be compared with the bias estimation, defined as the mean of the foreground residual over all the simulations: it is shown on the top of the single realisation, as indicated by the labels. It is interesting to note the effect of discontinuity due to the sharp cut of the regions. There are some cases where  $\text{ICA}$  does not reach the convergence: these are shown as null maps.

better description of the *WMAP* observations, although also in this case we do not know the impact of the subtraction of the Needlet ILC map (hereafter NILC), done in their analysis.

To check how far the mean estimation of the bias is from the single cases, we had a look at some single realisations randomly chosen among the set of 200 available. For each of them, we compared the bias correction to the mean value used to define the debiased maps (see Figures 6.13 and 6.14, as an example). Such an analysis is shown for the PSM simulations only, since the results derived with the other foreground models are analogous.

Generally, the mean bias seems to be a realistic estimation of the residuals present in the single recovered CMB maps, both for ILC and  $\text{ICA}$ . However, depending on the method used in the

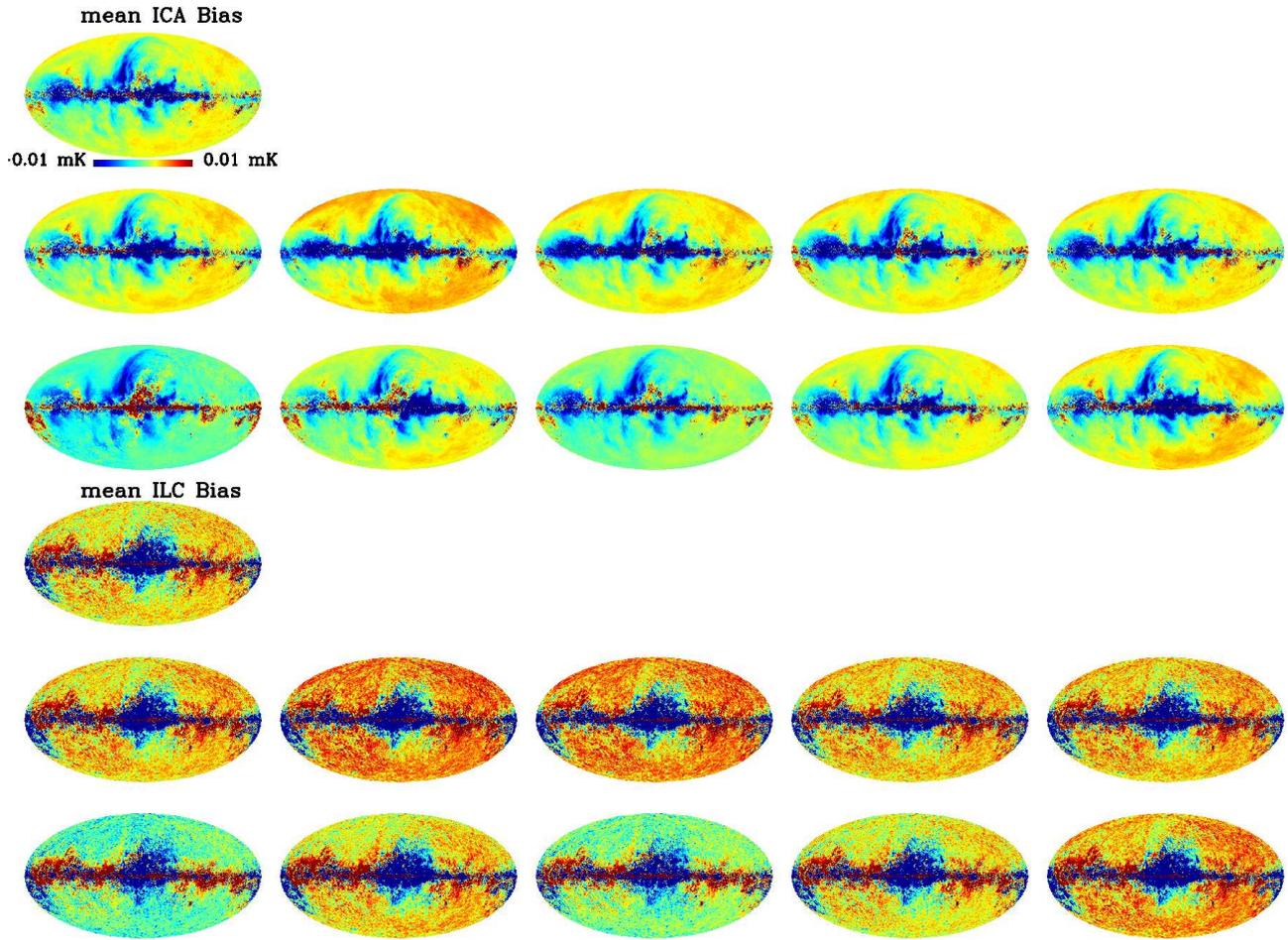


Figure 6.14: **Component separation analysis on three regions of the sky, selected with respect to the Ecliptic plane.** We show the foreground residual map of some simulations randomly selected among 200 available. They are derived from the  $\text{MEM}$  and ILC analysis: the two rows of maps on the top of the pictures are those derived with  $\text{MEM}$ , while the rows on the bottom are those from the ILC analysis. These maps can be compared with the bias estimation, defined as the mean of the foreground residual over all the simulations: it is shown on the top of the single realisation, as indicated by the labels. As for the full-sky analysis, the foreground residuals show different features depending on the method adopted in the analysis. Again, there are some cases where  $\text{MEM}$  does not reach the convergence: these are shown as null maps.

analysis, the dominant structures change. Specifically, with  $\text{MEM}$ , the bias is mostly dominated by negative dust-like structures along the Galactic plane and by the North Polar Spur. On the other hand, adopting the ILC code, the residuals are defined by a spherical negative structure concentrated around the Galactic center and some features along the Galactic plane, which can be easily associated with the free-free emission. These discrepancies are not visible for the MEM and GFM models, as can be seen in the bias maps derived (see Figure 6.15). The bias maps derived with the MEM and the GFM simulations look quite different from the PSM ones (see Figure 6.15). The bias maps are basically dominated by strong emissions along the external regions of the Galactic plane and a negative contribution in the Galactic Center, whose shape depends on the foreground model adopted. It is interesting to note that

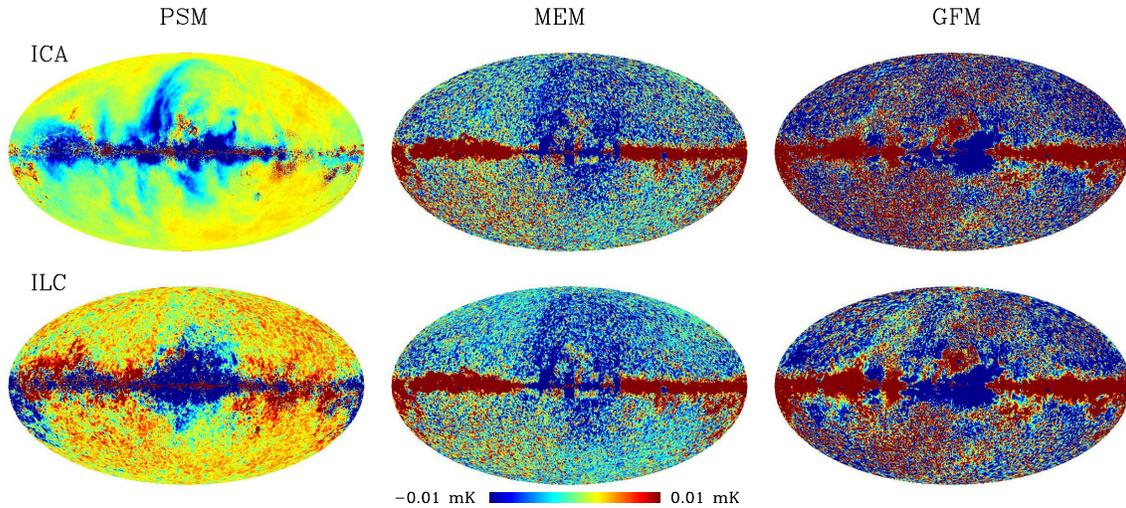


Figure 6.15: Comparison of the bias maps estimated from simulations generated using the three foreground models (PSM, MEM maps and GFM). Both ICA and the ILC code have been considered for the analysis with three regions defined with respect to the Ecliptic plane.

with the MEM maps, the bias seems to reflect the regions used by the ILC that is subtracted before the MEM analysis. Furthermore, as for the PSM, the North Polar Spur is evident and traced as a negative contribution. In the case of the GFM, instead, it is difficult to see evidence of some NILC partitions: the North Polar Spur is not present anymore, while the strongest negative contribution comes from a spherical region around the center, plus the Cyn region.

Finally, it is easy to note that the maps clearly show the effect of the partition of the sky used in the analysis, when the cut parallel to the Galactic plane is adopted (see Figure 6.13). Both the methods of analysis derive solutions which enhance the discontinuity between the area closer to the plane and those at higher latitudes. It is also interesting to observe that the same effect is not visible when other partitions of the sky are considered: since they follow better the natural spatial distributions of the foreground structures. We have tested that this effect does not depend on the foreground model considered to generate the simulated data: using either the MEM foreground templates or the GFM maps, these anomalies remain, although less evident because of a better CMB estimation in the Galactic regions of the sky and therefore, a lower noise contribution.

### 6.3.1 Statistical study of the simulations results

The comparison itself between the CMB maps is carried out in terms of their global statistical properties: for all the maps and for each simulations, we computed the variance, the skewness and the kurtosis over the total number of pixels of the maps. Furthermore, we evaluated the amplitude of the quadrupole moment of the maps. All these quantities have been used as a measurement of non-Gaussianity. Any deviation of the CMB estimated maps with respect to the input one (which is a Gaussian random realisation of the cosmological signal according to the theoretical estimation of the angular power spectrum) can be detected by means of the values of the statistics computed. Indeed,

a temperature map which includes deviations from a Gaussian pattern will have different statistical properties than a Gaussian field. We did not use the power spectrum as a figure of merit of a good estimation of the CMB, since we know that it is remarkably robust against residuals. This point will be explicitly shown later in Section 6.4.1.

Figure 6.16 shows as an example the scatter plots of the values of these statistics computed for 200 simulations with the PSM, for one of the several subdivisions of the sky adopted in the analysis and for both the ICA and ILC methods. Note that, since ICA does not reach the convergence with a number of regions larger than six, the nine and ten regions analysis has been carried out only with the ILC code. The variance values have been corrected for the noise contribution, which is computed using the simulations themselves.

This statistical analysis provides interesting information.

Generally, for both the methods, the results derived from the full-sky analysis show different properties than those derived using a partition of the sky, highlighting the impact of the new approach on the results. However, for the regional analysis, ICA is characterized by a weaker dependence on the type of partition considered, than the ILC code.

Furthermore, Figure 6.16 shows that:

1. The ICA variance values of the output maps are the same than the noisy and debiased maps. Moreover, there is not any offset with respect to the input maps values. This seems to suggest that the output maps are generally not contaminated by residual foregrounds and that the correction itself is negligible: at most, it reduces the spread of the points distribution.
2. On the contrary, the ILC values of the variance of the output maps are different with respect to the other maps: they generally show an offset with respect to the values of the noisy and bias corrected maps, although the amplitude of the offset depends on the partition of the sky. The larger the number, the better the CMB reconstruction, the lower the residuals, whose estimation is given by the bias.
3. Correcting for the residuals, the values of the debiased maps become consistent with those of the CMB input maps and the noisy ones. Therefore, we can assume the bias to be a good estimation of the foreground leftovers in the CMB maps, at least for the simulations.
4. The distribution of the skewness of the returned maps is generally broader than the input ones. This is clear for the CMB output map since it contains residuals. The other two maps present the same behaviour because the noise contribution is not subtracted in the skewness computation. The residuals also affect the kurtosis values. Again, the noise is included in the contamination, since we did not apply any correction for it: the noisy maps and the bias corrected ones indeed show an offset as well, although smaller.
5. For both the methods, the offset shown by the skewness and kurtosis is larger for the full-sky (not shown) than for the regional analysis, which demonstrates that as expected the regional analysis has less residuals.

6. The kurtosis is particularly sensitive to the presence of residuals endorsing the utility of using different statistics to evaluate the quality of the CMB reconstruction. This is clear looking at the results derived with the PSM: the variance of the output maps does not show any offset, while instead, an offset is visible in the kurtosis distributions.
7. The kurtosis offset becomes lower when the bias correction is applied, although it does not completely disappear due to the noise contribution. Such a noise contribution to the offset is also visible looking at the ILC results, although it is considerably lower than one. While is more affected by noise, the ILC is more affected by foregrounds residuals.
8. Finally, the values of the quadrupole amplitude show a broad distribution without offset. The presence of residuals is reflected more in the spread of the points, than in a possible offset.
9. is unstable when the sky is subdivided in more than three regions. In particular, with nine and ten regions, it reaches the convergence only for a few percent of simulations. This instability is the signature of the fact that the component separation, when performed with the ICA approach, is not actually improved by selecting regions.

The results of the statistical analysis of the simulations derived using either the MEM foreground maps or those proposed by Ghosh et al. (2010) are quite different with respect to the PSM case. Figures 6.17 and 6.18 show the plots obtained from the three regions analysis, as an example of what happens also with all the other partitions.

In general, the and ILC results look very similar: there is always an offset in the distribution of the statistical values derived from the output CMB maps with respect to the noisy maps and the bias corrected ones. The MEM and GFM foreground models make the bias larger in amplitude than before.

Besides the noise is clearly less significant than for the PSM model, especially for the analysis: the values of the four statistics of the noisy maps are always consistent with those of the input maps. This is true also for the skewness and kurtosis values which were more spread for the PSM.

When the bias correction is applied to the and ILC output maps, the statistics give results in agreement with the input ones, as for the PSM case. Although the bias maps have different features due to the different models adopted, the computation is always good enough to evaluate the actual residual in the CMB.

### 6.3.2 Local evaluation of skewness and kurtosis

Attempting to estimate the impact of the foreground residuals on small angular scales, we have studied in further detail the values of the skewness and kurtosis of the CMB estimators derived from the simulations: these statistics have demonstrated to have a significant sensitivity to non-Gaussianities.

Specifically, instead of computing the global value, we computed the local value of these statistics, over the pixels of the map (at  $N_{side} = 256$ ) which lie in a given pixel of a low resolution map of  $N_{side} = 32$ . The analysis has been carried out only for the PSM simulations, whose CMB maps are largely contaminated by residuals.

For each CMB map recovered in a single simulation (input, output, noisy and bias corrected map), we generated the maps of the skewness and the kurtosis. Finally, we averaged them over the entire ensemble of simulations. The results for a selection of cases of analysis are shown in Figures 6.19 and 6.20. The analogous plots for the ILC code are shown in Figures 6.21 and 6.22.

Looking at them, it is possible to get new hints about the properties of the CMB maps, especially for what concerns the impact of residuals on angular scales smaller than 2 degrees and the applied bias correction.

As expected, given the Gaussian nature of the simulated signal, the input CMB generates regular and uniform skewness and kurtosis maps: the values of both the statistics, indeed, are uniformly distributed over the sky. Of course, this is not the case for the output maps which are contaminated by foreground residuals. The latter are reflected in larger values of the skewness and kurtosis, mainly along the Galactic plane. It is interesting to note, that such a contamination is larger for the ILC analysis than for the one (when it does not have convergence problems), for both the skewness and kurtosis. As already seen with the scatter plots, the CMB maps returned by the ILC are generally more contaminated by spurious emissions, than the maps derived with .

For the partitions of the sky where is unstable, the amplitude of the kurtosis increases even more when computed close to the Galactic plane, due to the foreground residual. Moreover, the edges of the partition of the sky are clearly visible, significantly so when the cut parallel to the Galactic plane is adopted. It is remarkable that the skewness does not show the same irregularities due to the cut, confirming the kurtosis as a more sensitive tool to detect residuals and anomalies with respect to a Gaussian random field. A similar pattern is visible in the kurtosis maps of the noisy CMB, since the noise is evaluated according to the same weights.

Another interesting point is given by the statistics maps of the bias corrected CMB. For both the methods, they are similar to those derived from the noisy maps. Furthermore, when the ILC is used, the similarity is extended to the input CMB as well. This tells us that the bias correction is a pretty good estimation of the residuals in the CMB maps, as already noted studying the global values of the statistics. Though, we need to stress again that it is not necessarily a good estimation for the residuals in the real data analysis results. Indeed, in Section 6.6, we will demonstrate the limitations of the PSM in evaluating the bias, with respect to the MEM and GFM maps.

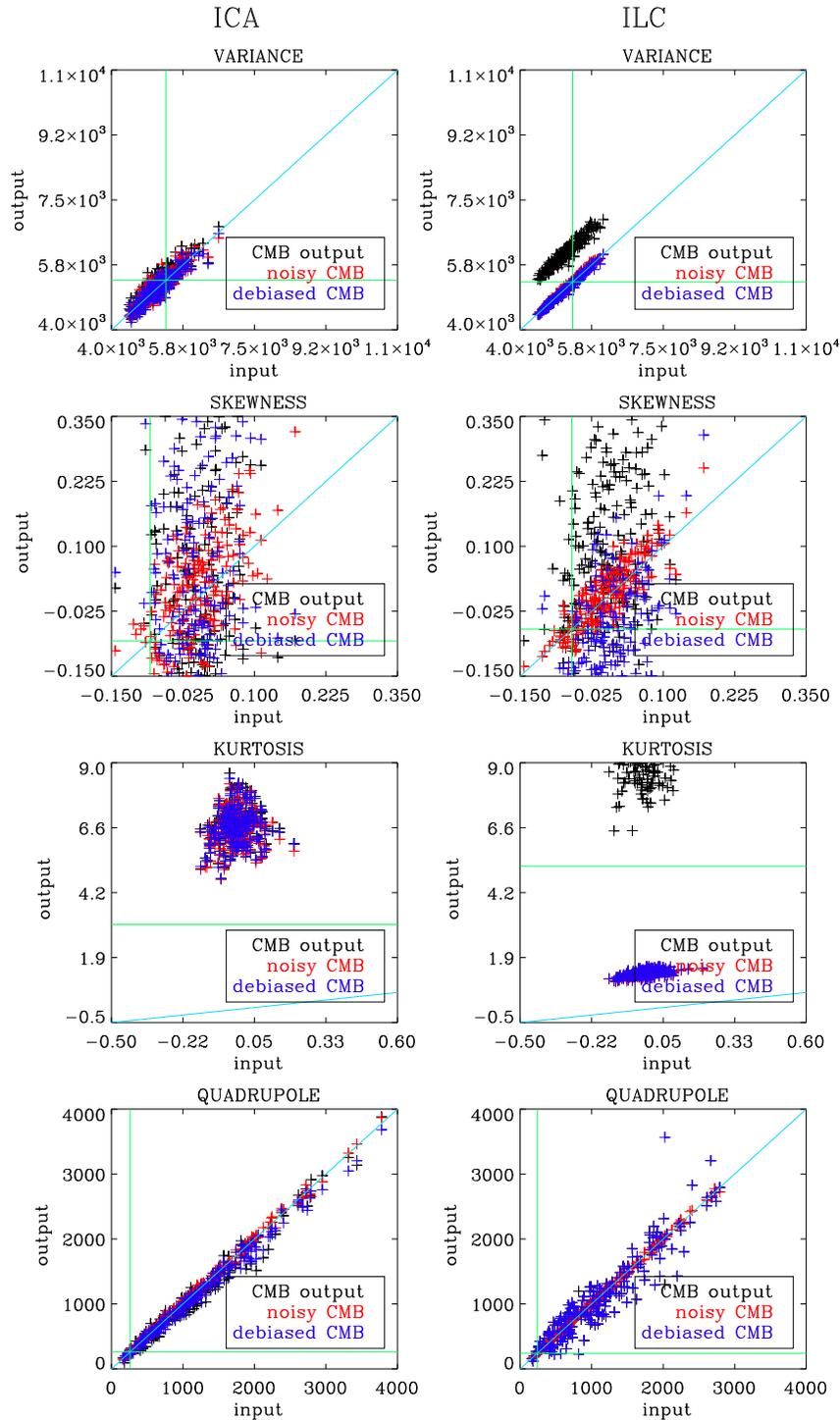


Figure 6.16: **Three regions component separation analysis with respect to the Ecliptic plane (fifth map of Figure 6.1).** Scatter plots of the variance, skewness, kurtosis and quadrupole amplitude of the CMB maps produced with 200 simulations. The values for the output CMB map, as well as the noisy and bias corrected output CMB maps are plotted against the values of the input CMB maps. The plots on the left are relative to the analysis performed with  $\dots$ , while those on the right are the analogous for the ILC method. The values derived from the real data analysis are shown in green.

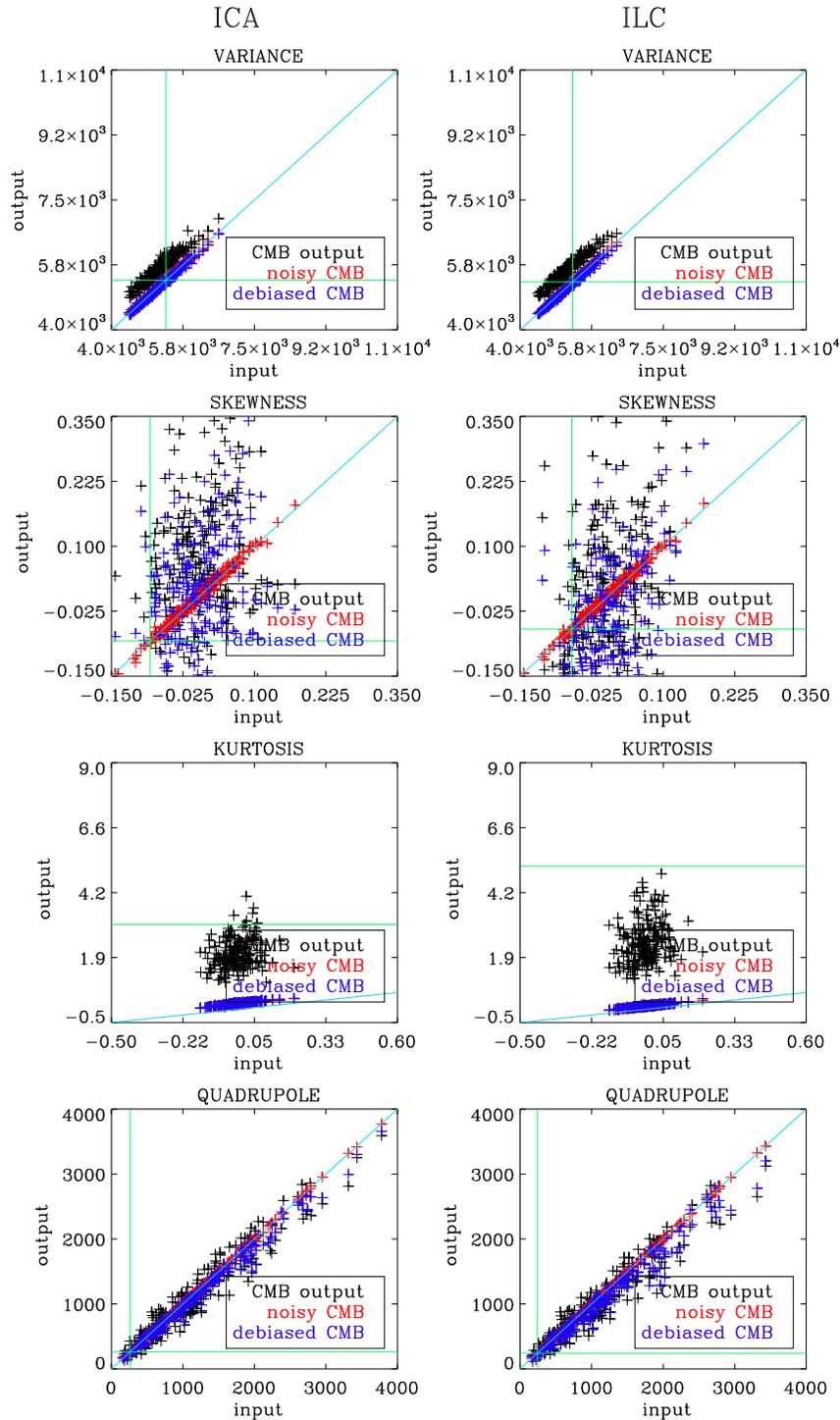


Figure 6.17: MEM simulations - Three regions component separation analysis with respect to the Ecliptic plane (fifth map of Figure 6.1). Scatter plots of the variance, skewness, kurtosis and quadrupole amplitude of the CMB maps produced with 200 MEM simulations. The values for the output CMB map, as well as the noisy and bias corrected output CMB maps are plotted against the values of the input CMB maps. The plots on the left are relative to the analysis performed with  $\beta = 0.5$ , while those on the right are the analogous for the ILC method. The values derived from the real data analysis are shown in green.

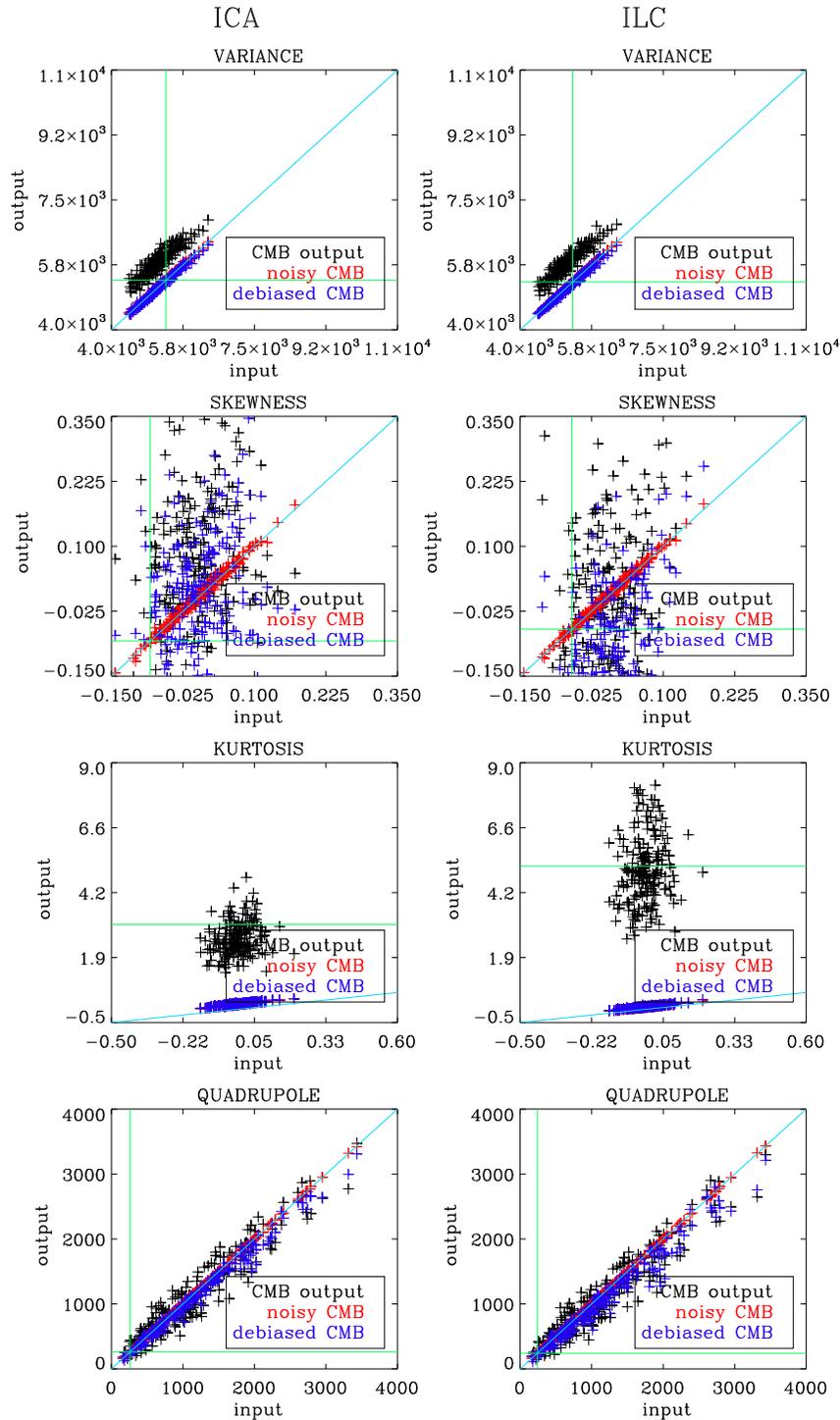


Figure 6.18: **GFM simulations - Three regions component separation analysis with respect to the Ecliptic plane (fifth map of Figure 6.1).** Scatter plots of the variance, skewness, kurtosis and quadrupole amplitude of the CMB maps produced with 200 GFM simulations. The values for the output CMB map, as well as the noisy and bias corrected output CMB maps are plotted against the values of the input CMB maps. The plots on the left are relative to the analysis performed with  $\dots$ , while those on the right are the analogous for the ILC method. The values derived from the real data analysis are shown in green.

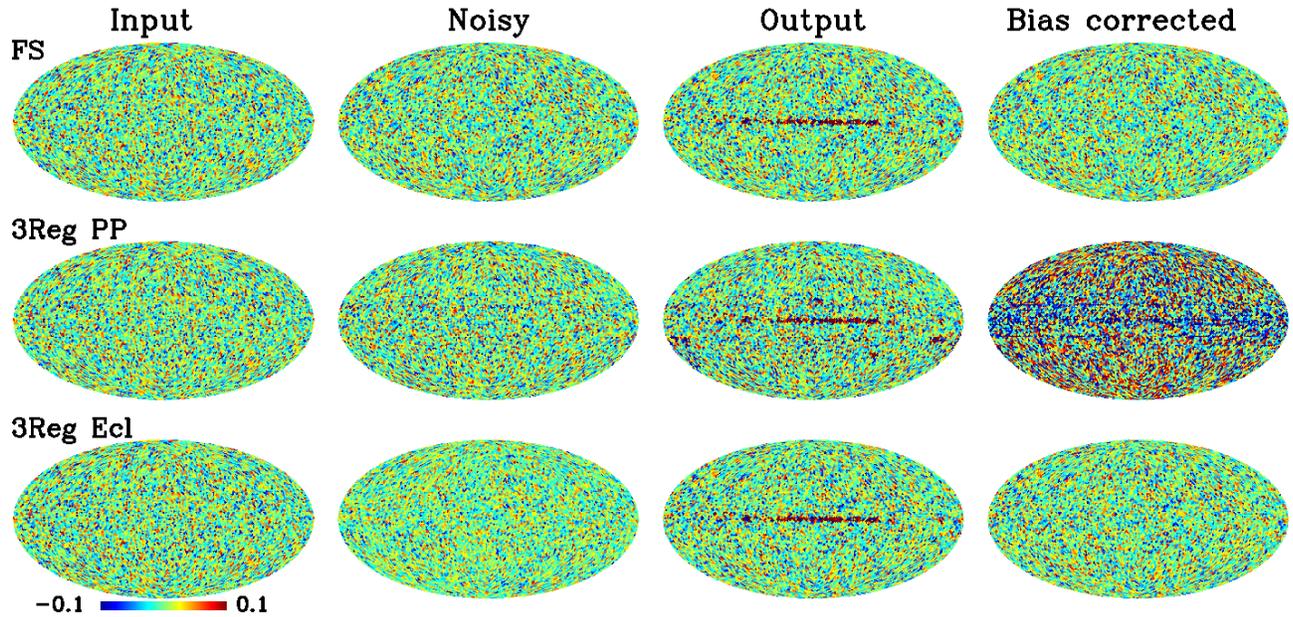


Figure 6.19: Mean maps of the skewness computed from the CMB maps derived from the simulations. The skewness values are computed for each pixel of a map at a resolution of  $N_{side} = 32$ . The CMB maps are derived from the analysis of simulated data. As example, we show the results of the full-sky analysis, and those obtained with three regions, i.e. the three regions defined with respect to the Galactic and Ecliptic planes.

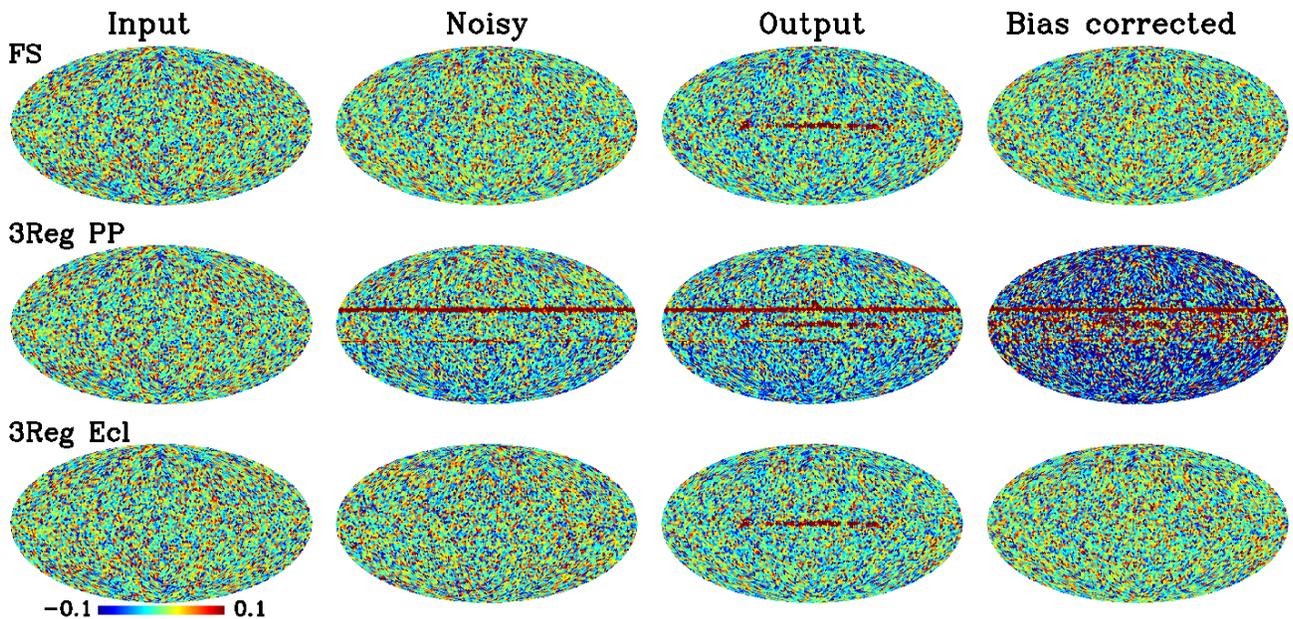


Figure 6.20: Mean maps of the kurtosis computed from the CMB maps derived from the simulations. The skewness values are computed for each pixel of a map at a resolution of  $N_{side} = 32$ . The CMB maps are derived from the analysis of simulated data. As example, we show the results of the full-sky analysis, and those obtained with three regions, i.e. the three regions defined with respect to the Galactic and Ecliptic planes.

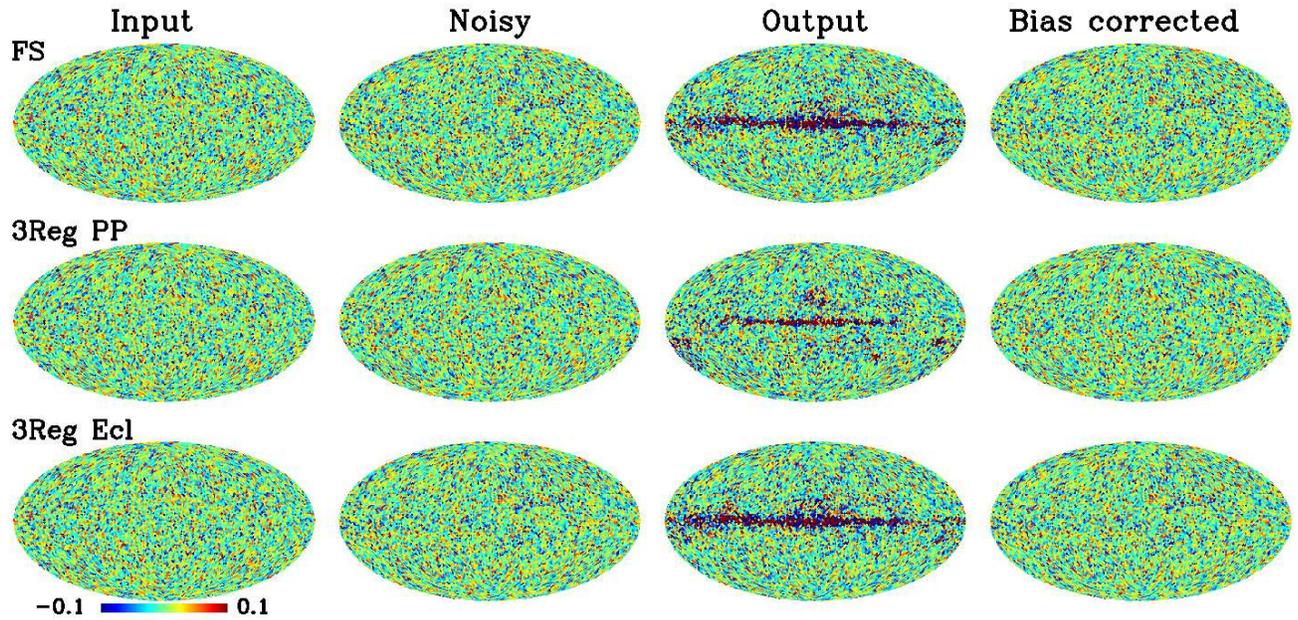


Figure 6.21: Mean maps of the skewness computed from the CMB maps derived from the simulations. The skewness values are computed for each pixel of a map at a resolution of  $N_{side} = 32$ . The CMB maps are derived from the analysis of simulated data. As example, we show the results of the full-sky analysis, and those obtained with three regions, i.e. the three regions defined with respect to the Galactic and Ecliptic planes.

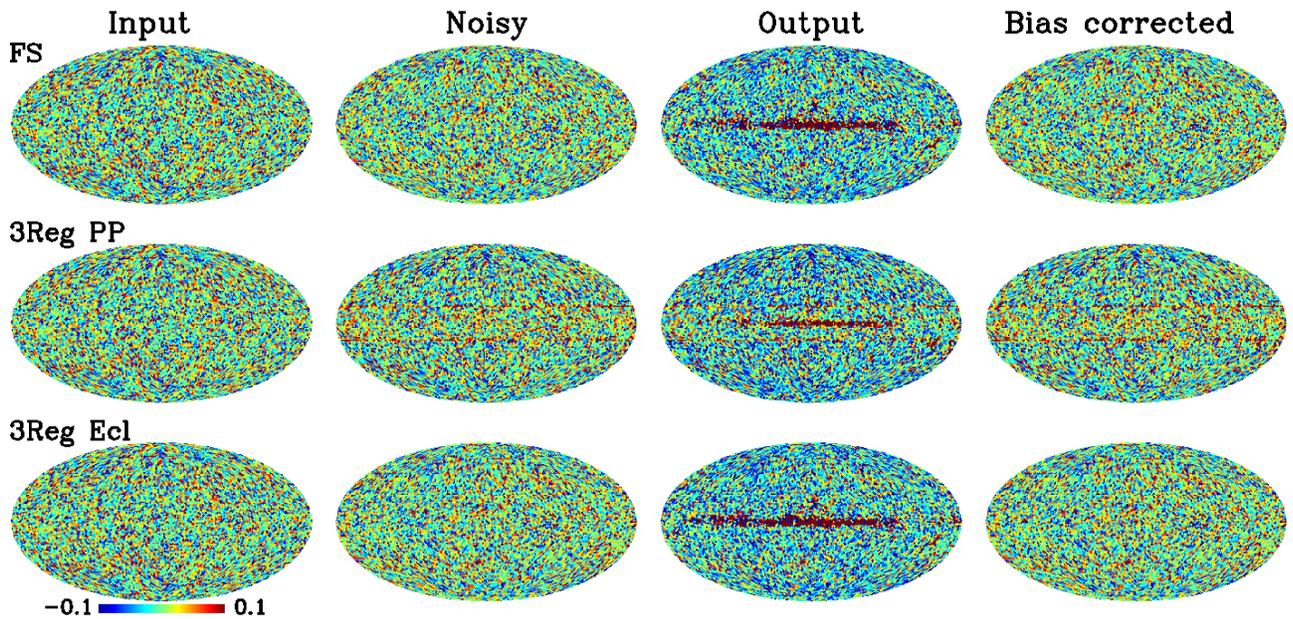


Figure 6.22: Mean maps of the kurtosis computed from the CMB maps derived from the simulations. The kurtosis values are computed for each pixel of a map at a resolution of  $N_{side} = 32$ . The CMB maps are derived from the analysis of simulated data. As example, we show the results of the full-sky analysis, and those obtained with three regions, i.e. the three regions defined with respect to the Galactic and Ecliptic planes.

## 6.4 Analysis of the *WMAP* five-year data

Our data set consists of the *WMAP* 5-year maps. As in the case of the simulations, they have been analysed on patches of the sky, accordingly to different selections of regions. Specifically, we have considered six different subdivisions of the sky: four of them have been chosen among those used in the testing phase, while two of them have been selected from the literature, namely the 12 ILC regions from *WMAP* (Hinshaw et al., 2007) and the 400 regions defined by Park et al. (2007). The analysis has been performed both with  $\text{ICA}$  and ILC as implemented by Eriksen et al. (2004).

Unlike previous works on this data set (see Maino et al. (2007) and Chapter 4), no masks have been applied to the data, nor has there been a subtraction of the point sources, the main reason for this choice being the possibility of obtaining a full-sky reconstruction of the CMB. In fact, we already know from Chapter 4 that the algorithm is able to achieve good results even under these conditions.

As for the simulations, we computed the CMB component for each region, and afterwards we merged them in a full-sky map. The merging has been done simply adding the maps together, as well as previously smoothing the edges of the regions with an effective beam of  $1.5^\circ$ . As example, in Figures 6.23, 6.24, 6.25 and 6.26, we show the resultant full-sky CMB maps together with the CMB components of each single region, in the case where  $\text{ICA}$  has been used for the analysis and for different cases of partitioning of the sky. In addition, Figure 6.27 shows the merged CMB maps obtained with the two codes, for the four subdivisions of the sky, when the edges of the regions are smoothed.

A first glance at the full-sky CMB maps is sufficient to observe some differences between the  $\text{ICA}$  and ILC results. With the first method indeed, there are generally some residuals left along the Galactic regions which increase in amplitude when the sky is divided in 6, 9 and 10 regions. Indeed, the Cyn A region as well as a thinner layer in the central part of the Plane are particularly prominent when the sky is divided in more than 6 regions. Visually speaking, only the CMB derived from the three regions analysis with  $\text{ICA}$  is comparable with the ILC one. This result confirms what already found with simulations:  $\text{ICA}$  is unstable when used on small regions of the sky and gets worse in performance with respect to the ILC code. This is probably also connected to the small extension of the regions closer to the Galactic plane (for example, region 5 and 6 in Figure 6.23). Due to the strong mixture of foregrounds of these regions, it is very difficult to disentangle the component and it is not perfectly achieved by the code. In fact, the statistics adopted in the computation itself enhance those features of the sky which are already strong, to the detriment of the CMB emission. At higher latitudes, instead, where the contamination becomes less significant, the CMB pattern looks clean and properly defined.

On the other hand, the ILC estimation is remarkably improved by patching the sky: this is clear when comparing the CMB map derived from the full-sky analysis with those obtained using the regional approach. Moreover, the larger the number of regions, the better the results: the residuals along the Galactic plane gradually reduce with an increasing number of subdivisions.

To better compare the ILC and ICA results, we have computed the difference between the respective CMB maps (see Figure 6.28). For the full-sky and three regions analysis, the disparity is mostly constituted by a residual along the Galactic plane. With a larger subdivisions of the sky, instead, the difference increases in amplitude.

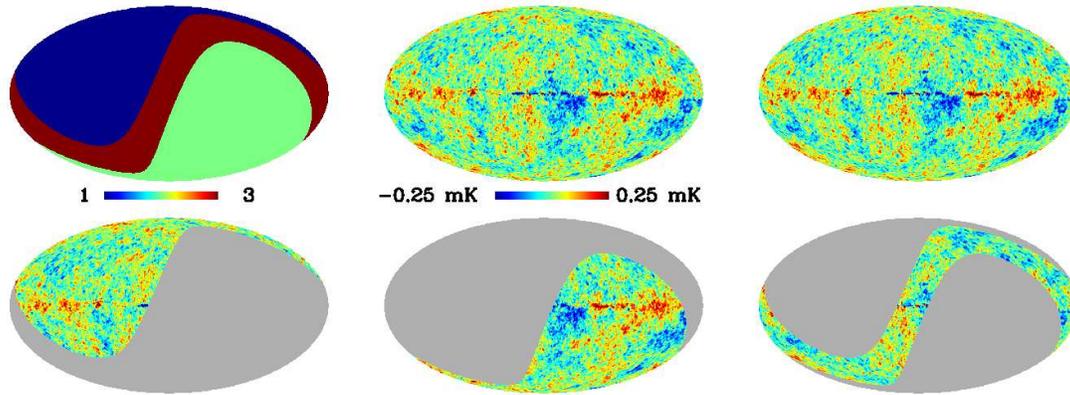


Figure 6.23: We show the three regions map (upper left). The CMB components returned by [Equation 6.10](#) for each region are also shown (second row maps) as well as the one obtained merging the regions. We merge the maps simply adding them together, but also previously smoothing the edges of the regions at an angular resolution of 1.5 degrees (see the middle and right maps in the top row, respectively).

The stronger signal comes from the Galactic plane, although the structure of the North Polar Spur is more visible. Finally, it is interesting to note that, when the sky is subdivided in 6 or 9 regions, the area around the Gum Nebula is well resolved by both the algorithms. It is not the same when the sky is divided in 10 patches and the region is analysed together with others closer to the Center. This is probably due to a stronger cross-correlation among the components and changes in the values of spectral indices.

The different behaviour of the two methods is also visible looking at the weights of the CMB, returned by the two algorithms. They reflect the effect of the constraint imposed on the CMB when using ILC, and the fact that both method employ different statistics.

As seen with simulations, a limitation of [Equation 6.10](#) is the fact that it does not reach the convergence if the regions of the sky are too small. This happens with the 12 regions proposed by *WMAP* and the 400 regions defined by Park et al. (2007). The number of pixels included in a single region is not statistically sufficient to make the code converge. For this reason, for these two partitions of the sky, we employed only the ILC algorithm: as mentioned before, these cases are interesting as a comparison with other works and because they give us the opportunity to explore more largely the dependence of the ILC results on the number of regions (see Section 6.4.2). The recovered full-sky CMB maps are presented in Figure 6.29: the CMB pattern looks clean of any residuals.

### 6.4.1 Power spectrum

Another figure of merit for evaluating the final CMB map is the power spectrum estimation. In Figure 6.30, we compare the power spectra of the full-sky CMB maps merged smoothing the edges of the regions. The results from the simple full-sky analysis are also included as reference, as well as the measured *WMAP* power spectrum.

As anticipated in the previous discussion (Section 6.4), the only case where it seems reasonable to

	K		Ka		Q	
	ILC		ILC		ILC	
	weights - FS					
	-0.01625	-0.01947	-1.22580	-0.01931	14.03021	-0.65080
	weights - 3 Regions					
1	-0.07322	-0.08682	0.31048	0.06388	-1.07516	-0.42592
2	-0.02643	0.02943	0.09745	-0.20285	-0.82779	-0.52696
3	0.10621	0.04027	-0.37649	-0.20133	-0.67149	-0.62444
	weights - 6 Regions					
1	0.26983	-0.04028	-1.09501	-0.39777	-0.21328	0.36204
2	0.16486	-0.18363	-1.02409	0.10310	0.29037	0.29051
3	0.35934	-0.08428	-1.44334	-0.15584	-0.07637	0.20719
4	0.15344	0.17263	-0.65040	-0.45023	-0.44628	-1.03254
5	0.12542	0.17570	-0.41078	-0.61689	-0.65726	-0.55256
6	-0.02354	-0.17896	1.54971	0.69843	-3.90176	-1.17644
	weights - 9 Regions					
1	-0.12033	-0.26980	-0.92638	0.25772	2.05783	0.29352
2	0.43760	-0.03364	-2.55416	-0.43984	2.07979	0.45829
3	-0.69573	0.04325	2.82638	-0.49383	-2.73613	0.05063
4	0.16223	-0.18379	-1.01775	0.10392	0.29491	0.28950
5	0.37001	-0.03362	-1.44859	-0.27517	-0.16775	0.06867
6	-0.00837	0.10286	0.01746	-0.28553	-0.87007	-0.93630
7	0.14200	0.15787	-0.36083	-0.56462	-1.13101	-0.69950
8	0.23731	0.14290	-0.84902	-0.18432	-0.34356	-1.51762
9	-0.00593	-0.03730	0.62716	0.20907	-2.05310	-0.99335
	weights - 10 Regions					
1	-0.12033	-0.26980	-0.92638	0.25772	2.05783	0.29352
2	0.47421	-0.11945	-2.75902	-0.05226	2.30928	0.12126
3	-0.69573	0.04325	2.82638	-0.49383	-2.73613	0.05063
4	0.17895	0.33723	0.12974	-1.54008	-2.47513	0.32035
5	0.16223	-0.18379	-1.01775	0.10392	0.29491	0.28950
6	0.37001	-0.03362	-1.44859	-0.27517	-0.16775	0.06867
7	-0.00837	0.10286	0.01746	-0.28553	-0.87007	-0.93630
8	-0.02712	-0.06103	-0.00655	-0.08511	-0.90176	-0.62822
9	0.23731	0.14290	-0.84902	-0.18432	-0.34356	-1.51762
10	-0.00593	-0.03730	0.62716	0.20907	-2.05310	-0.99335

Table 6.1: Values of the K-, Ka- and Q- bands weights returned by the and ILC codes for the single regions.

and ILC codes for the single regions.

	V		W	
	ILC		ILC	
	weights - FS			
	-28.06411	1.69522	1.76175	-0.00565
	weights - 3 Regions			
1	1.93295	1.37201	-0.09505	0.07685
2	1.65879	1.65676	0.09798	0.04362
3	2.12949	1.84492	-0.18772	-0.05942
	weights - 6 Regions			
1	3.39914	1.63095	-1.36069	-0.55493
2	1.38278	0.61614	0.18609	0.17388
3	3.08146	0.99252	-0.92109	0.04040
4	2.32883	2.89490	-0.38559	-0.58476
5	2.23501	2.23527	-0.29238	-0.24153
6	3.68580	1.36435	-0.31020	0.29262
	weights - 9 Regions			
1	0.29953	0.80988	-0.31065	-0.09131
2	2.18614	1.16352	-1.14936	-0.14833
3	1.14578	2.16641	0.45969	-0.76647
4	1.36927	0.61507	0.19134	0.17531
5	2.91220	1.02702	-0.66587	0.21309
6	1.99765	2.67398	-0.13668	-0.55501
7	3.21871	2.53187	-0.86886	-0.42563
8	1.90855	3.47086	0.04673	-0.91182
9	2.53081	1.70833	-0.09894	0.11325
	weights - 10 Regions			
1	0.29953	0.80988	-0.31065	-0.09131
2	2.17080	0.94815	-1.19527	0.10229
3	1.14578	2.16641	0.45969	-0.76647
4	4.15225	2.09217	-0.98580	-0.20967
5	1.36927	0.61507	0.19134	0.17531
6	2.91220	1.02702	-0.66587	0.21309
7	1.99765	2.67398	-0.13668	-0.55501
8	2.47438	2.48274	-0.53895	-0.70839
9	1.90855	3.47086	0.04673	-0.91182
10	2.53081	1.70833	-0.09894	0.11325

Table 6.2: Values of the V- and W-band weights returned by the `weights` and ILC codes for the single regions.

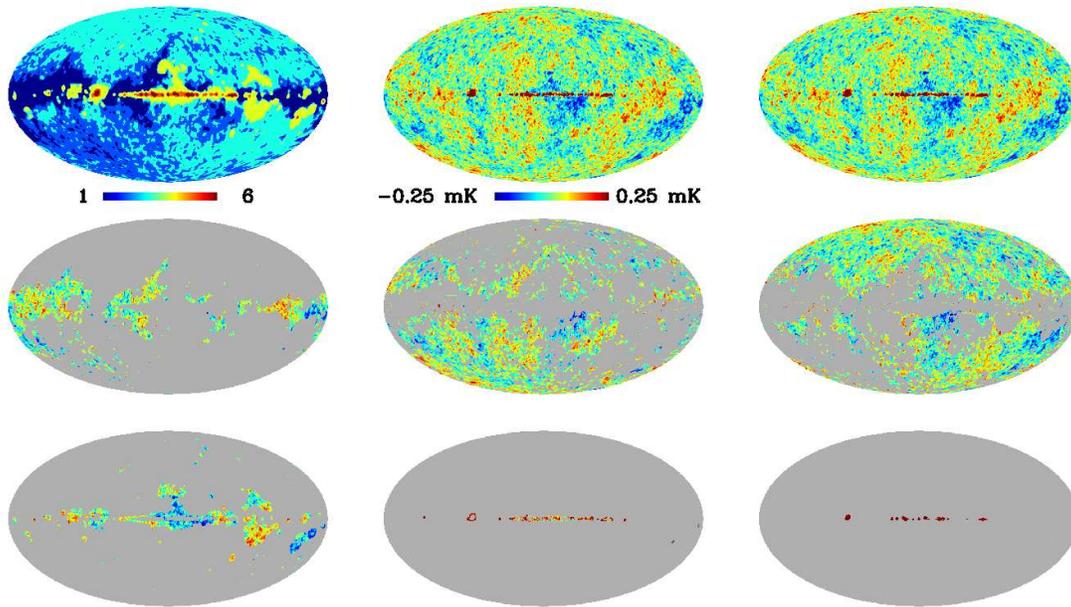


Figure 6.24: We show the six regions map (upper left) derived from the first component returned when the real data are analyzed internally with the  $g$  non-linear function. The regions are selected in order to have the same integrated temperature. Negative and positive pixels are treated separately. The CMB components returned by  $g$  for each region are also shown (second and third rows) as well as the one obtained merging the regions. We merge the maps simply adding them together, but also previously smoothing the edges of the regions at an angular resolution of 1.5 degrees (see the middle and right maps in the top row, respectively).

implement a regional analysis with  $g$  is the one where the sky is subdivided in three regions: the power spectrum is consistent (within  $1\sigma$ ) with the estimation of the *WMAP* science team and, at small angular scales, shows the same behaviour with respect to the full-sky analysis. On the other hand, the other three cases (6, 9 and 10 regions) show an excess of power for multipoles larger than 10, with respect to the full-sky analysis and even larger with respect to the *WMAP* result.

Any excess disappears as soon as the minimal mask defined in Chapter 4 is applied to the maps (see Figure 6.31). This demonstrates the fact that it is actually the presence of residuals along the Galactic plane that is responsible for the excess in the power spectra rather than, for example, an underestimation of the noise.

A further confirmation of this point is given when using the cross-power spectrum estimator, already defined in our previous analysis: as described in Chapter 4, we have adopted the two input CMB sky maps derived from subsets of the individual DA sky maps to compute it. In Figure 6.32, we compare the auto-power spectra of the maps derived from the regional analysis (and already shown in Figure 6.30) with the cross-power spectra derived using the same partitions of the sky. Since the cross-power spectra follow the same curve than the auto-power spectra, it seems clear that the excess of power is connected to the foreground residuals along the plane, rather than to the noise.

Finally, Figure 6.33 compares the  $g$  results with the ILC ones. As expected looking at the maps, the power spectra of the ILC analysis do not show any further excess of power, apart from the

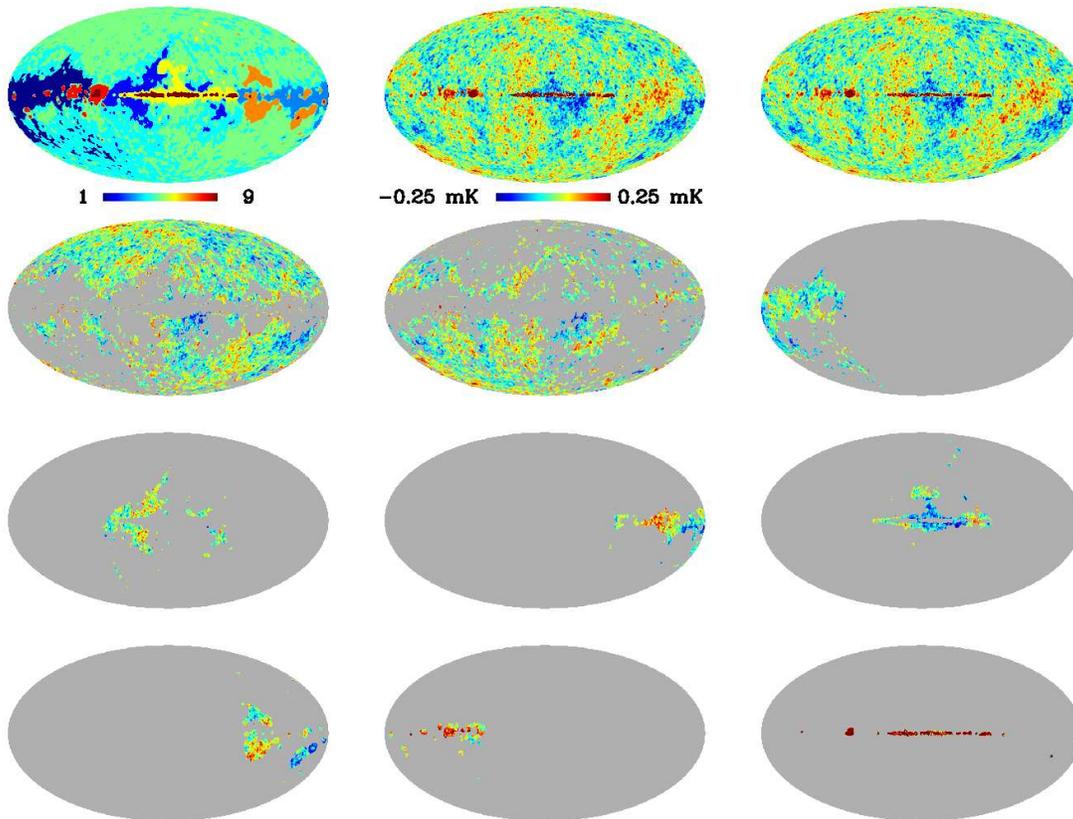


Figure 6.25: The same regions presented in Figure 6.24 are now further subdivided in order to separate the main structures of the plane. The final map has 9 regions (upper left). The CMB components returned by for each region are also shown (second, third and fourth rows) as well as the one obtained merging the regions. We merge the maps simply adding them together, but also previously smoothing the edges of the regions at an angular resolution of 1.5 degrees (see the middle and right maps in the top row, respectively).

usual one at small angular scales due to the vicinity to the resolution limit of the map. All the partitions of the sky return CMB maps whose power spectrum follows quite well the *WMAP* result and the full-sky estimation. However, the power spectrum alone is not sufficient to appreciate the improvements due to a large partition of the sky. Indeed, it is not easy to see any difference in the power spectra, connected to the smaller impact of the residuals in the maps. The power spectrum is remarkably robust to small changes in the foreground correction, as achieved by different partitions. This emphasizes why other figures of merit are necessary for comparisons. The simplest ones are the variance, skewness and kurtosis of the CMB maps, both globally and locally evaluated, which will be addressed in Sections 6.4.2 and 6.4.3. More sophisticated ones instead are measurements of non-Gaussianity as described in Section 6.4.4.

Note that all these power spectra have not been corrected for the unresolved point sources contribution, unlike in the analysis described in Chapters 3 and 4. In these previous works indeed we applied the correction as done by Wright et al. (2009), which provides an estimate according to global

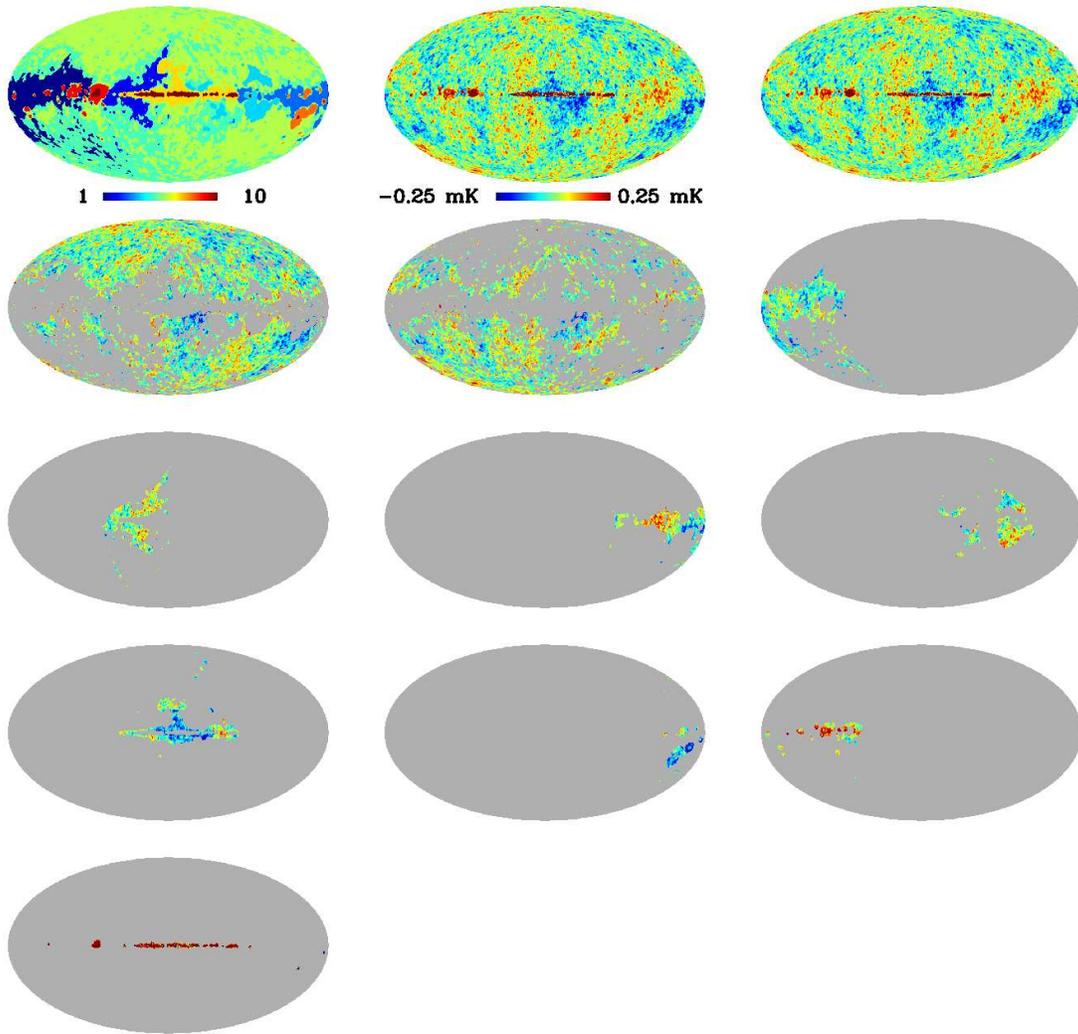


Figure 6.26: The same regions presented in Figure 6.24 are now further subdivided in order to separate the main structures of the plane. The final map has 10 regions (upper left). The CMB components returned by  $\text{CMB\_COMPONENTS}$  for each region are also shown (second to fifth rows) as well as the one obtained merging the regions. We merge the maps simply adding them together, but also previously smoothing the edges of the regions at an angular resolution of 1.5 degrees (see the middle and right maps in the top row, respectively).

linear weights and, therefore, can not be adopted again for the regional analysis. Nevertheless, we have computed the correction (following the same prescriptions) for the single regions used for partitioning the sky: it is useful in order to have an idea of the magnitude of the contribution of unresolved point sources and to identify the regions of the sky where this becomes more significant. The values derived for both the  $\text{CMB\_COMPONENTS}$  and ILC analysis are generally small enough not to affect the curves shown in the previous plots and the conclusions derived from them. However, the correction is significant when the weights are anomalous, such as in the case where  $\text{CMB\_COMPONENTS}$  is applied to full-sky data or it is used for the analysis with a number of regions larger than 6. Furthermore, there is not a specific trend in the order of the regions shown by the  $\text{CMB\_COMPONENTS}$  and ILC analysis: the largest contribution derives from different

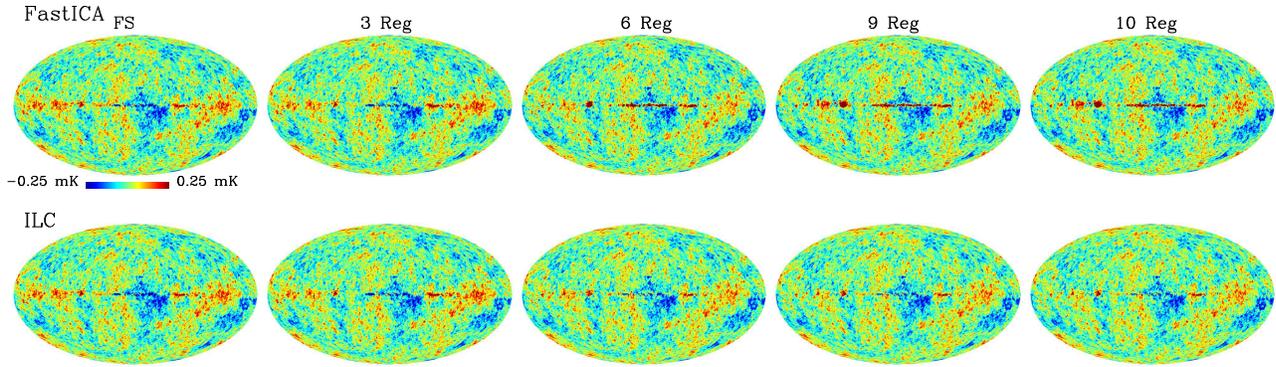


Figure 6.27: We compare the merged CMB maps derived from (already shown in Figures 6.23, 6.24, 6.25 and 6.26) with those derived with the ILC code. The edges of the regions are smoothed at the resolution of  $1.5^\circ$ .

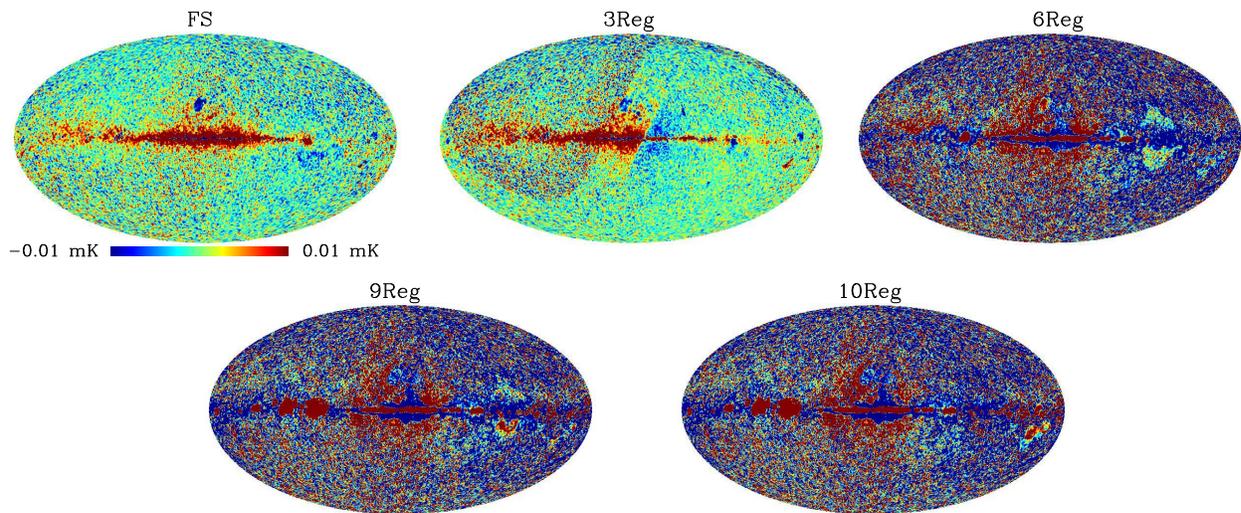


Figure 6.28: Difference between the and ILC merged CMB maps derived from the full-sky analysis and from the regional analysis, shown in Figure 6.27.

regions for the two methods.

## 6.4.2 Variance

For each CMB map derived with ICA and ILC, we evaluated the variance over the total number of pixels of the map (see Table 6.3). The computation has been made both for the case where the edges of the regions were smoothed before merging them, and where they were not. Moreover, we corrected for the noise variance, evaluated using noise simulations of the single bands of *WMAP* and combining them according to the weights of the specific CMB map considered.

The first interesting point to note is the fact that the values computed from the merged maps with the ILC code are smaller than the full-sky ones, indicating the regional analysis as a good tool to improve

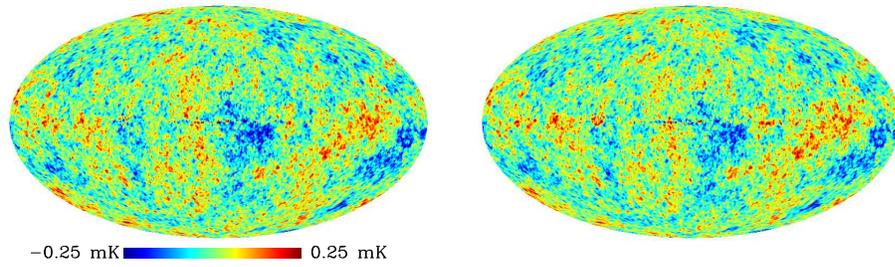


Figure 6.29: Plot of the full-sky CMB maps derived with ILC when the *WMAP* partition of the sky is adopted as well as the one proposed by Park et al. (2007).

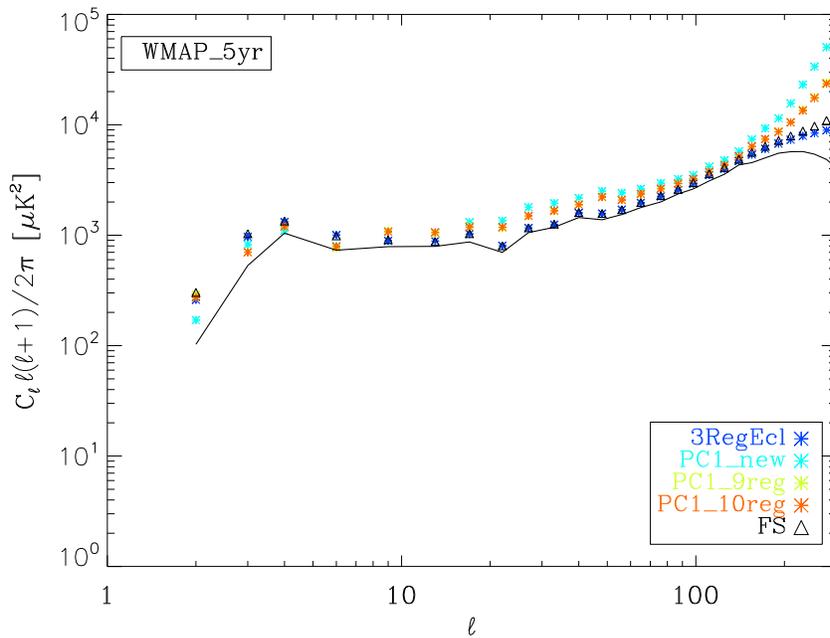


Figure 6.30: Power spectrum of the merged maps shown in the previous figures. We considered only the case where the edges of the regions have been smoothed.

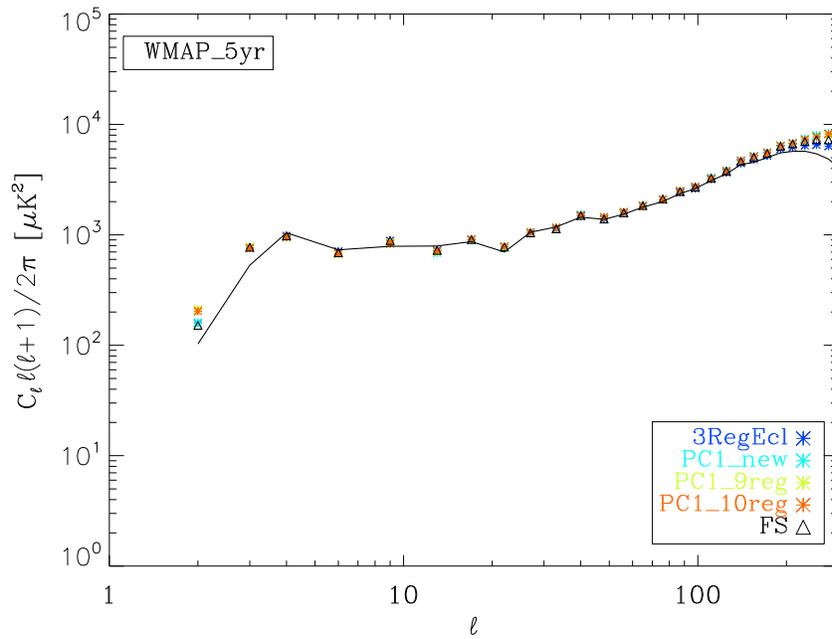


Figure 6.31: Power spectrum of the merged maps shown in the previous figures. The minimal mask is previously applied on the maps.

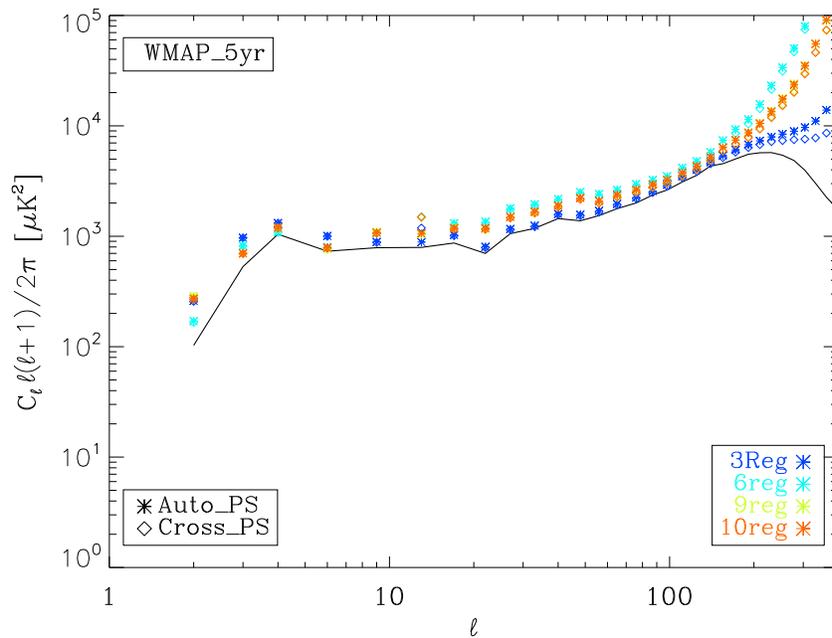


Figure 6.32: Comparison between the auto- and cross-power spectra of the merged maps shown in Figure 6.30. They are all derived from a full-sky analysis. The auto-power spectra are corrected for the noise. They basically overlap with the cross-spectra, suggesting that the noise computation is actually correct.

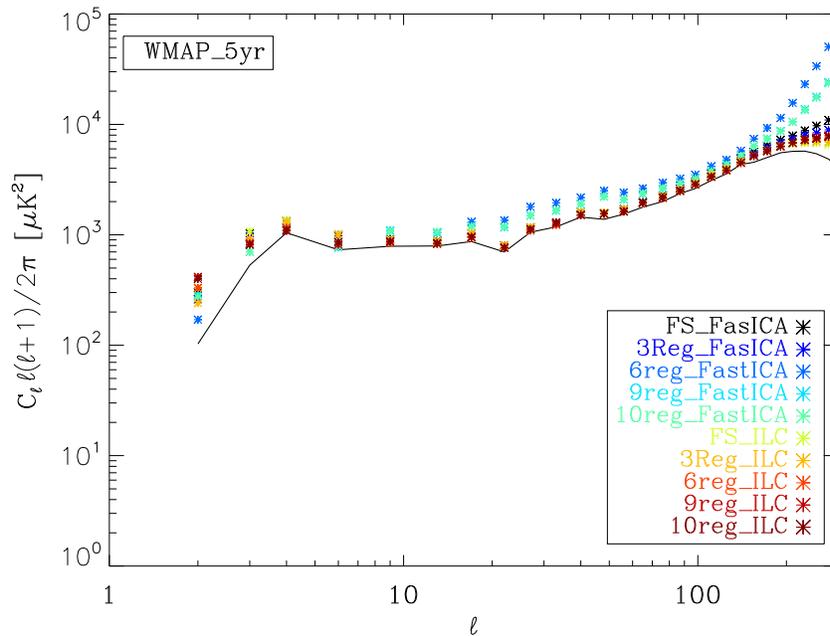


Figure 6.33: The same power spectra shown in Figure 6.30 are compared with the power spectra of the ILC maps. They are all derived from a full-sky analysis.

the quality of the CMB reconstruction. Using  $\chi^2$ , this is still the case for the three regions analysis, but not for the larger subdivisions of the sky. We have again the confirmation of what has been found with the simulations study:  $\chi^2$  does not dramatically improve with a regional analysis, while such an approach is more interesting and useful for the ILC method. Indeed, the variance values derived from the ILC results are generally lower than those computed from the  $\chi^2$  maps. This is also connected with the instability of the code, when it works on small areas of the sky, as seen with the simulations. Furthermore, this is a consequence of the different statistic used to find the CMB: the neg-entropy is approximated with an exponential function of the temperature so that bright regions become dominant in the sky.

Finally, it is interesting to note the trend in the variance of the ILC maps, which is such that the larger the number of regions, the lower the variance. The 9 and 10 regions analysis are the only exceptions, since the variance increases again. However, the two sets of regions are those derived by further subdividing by hand the 6 regions map, so that specific structures could be treated separately. The modification in the trend of the variance values suggests that these arbitrary partitions are not optimal: possibly, they do not reflect common properties of the structures. Respecting the actual spatial variations of the properties of the foregrounds is therefore mandatory to improve the performance in reconstructing the CMB. A large partition of the sky, does not necessary improve it, if such a requirement is not fulfilled.

Table 6.4 reports the values of the variance computed over the pixels of the single regions, with the

	with smoothed regions			
	ILC			
	<i>noise corrected</i>	<i>only map</i>	<i>noise corrected</i>	<i>only map</i>
<i>3Reg</i>	5335.31	5377.21	5296.31	5335.41
<i>6Reg</i>	6893.71	6984.89	4990.35	5028.29
<i>9Reg</i>	6154.47	6234.22	5009.29	5046.15
<i>10Reg</i>	6160.00	6239.89	5021.44	5058.53
<i>12Reg</i>	–	–	4800.19	4988.60
<i>400Reg</i>	–	–	4612.77	4677.68
	No smoothed regions			
	ILC			
	<i>noise corrected</i>	<i>only map</i>	<i>noise corrected</i>	<i>only map</i>
<i>3Reg</i>	5336.52	5378.42	5297.10	5336.21
<i>6Reg</i>	7577.02	7668.20	5049.48	5087.44
<i>9Reg</i>	6524.10	6603.85	5065.20	5102.06
<i>10Reg</i>	6533.63	6613.52	5081.46	5118.55
<i>12Reg</i>	–	–	4803.56	4991.97
<i>400Reg</i>	–	–	4683.42	4720.51

Table 6.3: Values of the variance (in  $\mu K^2$ ) of the merged CMB maps derived from the regional analysis performed with  $\chi^2$  and the ILC code. We show the values with and without the noise correction, for the maps derived smoothing the regions as well as those derived without applying the smoothing. For reference, the variance value of the CMB map derived from the full-sky are 5431.97 and 5390.41 respectively for  $\chi^2$  and ILC: the values computed from the merged maps are smaller than the full-sky ones, indicating the regional analysis as a good tool to improve the quality of the CMB reconstruction.

two component separation methods. This local computation is interesting to check which region has the largest variance value and, therefore, has the largest contamination by a spurious emission. We also compared these values with those derived from the CMB maps directly returned by a full-sky analysis. The difference between the variance evaluated on the single regions for both the maps, should give an idea of the possible improvement of the regional analysis: the ICA and ILC solution for the single area, indeed, are expected to be better than the one in the same region, but derived for the entire sky. In the latter case in fact, the solution is likely to be affected by compensations due to spatial variations of the foregrounds. Indeed, the values derived from the regional analysis are generally smaller than those from the full-sky analysis, although some  $\chi^2$  values are exceptions, due to the bad performance of the code with a number of regions larger than three. Furthermore, as predicted, the regions closer to the Galactic plane show the larger variance. The only exception is the three regions analysis, since all the patches include a fraction of the plane. In addition, this partition of the sky highlights the existence of an asymmetry between the northern and southern Hemisphere with respect to the Ecliptic plane: the southern hemisphere indeed has a larger variance than the northern one. Finally, the ILC values are generally lower than the  $\chi^2$  ones, as expected from the global values commented above.

	Reg	ILC Reg	FS	FS ILC
variance - 3 Regions				
1	4518.57	4459.84	4686.66	4571.69
2	6308.00	6288.65	6329.76	6337.36
3	5171.59	5128.14	5187.50	5168.01
variance - 6 Regions				
1	5249.85	4964.53	5535.66	5429.92
2	4794.53	4645.00	4811.95	4741.59
3	4669.69	4369.24	4485.33	4433.86
4	7912.21	7869.50	8605.50	8908.42
5	25969.79	17485.97	22203.82	24055.39
6	205724.72	46649.30	75595.66	72360.07
variance - 9 Regions				
1	4447.57	4264.94	5318.89	5170.01
2	4392.84	3884.38	4165.76	4043.50
3	8296.39	7406.10	7780.90	7732.63
4	4781.03	4634.58	4809.71	4739.33
5	4647.85	4387.68	4493.99	4442.38
6	7810.04	7237.90	7991.96	8088.57
7	8449.41	8342.66	8626.28	8716.38
8	8581.23	5891.20	7837.29	8415.88
9	56797.98	30324.54	38826.14	39327.83
variance - 10 Regions				
1	4447.57	4264.94	5318.89	5170.01
2	4468.48	3674.16	4002.02	3858.58
3	8313.26	7446.32	7780.90	7732.63
4	5451.64	5145.44	5809.27	5955.28
5	4783.29	4633.38	4809.71	4739.33
6	4649.75	4389.51	4493.99	4442.38
7	7817.43	7215.36	7991.96	8088.57
8	7449.21	7365.39	9245.57	9274.25
9	8581.26	5893.38	7837.29	8415.88
10	57646.19	30886.14	38826.14	39327.83

Table 6.4: Values in  $\mu K^2$  of the variance of the CMB maps derived from the regional analysis performed with and ILC. We show the values for the single regions, computed using the integrated CMB maps, as well as the maps derived from the full-sky analysis.

### 6.4.3 Local skewness and kurtosis

Together with the variance, we also studied the skewness and kurtosis of the CMB maps derived from the regional analysis. We evaluated them locally, repeating the procedure followed with the simulations (Section 6.3.2). For each CMB map, we computed the skewness and kurtosis values over the ensemble of points within each pixel of a map of resolution  $N_{side} = 32$ . The results are shown in Figure 6.34.

When the ILC method is used, the values of both the statistics are uniformly distributed on the sky, sign of the fact that on a small angular scale the presence of the foreground residual does not impact the properties of the CMB. This is particularly remarkable for the full-sky analysis, where the residuals are certainly present and strong. In the case of  $\ell < 10$  instead, both the skewness and kurtosis maps show an enhancement in amplitude along the Galactic plane, exactly where the residuals get stronger. This is true in particular for the analysis with a number of regions equal or larger than six. Thus also for the real data analysis it is true that the ILC applied on small regions of the sky improves its performance while it is not the case for  $\ell < 10$ , the CMB reconstruction of which is generally more contaminated by foreground residuals. The latter analysis hence, affects the statistics of the CMB signal even at small angular scales.

It is also interesting to note that the same study with simulations showed that the ILC method was more contaminated than  $\ell < 10$ . Such discrepancy, again, must be connected to the inadequacy of the simulated foreground sky compared to observation.

The effect of the residuals is more evident when considering a lower resolution ( $N_{side} = 8$ ) for the maps (see Figure 6.35). The number of pixels with a large value of skewness and kurtosis is generally lower for the ILC results than when  $\ell < 10$  is used to derive the CMB emission. The only exception, as noted before, is the analysis of the sky divided in three regions: this is the only case where the result are better than the ILC one.

### 6.4.4 Further non-Gaussianity tests

Other measures of non-Gaussianity have been used for comparing different CMB estimations, in order to better appreciate improvements achieved by the performed regional analysis with  $\ell < 10$  and ILC. Several anomalies have been claimed to affect the *WMAP* data and they have been partially associated with the presence of foreground residuals. Therefore, generally, better results in cleaning CMB data should be reflected in a smaller evidence of such anomalies, which otherwise could be considered genuine features of the data, if statistically significant.

Following Jaffe (2006), we focused our attention on large scale anomalies evaluated by means of different statistics. The T-statistic and the angular momentum dispersion (hereafter D-statistic) have been used as measure of planarity of a given multipole, as proposed by de Oliveira-Costa et al. (2004): they are two measures of the maximal percentage of the power of a given multipole  $\ell$  that can be attributed to  $|m| = \ell$ . They are defined respectively as:

$$t \equiv \max_{\hat{\mathbf{n}}} \frac{|a_{\ell-\ell}(\hat{\mathbf{n}})|^2 + |a_{\ell\ell}(\hat{\mathbf{n}})|^2}{\sum_{m=-\ell}^{\ell} |a_{lm}(\hat{\mathbf{n}})|^2} \quad (6.1)$$

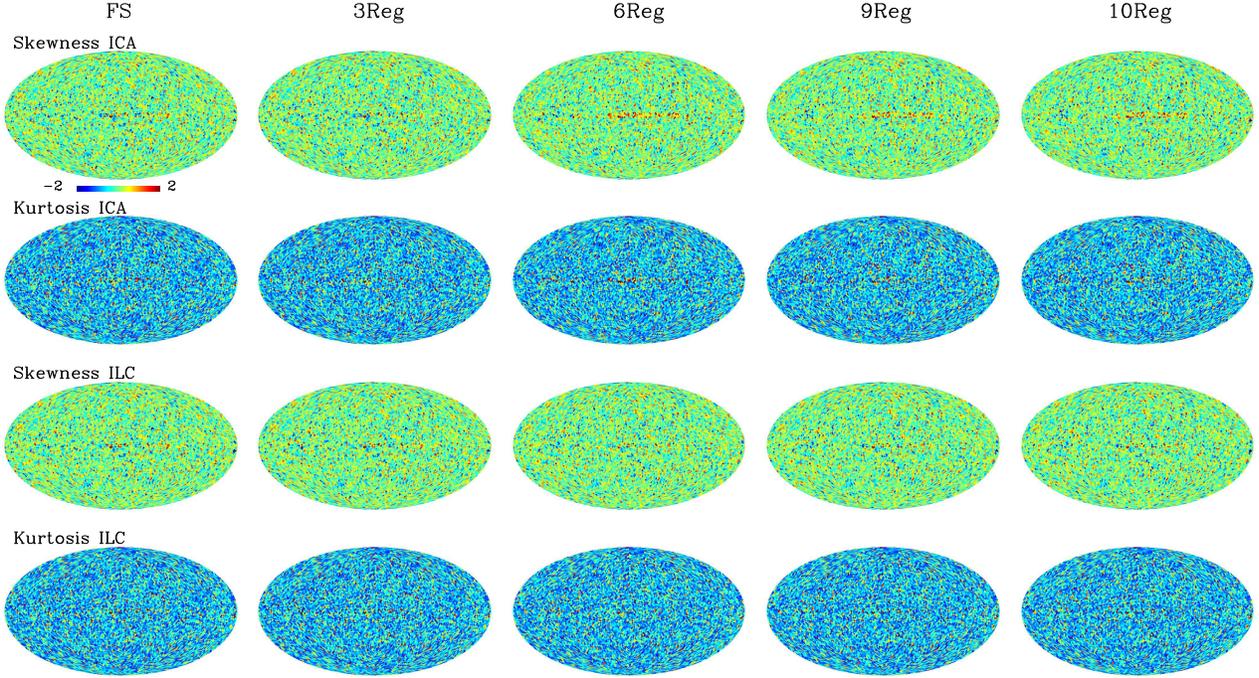


Figure 6.34: Maps of the local skewness and kurtosis estimation for all the cases of analysis (full-sky and using regions) with both the ICA and ILC codes. The maps have a resolution of  $N_{side} = 32$ .

and

$$\langle \psi | (\hat{\mathbf{n}} \cdot L)^2 | \psi \rangle = \sum_m m^2 |a_{\ell m}(\hat{\mathbf{n}})|^2 \quad (6.2)$$

where  $\frac{\delta T}{T}(\hat{\mathbf{n}}) \equiv \psi(\hat{\mathbf{n}})$  and  $L$  is the angular momentum.

Furthermore, the S-statistic is used as a measure of symmetry of CMB maps, as done by Starobinskij (1993) and more recently by de Oliveira-Costa et al. (1997):

$$S(\hat{\mathbf{n}}_i) \equiv \frac{1}{N_{pix}} \sum_{j=1}^{N_{pix}} \left[ \frac{\delta T}{T}(\hat{\mathbf{n}}_j) - \frac{\delta T}{T}(\hat{\mathbf{n}}_{ij}) \right]^2, \quad (6.3)$$

where  $N_{pix}$  is the number of pixels in the map and  $\hat{\mathbf{n}}_{ij}$  denotes the reflection of  $\hat{\mathbf{n}}_j$  in the plane whose normal is  $\hat{\mathbf{n}}_i$ , i.e.,

$$\hat{\mathbf{n}}_{ij} = \hat{\mathbf{n}}_j - 2(\hat{\mathbf{n}}_i \cdot \hat{\mathbf{n}}_j)\hat{\mathbf{n}}_i. \quad (6.4)$$

$S(\hat{\mathbf{n}})$  is a measure of how much reflection symmetry there is in the mirror plane perpendicular to  $\hat{\mathbf{n}}$ . The more perfect the symmetry is, the smaller  $S(\hat{\mathbf{n}})$  will be.

In addition, we also visually tested the Gaussianity of our maps, by using the color coded phase-gradient  $D_\ell = \varphi_{\ell+1,m} - \varphi_{\ell,m}$  (where  $\varphi_{\ell,m}$  is the phase, assuming  $a_{\ell,m} = |a_{\ell,m}| \exp(i\varphi_{\ell,m})$ ) proposed and adopted by Chiang et al. (2003): this is a visual technique which allows us to look for correlations that may or may not be quantifiable by the other three statistics. If the phases are randomly distributed the

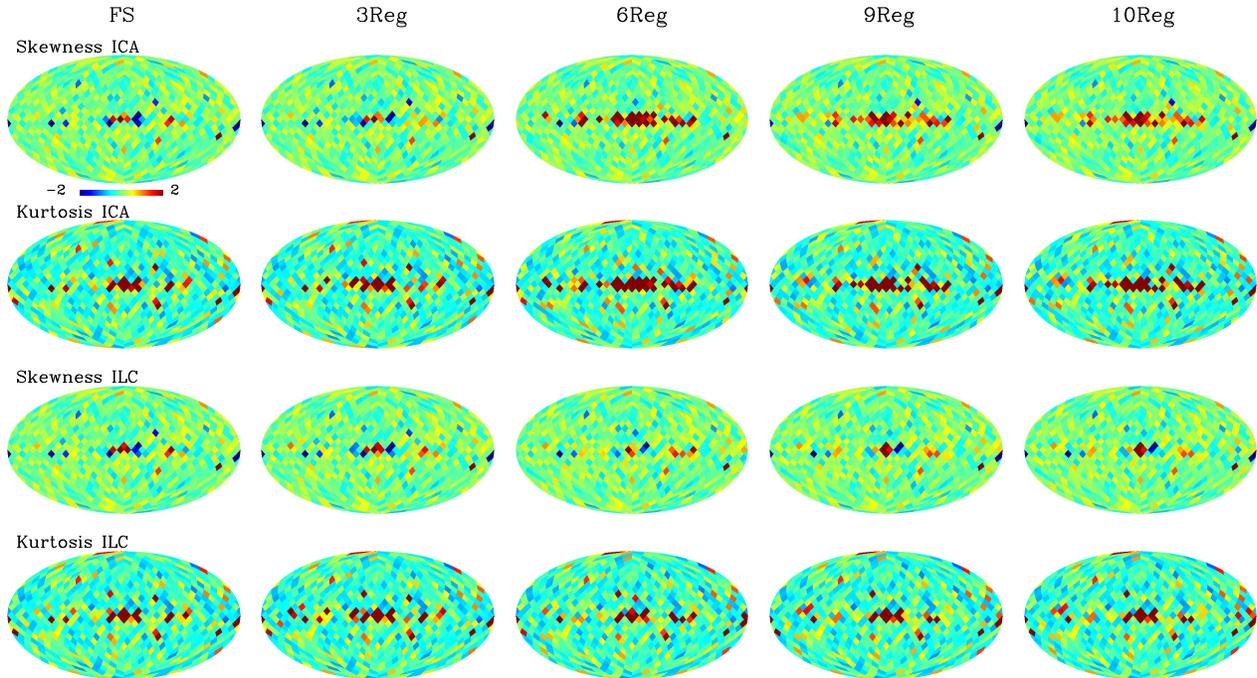


Figure 6.35: Maps of the local skewness and kurtosis estimation for all the cases of analysis (full-sky and using regions) with both the ICA and ILC codes. The maps have a resolution of  $N_{side} = 8$ .

plot should be filled uniformly. Deviation thereof gives a measure of Gaussianity.

All of them have been computed for the ILC and ICA CMB maps derived from the regions analysis and a simple full-sky analysis. The phase-gradient plots have also been generated.

We first evaluated the T-Statistic (i.e. planarity), and the D-Statistic (i.e. the angular momentum dispersion), for each multipole  $\ell$ , with a particular attention on the large scales (e.g.  $\ell = 2 - 17$ ). Furthermore, we determined the direction in the sky where the statistic reaches its maximum value, as well as the degree of alignment with the preferred direction of the previous multipole. To quantify the statistical significance of any anomaly, these results have been compared with those derived from a set of Gaussian Random field simulations. In practice, whenever the real data analysis returns a value larger than 95% of the simulations, it is considered an interesting indication of non-Gaussianity. Note, however that this non-Gaussianity reflects foreground residuals in the case of simulations. It is difficult however to assess the true non-Gaussianity in the real data since it is not certain whether the source of non-Gaussianity signals is residuals or otherwise. A perfect foreground model in the simulated data would allow us to quantify the relevant residuals, although it is obviously hard to provide such a model.

We first analysed the results derived with the ILC method, using different partitions of the sky (including a simple full-sky analysis): the returned maps, in this case, are expected to be closer to a Gaussian random field, since they are less contaminated by foregrounds. Then, we compared the results with those obtained using ICA.

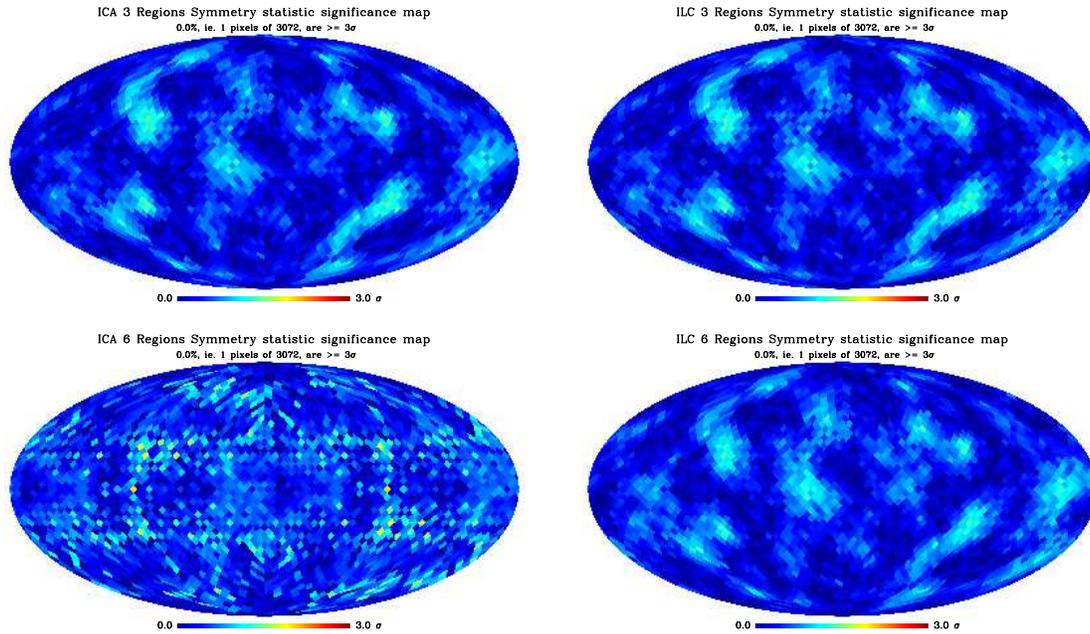


Figure 6.36: We show the symmetry statistic S-maps derived from the CMB ICA (left) and ILC (right) maps obtained partitioning the sky with 3 (top row) and 6 (bottom row) regions. All the maps are shown in the initial orientation of the input CMB map. The stronger foreground residual along the Galactic plane which affects the CMB estimation with 6 regions is likely to be the explanation for the different pattern on the S-map compared to the other cases.

## ILC MAPS RESULTS

Generally, particular attention has been given to the statistics results obtained for the quadrupole and the octopole, given the claims of alignment and planarity that concern them (see Section 1.3).

The analysis of the ILC maps shows that the quadrupole and the octopole have always a quite large value of the T-statistic, even though this is not significant at a level of 95% with respect to Gaussian simulations. For almost all the ILC maps, the maximum of planarity occurs at the same direction of  $b = 60^\circ$  and  $l = 238^\circ$ . Exceptions are the full-sky case the direction of which is rotated of 3 degrees, and the 3 regions case where the maximum occurs at slightly different position ( $b = 60^\circ$  and  $l = 229^\circ$ ).

On the other hand, the T-statistic is larger than 95% of Gaussian simulations for  $\ell = 6$  and  $\ell = 17$ , for all the ILC input maps. In the same way, the  $\ell = 13$  multipole has a significant level of planarity for all the partitions of the sky, but for the case where 10 and 12 regions are used. The octopole is aligned with the quadrupole for the 3, 6, 9, 10 and 12 regions analysis. Finally, the multipole  $\ell = 6$  is aligned with  $\ell = 5$  when the 12 regions from *WMAP* are adopted.

The D-statistic shows similar results. In particular, the direction of the alignment between the quadrupole and the octopole is the same than the one derived using the T-statistic (the only exceptions are the ILC maps derived using 6 and 12 regions).

The symmetry statistic S-maps derived from the CMB estimation using ILC are presented in Figure 6.36: here, we show the maps derived from the 3 and 6 regions analysis. The S-map has been used by

de Oliveira-Costa et al. (2004) as a figure of merit to evaluate the symmetry of the CMB pattern. They showed that there is a dark spot in the supposed direction of suppression of the original map, which is the same preferred direction of the alignment between the quadrupole and the octopole. All our ILC maps seem to return very similar S-maps, which show the same dark spot than de Oliveira-Costa et al. (2004), therefore confirming their observations.

Finally, Figures 6.37 and 6.38 show the phase gradients plots of the ILC maps derived using all the partitions. They are always affected by very large correlations between phases, except than when the 12 and 400 regions are adopted. Strong stripes are in fact visible in all the plots, particularly so for the ILC maps derived using 6, 9 and 10 regions. The analysis made by Jaffe (2006) with different toy models showed that these features are generally produced by point sources residuals along the Galactic plane. In fact, the residuals which affect our CMB maps along the Galactic plane look like point sources, although slightly broadened. Moreover, the larger the number of regions, the lower the residual which settles itself in this point source shape. Therefore, it is not surprising to see such strong features in the plots derived using the 6 regions ILC map, which in principle should be less affected by residuals, based on the variance study (see Section 6.4.2). In this case, we have a signature of non-Gaussianity which is most likely induced by the specific morphology of the residual, rather than the amount of it.

Finally, the ILC maps obtained from 12 and 400 regions, show remarkably less correlations and therefore, signature of non-Gaussianities due to foreground residuals: this is not surprising since they are generally less contaminated by spurious emissions and do not show any point-source like structures.

## ICA MAPS RESULTS

Although, the ICA CMB estimations show generally similar results in terms of high order statistics as those derived with the ILC maps, some differences have to be highlighted.

We know already that the CMB maps returned by ICA using 6, 9 and 10 regions are actually more contaminated by foregrounds than the ILC maps. Such a contamination shows its effect in the S-maps (see Figure 6.36), which have some spurious features with respect to those derived from the ILC maps. Furthermore, the phase gradient of these maps shows very strong stripes (see Figures 6.39 and 6.40), although they are weaker than those shown by the ILC analysis, since the residuals are more extended.

Looking at the T- and D-statistics, as a general trend, the ICA results seems to be more consistent with simulations than the ILC ones. As before, according to the T-statistic, the octopole shows larger planarity although the value of the maximum of the statistic is never larger than 95% of the Gaussian simulations. However, the preferred direction equal to  $(b,l)=(60,248)$  is generally closer to the one found by de Oliveira-Costa et al. (2004) (they found  $(b,l)=(60,250)$ ). The longitude is 10 degrees closer than for the ILC case. Finally, the alignment with the quadrupole is significant at the level of 95% only for the 6 regions analysis.

## NON-GAUSSIANITY ANALYSIS OF SIMULATED CMB MAPS

The statistical significance of all the results presented in the previous section has been tested by processing the CMB maps derived from the simulations described in Section 6.3 in the same way

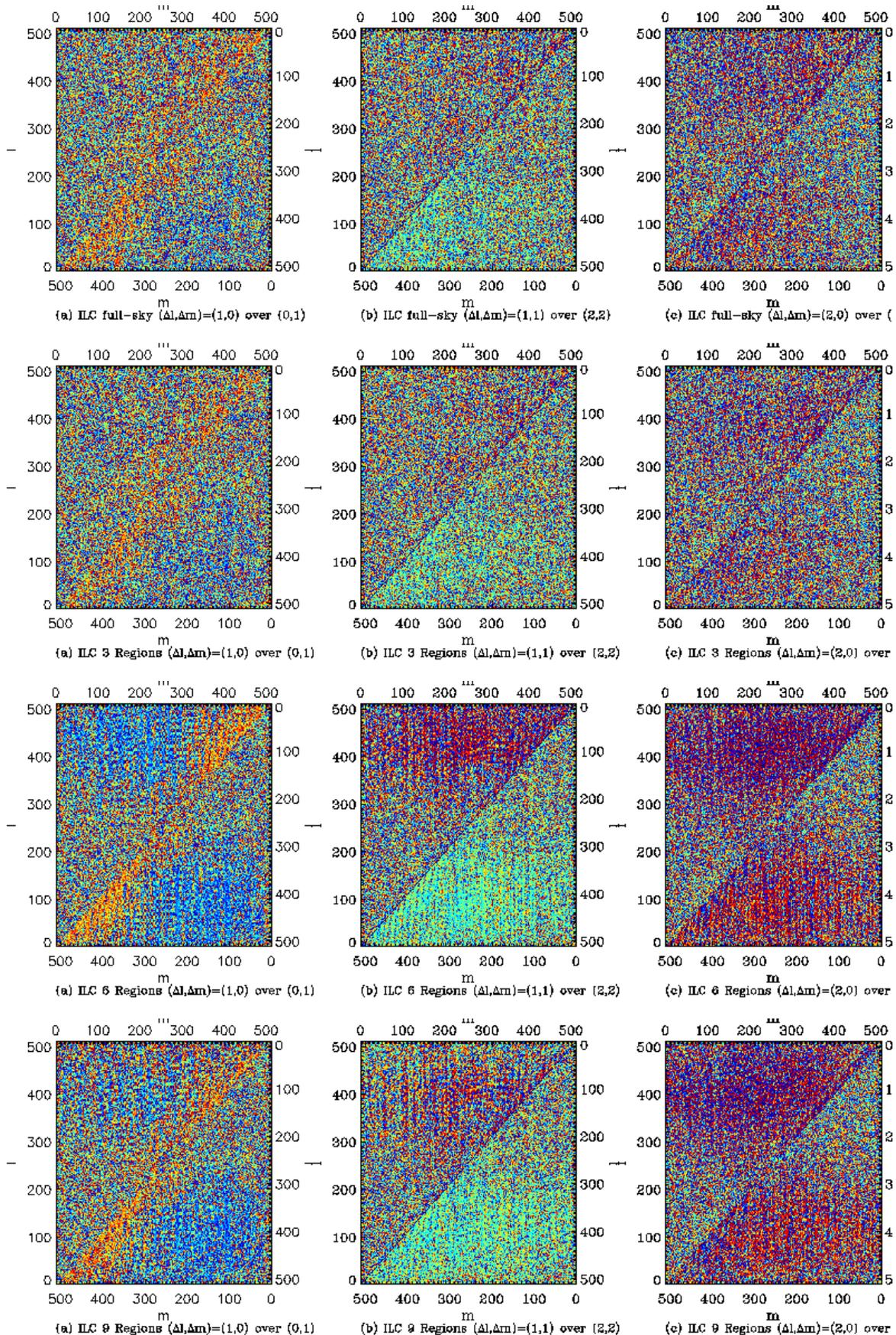


Figure 6.37: Color-coded phase gradients (following Chiang et al. (2003)) of the CMB maps derived with ILC, from the full-sky analysis (top), the 3 regions analysis (second row), the 6 regions analysis (third row) and 9 regions analysis (bottom).

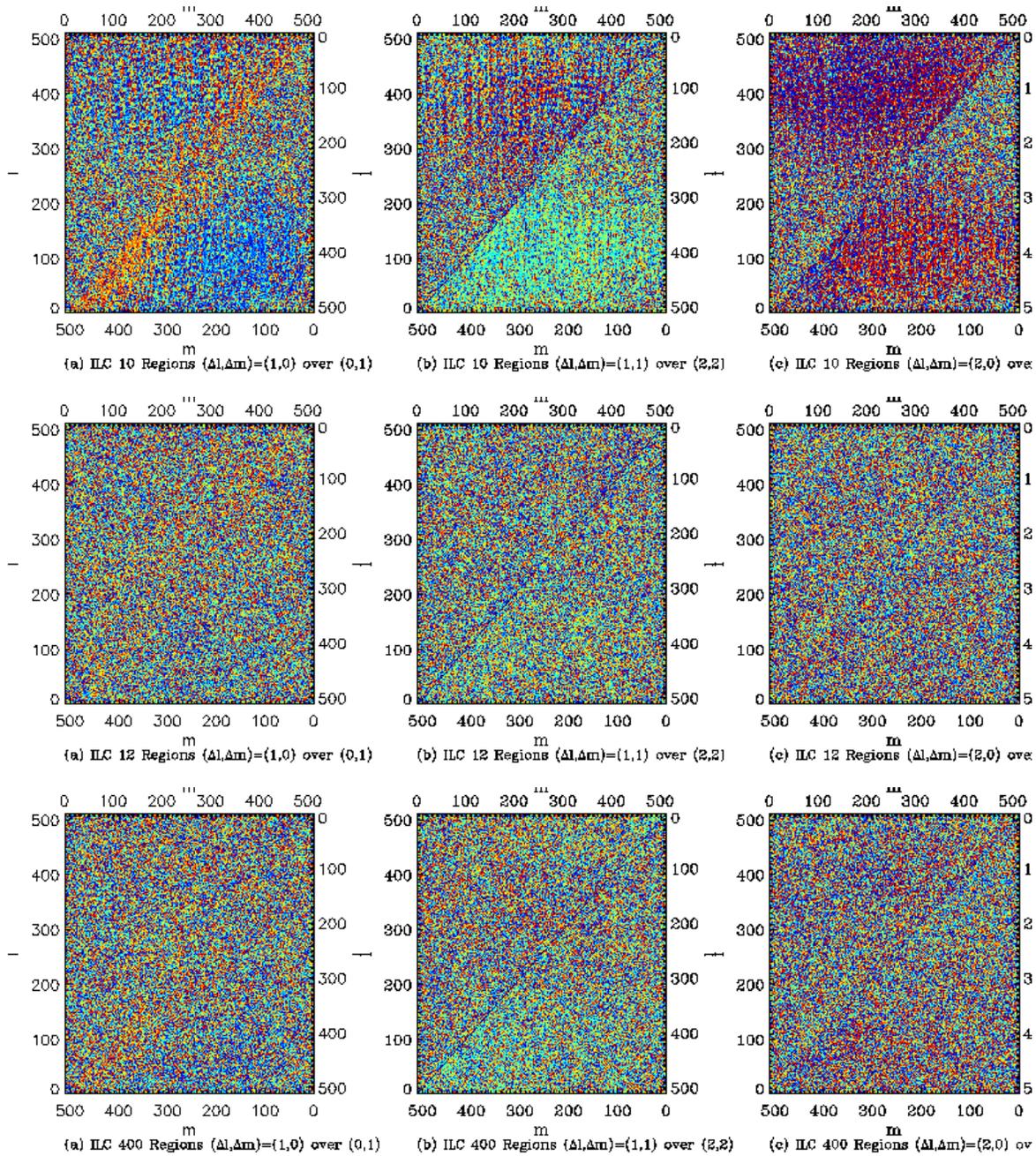


Figure 6.38: Color-coded phase gradients (following Chiang et al. (2003)) of the CMB maps derived with ILC and different partitions of the sky. On the top row we show the results from the 10 regions analysis, in the middle row the results from the 12 regions analysis, on the bottom row those from the 400 regions analysis. Dark blue along the diagonal or the border represent modes for which the given gradient is not defined.

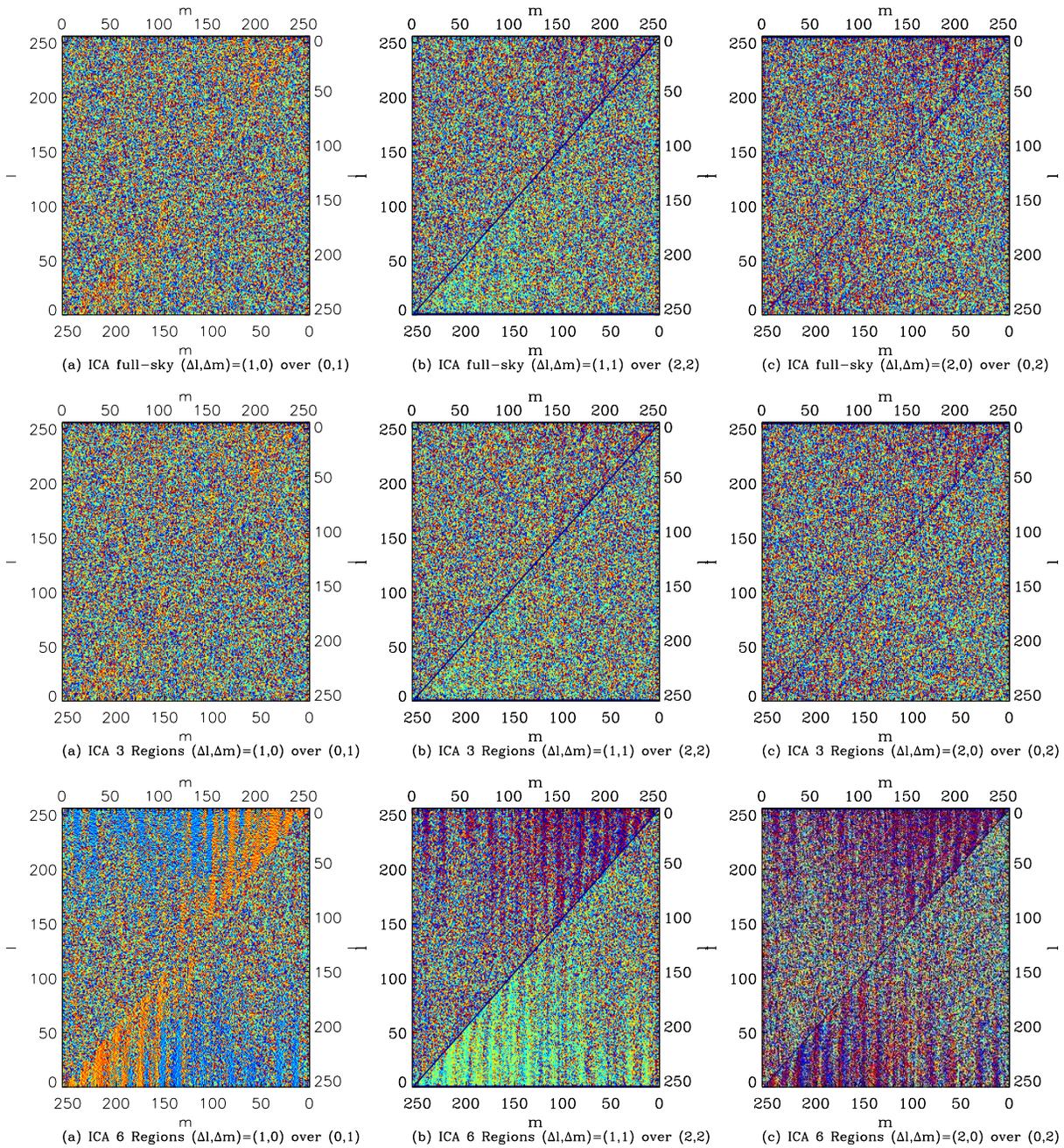


Figure 6.39: Color-coded phase gradients (following Chiang et al. (2003)) of the CMB maps derived with  $\dots$ . On the top row we show the results from the full-sky analysis, in the middle the results from the 3 regions analysis and on the bottom row those from the 6 regions analysis. Dark blue along the diagonal or the border represent modes for which the given gradient is not defined.

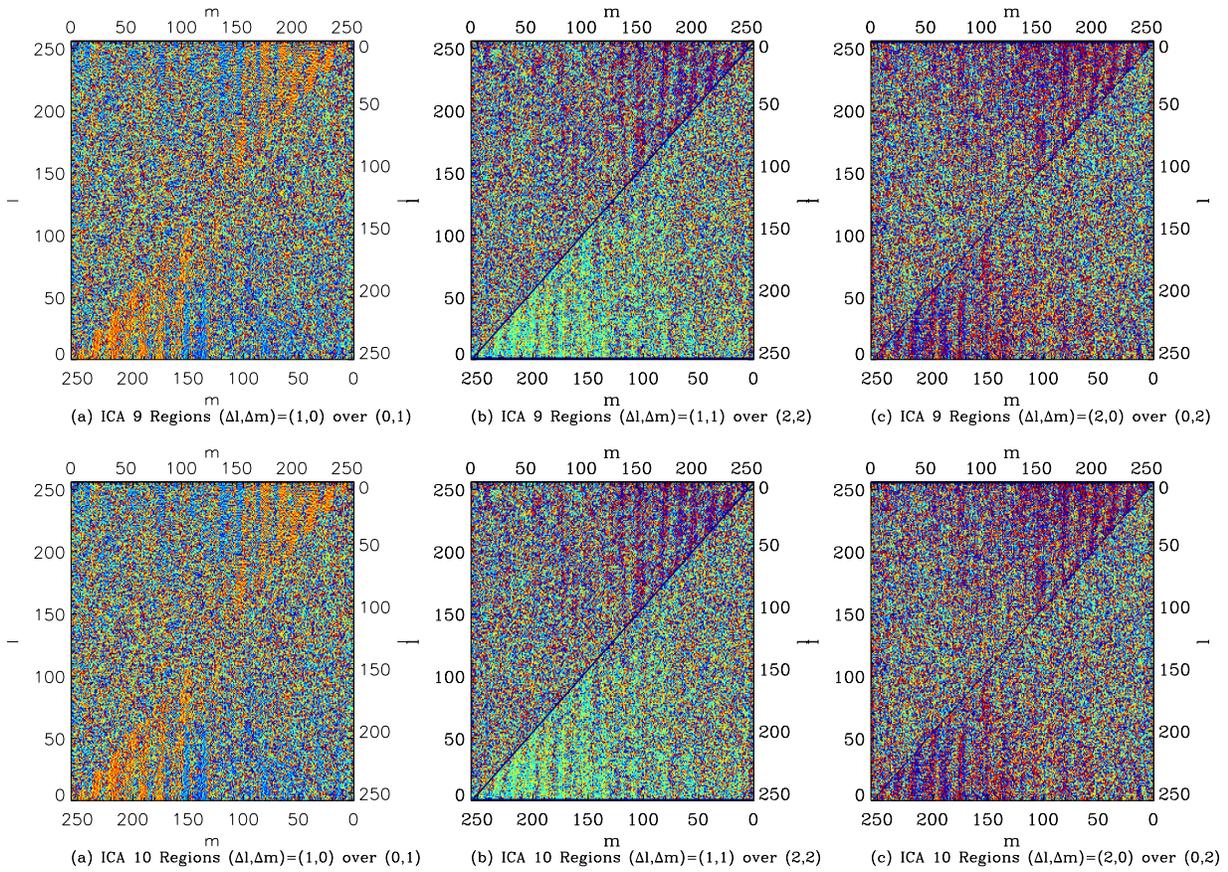


Figure 6.40: Color-coded phase gradients (following Chiang et al. (2003)) of the CMB maps derived with  $l_{max} = 250$ . On the top row we show the results from the 9 regions analysis and on the bottom row those from the 10 regions analysis. Dark blue along the diagonal or the border represent modes for which the given gradient is not defined.

we did for the CMB maps obtained from the real data analysis. The main point was to assess the significance of the real data results using simulations where the CMB maps have undergone the ILC or ICA analysis. Specifically, we have studied the distribution of the maximum value of the T- and D- statistics for  $\ell = 3$  and  $\ell = 6$  derived from the output, noisy and bias corrected CMB maps, with respect to the input ones (see Section 6.3 for the definition of the maps). These multipoles, indeed, are the most interesting ones in terms of anomalies detected using real data. As an example, the results of the three regions analysis with respect to the Ecliptic plane for both multipoles are shown in Figures 6.41 and 6.42. The real data results (green lines) have also been compared to these distributions.

The noisy and bias corrected maps generally have statistics values which are consistent with those derived from the input maps (i.e. CMB maps from Gaussian random field realisations). Besides, the output maps have more spread distributions, due to the effect of the residuals.

Furthermore, the T- and D- statistics values derived from the real data analysis for the two multipoles are always larger than the mean values of the simulations and lie on the tail of the distribution, remarkably so for  $\ell = 6$ . This is in agreement with the results of the real data analysis where the multipoles  $\ell = 3$  and  $\ell = 6$  generally return a planarity larger than 95% of the Gaussian simulations. Moreover, it indicates that the and ILC regional processing of the simulated data does not induce any spurious anomaly.

The effect of the residuals is particularly visible when inspecting the phase gradient plots: they generally have marked stripes, which get significantly weaker when the debiased maps are considered. The plots derived from the input and noisy maps, as expected, do not show any specific features.

Finally, simulations allow us to verify the hypothesis that the features of the phase gradients plots were actually due to residuals with a point sources like structure, rather than effect of real point sources. Indeed, by construction, the simulated data do not include point sources: therefore, the stripes visible in the plots can actually be associated with the residuals along the Galactic plane. Of course, this is with the caveat that the templates used to generate the diffuse emission are not still contaminated by sources - we are not sure to what extent this is true. There is also a noise contribution induced by the residuals, by means of the linear weights. In fact, we have carried out a statistical analysis of pure noise maps and we have tested that these maps give phase gradients affected by similar features, which are partially washed out by the CMB signal, when it is added.

## 6.5 ILC: a deeper study of the effect of partitioning the sky

Given the results concerning , we decided to concentrate on the ILC method to study the effect of the partition of the sky more in detail.

The simulations told us that the ILC performance improves with a regional approach. The real data analysis has confirmed this results and has also pointed out the importance of a subdivision of the sky which takes into account variations of the physical properties of the foreground emissions.

To strengthen this conclusion about the impact of non-Gaussianity depending on the kind of partition of the sky, we have carried out a systematic analysis of the CMB maps derived using a

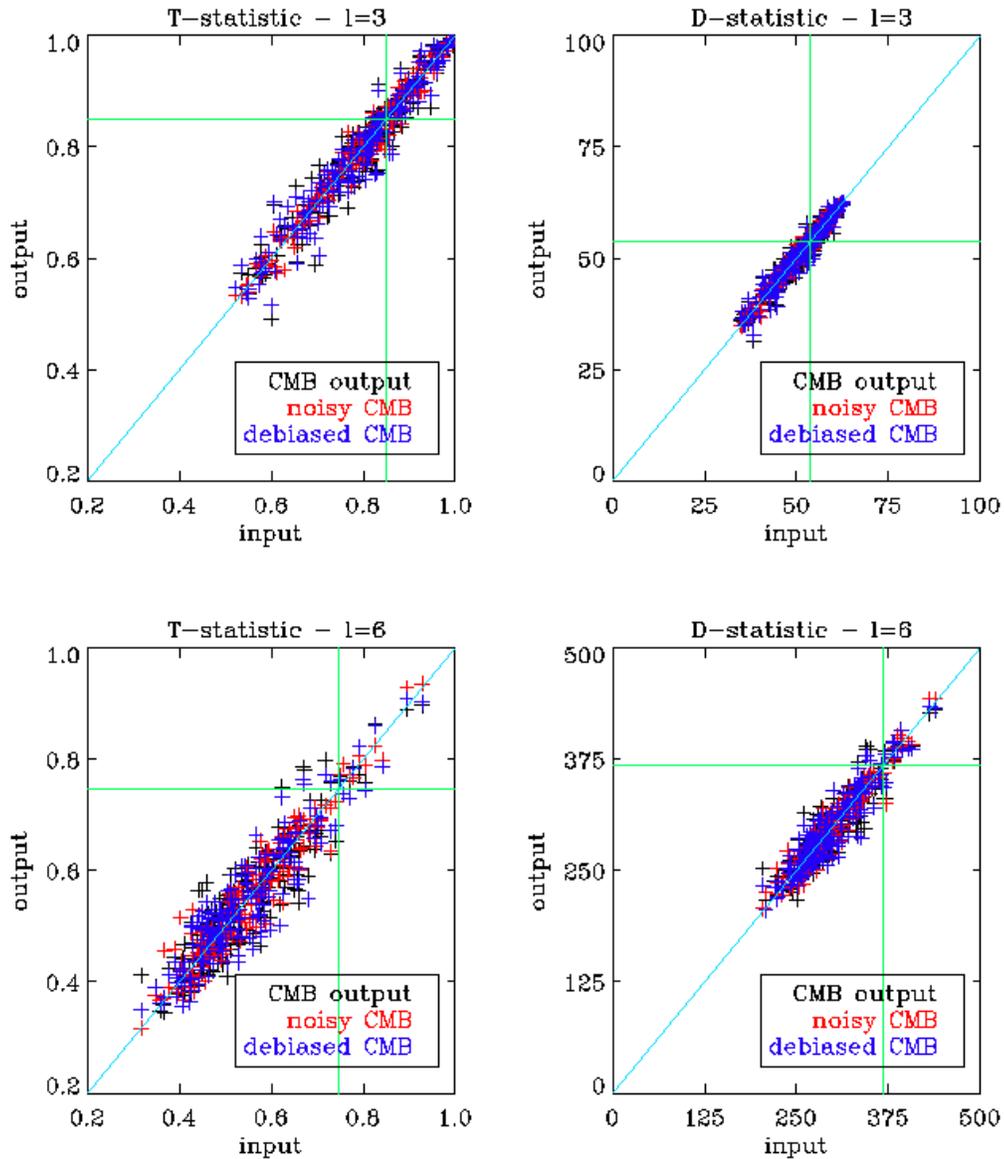


Figure 6.41: **three regions analysis:** scatter plots of the maximum values of the T- and D-statistics derived for the multipoles  $\ell = 3$  (left) and  $\ell = 6$  (right). The results are derived using the CMB maps obtained from simulations. The values from the output, noisy and bias-corrected maps are shown against the input ones. The green lines show the results derived from the real data analysis.

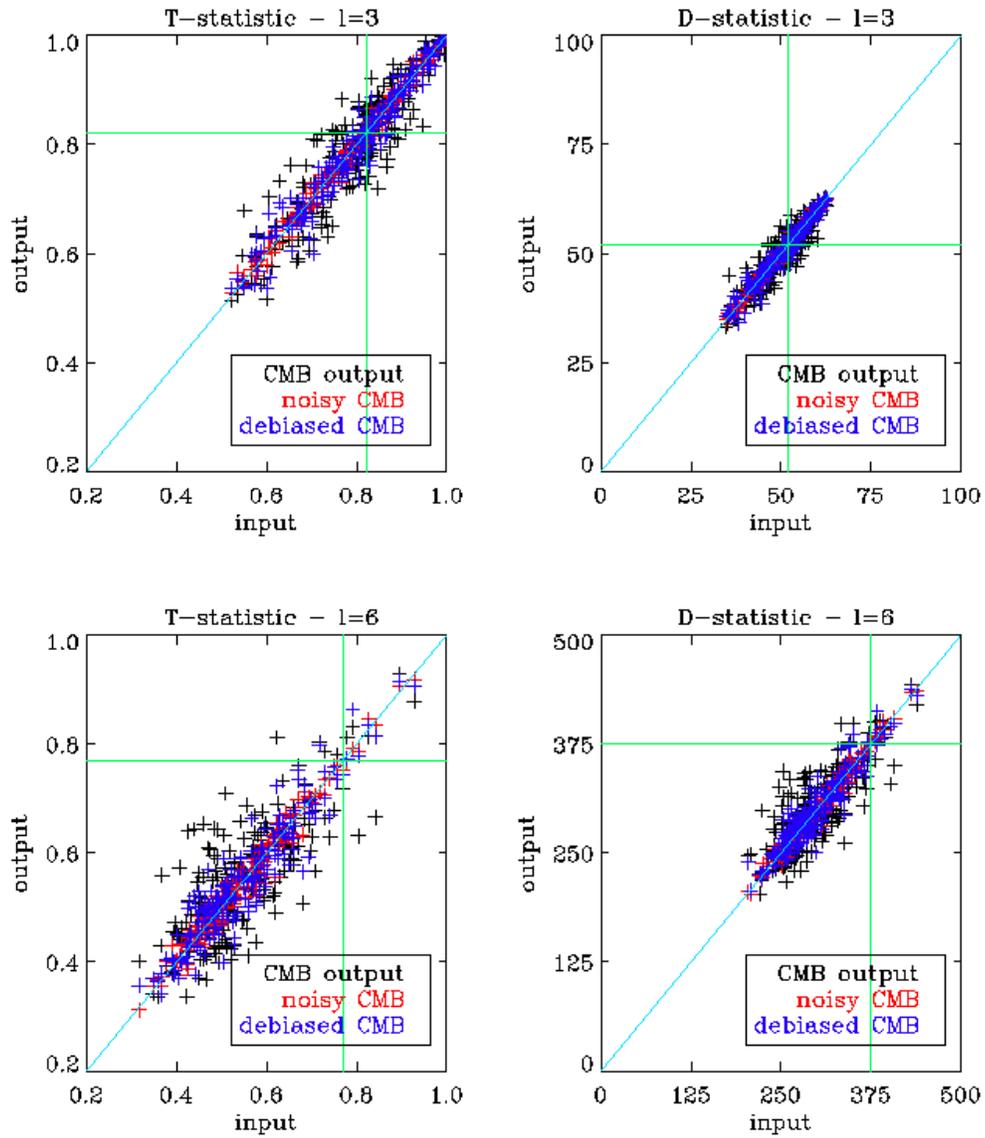


Figure 6.42: **ILC three regional analysis**: scatter plots of the maximum values of the T- and D-statistics derived for the multipoles  $\ell = 3$  (left) and  $\ell = 6$  (right). The results are derived using the CMB maps obtained from ILC simulations. The values from the output, noisy and bias-corrected maps are shown against the input ones. The green lines show the results derived from the real data analysis.

purely geometrical partition of the sky. We subdivided it in regions, which are defined as the pixels of an HEALPix map at the resolution of  $N_{side} = 1, 2, 4$  and  $8$ . In other words, we have defined new regions maps, with  $12, 48, 192$  and  $768$  regions respectively, the number of pixels being defined as  $N_{pix} = 12 N_{side}^2$ .

We applied the ILC analysis using the new partitions, and we computed the global variance of the merged CMB maps. In addition, we computed the variance over the pixels within each of the  $768$  regions defined with the map at  $N_{side} = 8$ . The same computation has been repeated using the cross-variance, to check again that the results were not actually biased by a poor estimation of the noise. As usual, in this case, the data are the fictitious ones (see Chapter 4).

The values of the variance and cross-variance are shown in Table 6.5:

	<i>variance</i>	<i>cross – variance</i>	<i>variance side A</i>	<i>variance side B</i>
<i>FS</i>	5390.41	5361.88	5494.43	5466.71
$N_{side} = 1$	4982.33	4972.39	5065.51	5035.30
$N_{side} = 2$	4843.32	4841.39	4917.80	4914.48
$N_{side} = 4$	4927.95	4836.83	4930.35	4908.85
$N_{side} = 8$	5106.37	4783.60	5033.96	4949.69

Table 6.5: Values in  $\mu K^2$  of the variance (first column) and cross-variance (second column) of the CMB maps derived from the regional analysis performed with the ILC code using the pixel based partitions of the sky. The values of the variance for the data used in the cross-variance computation are also included. Finally, for comparison, we also show the values of the variance of the CMB map derived from the full-sky analysis.

Looking at the cross-variance values, we have the expected trend: the larger the number of regions, the lower the cross-variance. However, this is not true for the variance itself. With  $192$  and  $768$  regions the value starts again to increase. This is most likely the consequence of the fact that the regions are very small and therefore, the noise becomes relatively important on small-scales. Indeed, the cross-variance does not show any anomaly, while the single fictitious maps used to compute it show the same behavior as the real data.

The larger value of the variance for small regions is also visible, looking at the CMB maps shown in Figure 6.43: the maps derived with  $768$  and  $192$  regions have some excess in power at high and low latitudes (the reader is warned that though visible, the effect is small), suggesting that the ILC solution is not optimal in those specific regions.

In Figure 6.44, we show the maps of the variance and cross-variance computed from the CMB maps obtained from the full-sky analysis, as well as from the analysis with  $12, 48$  and  $192$  regions. The variance (and cross-variance) is computed for each of the  $768$  regions, relatively to the value of the map generated from the analysis with the  $768$  regions themselves (the last one in Figure 6.43). Therefore, for example, a red color reflects a larger value of the variance with respect to the map computed using the  $768$  regions. These pixels are generally concentrated along the Galactic plane, where the analysis with a smaller subdivision of the sky leaves a more extended contamination from the foregrounds. The larger the number of the regions, the smaller the variance and the cross-variance, since the foreground residual decreases. However, for each case of analysis, there are regions where the variance (and cross-

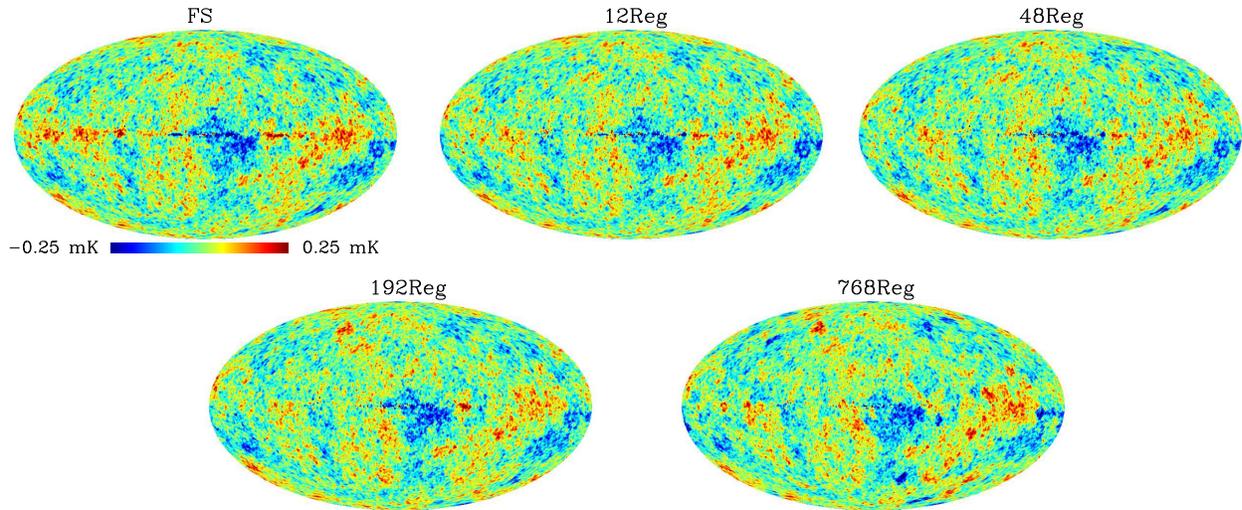


Figure 6.43: Plot of the CMB maps derived from the regional analysis performed with the ILC code, using 12, 48, 192, and 768 regions. For comparison we plotted also the CMB recovered from the full-sky analysis.

variance) is smaller than for the map derived with 768 regions. These are shown in the maps of the first column of Figure 6.44. They generally lie at high latitudes where the CMB pattern is normally well defined. This suggests the fact that a partition of the sky becomes more important for the regions which are closer to the Galactic plane. Indeed, this is what the *WMAP* science team has proposed and what our internal definition of the regions stands for.

The same results are shown in Figure 6.45: it presents the scatter plots for the variance and cross variance respectively. The values derived from the maps computed with a lower number of regions are plotted against the values derived from the map computed using 768 regions. The spread of the points decreases if the number of regions gets larger. Given a certain value of the variance for the map derived using 768 patches of the sky, the correspondent value of the maps derived with less regions is generally higher. It is the same for the cross-variance, meaning that the noise does not play an important role, while the variance seems to be properly derived.

In conclusion, this analysis confirms the importance of subdividing the sky to improve the CMB estimation. However, the comparison between the cross-variance value derived using 768 regions, with the one derived from the 400 regions analysis stresses the fact that a large partitioning of the sky also necessary requires it to be physically motivated. Indeed, the 400 regions analysis returns a CMB map which has a lower variance than the map derived with a simple geometrical subdivision.

## 6.6 Bias correction of the CMB estimation

Simulations implemented with the PSM showed that there is a significant discrepancy between the real observations and the mock ones, supporting the idea that the PSM is actually not optimal for describing the *WMAP* data. In fact, the direct comparison of the model with the real observations of the satellite

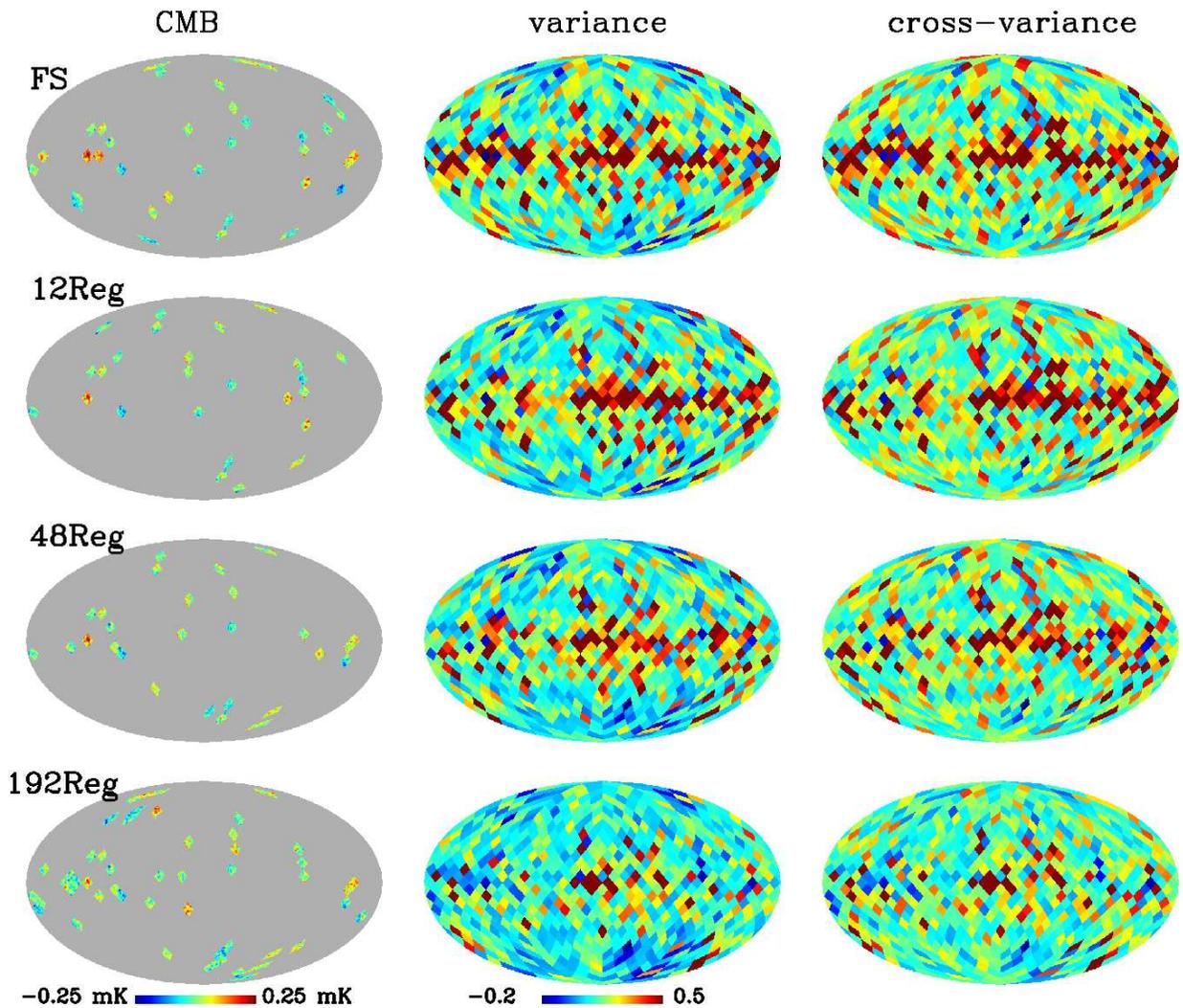


Figure 6.44: Plot of the maps of the variance and cross-variance computed from the maps obtained from the full-sky analysis as well as from the analysis with 12, 48 and 192 regions (already shown in Figure 6.43). The variance (and cross-variance) is computed for each of the 768 regions, relatively to the value of the map generated from the analysis of the 768 regions themselves. Therefore, a red color reflects a larger value of the variance with respect to the map computed using the largest number of regions. For each case of analysis, there are regions where the variance (and cross-variance) is smaller than for the map derived with 768 regions. These are shown in the maps of the first column of the figure.

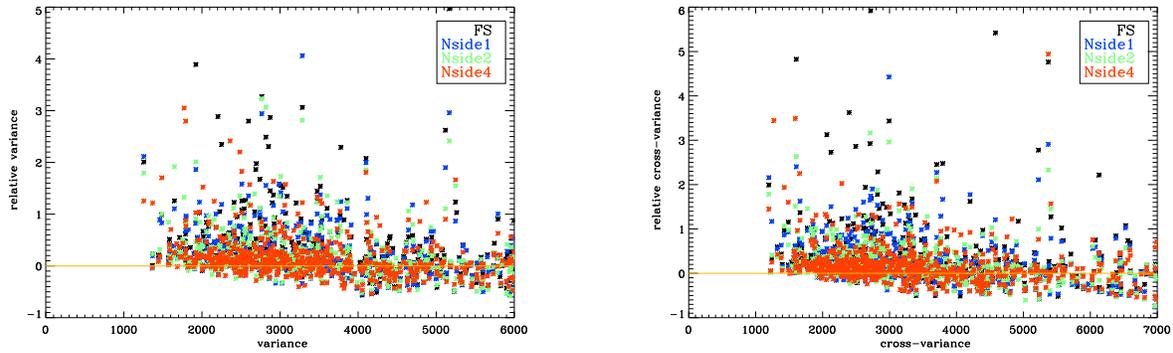


Figure 6.45: Scatter plots of the variance (left) and cross-variance (right) of the maps derived from the full-sky analysis (black points) as well as from the analysis with 12 (blue), 48 (green) and 192 (red) regions. The (cross-)variance is computed for each of the 768 regions, relatively to the value of the map generated from the analysis of the 768 regions themselves. The values derived from the maps computed with a lower number of regions are plotted against the values derived from the map computed using 768 regions.

(shown in Chapter 2), tells us that, although the PSM is still consistent with the *WMAP* data, the MEM solutions and the foreground maps proposed by Ghosh et al. (2010) are a better alternative as a tracer of the Galactic emissions.

The direct consequence of such a discrepancy is that the bias estimation derived with the PSM simulations is likely to be unreliable. Therefore, it is not advisable to use it for correcting the CMB maps returned either by  $\text{MEM}$  or ILC. It is better to adopt the bias correction computed from the simulations derived with the other two foreground models.

Ultimately, one faces the issue of whether it is better to choose a foreground model derived directly from the data after an ILC-like subtraction (such as the MEM and the GFM maps) and therefore affected by the corresponding foreground residuals, or to utilise a more independent model (e.g. the PSM), but risk that it is a poor representation of the true sky signal and introduce larger errors through an inadequate bias correction.

We practically demonstrate our statements applying the bias correction derived by means of simulations. Figures 6.46 and 6.47 compare the not corrected CMB maps derived from the regional analysis with  $\text{MEM}$  and ILC respectively, with those corrected from the bias as computed using the three foreground models available.

With  $\text{MEM}$ , the correction has been used for all the partitions of the sky, but with nine and ten regions: here, it was not possible to compute the bias due to the instability of the code. On the other hand, the six regions CMB map looks still quite contaminated regardless of the correction itself and of the model used. Again, the instability of the algorithm makes the bias estimation too poor to improve the quality of the CMB estimation.

Besides, the maps obtained from the full-sky and three regions analysis look much better when they are corrected for the bias, especially when it is computed with the MEM and GFM maps. The foreground contamination along the Galactic plane practically disappears. This is not the case with the

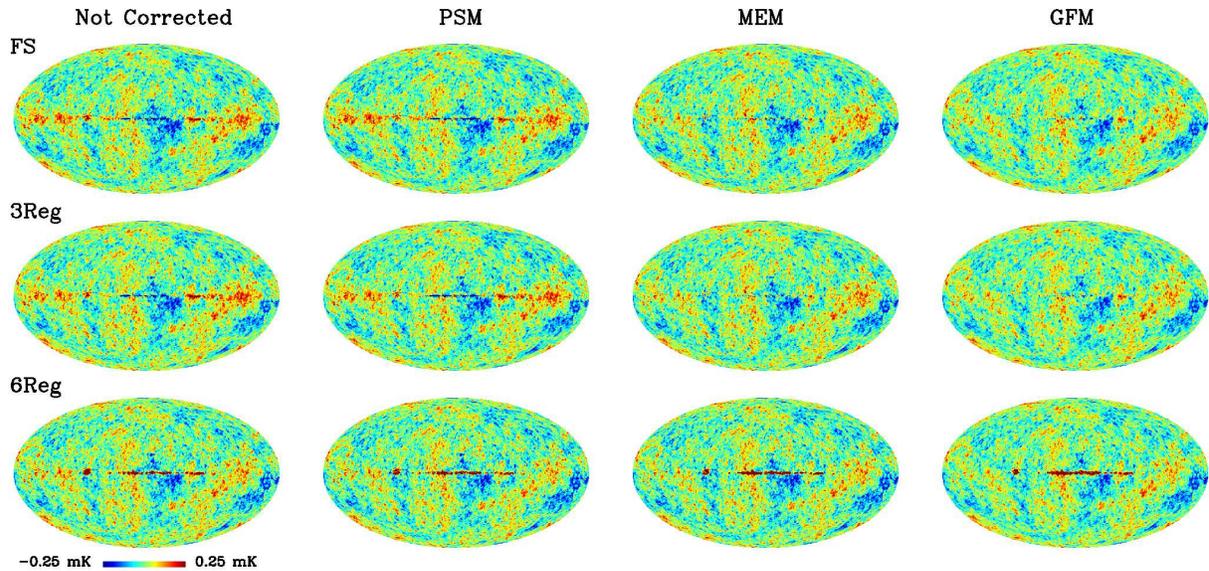


Figure 6.46: CMB maps derived with  $\text{MEM}$  and several partitions of the sky, after the bias has been removed. The latter has been computed by means of simulations with the PSM (second column), MEM maps (third column) and the GFM (fourth column). The non-corrected CMB maps are shown for comparison in the first column.

PSM: the CMB maps look even more contaminated than the not corrected ones, suggesting that the correction is not consistent with the data and therefore, not advisable.

The discrepancy between the PSM and the other two models is partially confirmed by the ILC results. Also in this case, indeed, the CMB maps corrected for the MEM and GFM bias do not show any contamination along the plane, while with the PSM correction, it is still present. Unlike the analysis, however, the residuals are visibly reduced for the six, nine and ten regions analysis, although the quality of the maps is still not optimal. This result is interesting and can be read as a hint in favour of the regional analysis.

### 6.6.1 Statistical properties of the bias corrected CMB maps

To give a more quantitative evaluation of the possible improvement due to the bias cleaning, we computed the variance of the corrected maps. Table 6.6 shows the values compared to those derived from the CMB maps without correction.

The qualitative assessments are confirmed: the variance of the ICA and ILC maps decreases with respect to the non-corrected maps when either the GFM or the MEM bias are used, while it increases with the PSM bias correction. The six regions map is an exception: only the GFM bias correction seems to work fine, although the variance value is significantly larger than those derived with the other partitions. Furthermore, it is interesting to note that the CMB maps derived from the six, nine and ten regions ILC analysis and cleaned with the PSM bias, have larger variance than the non-corrected maps, which is in contrast with what we could qualitatively judge from Figure 6.47: the former seemed apparently cleaner than the latter.

The dependence of the variance on the number of regions used in the analysis, follows the same

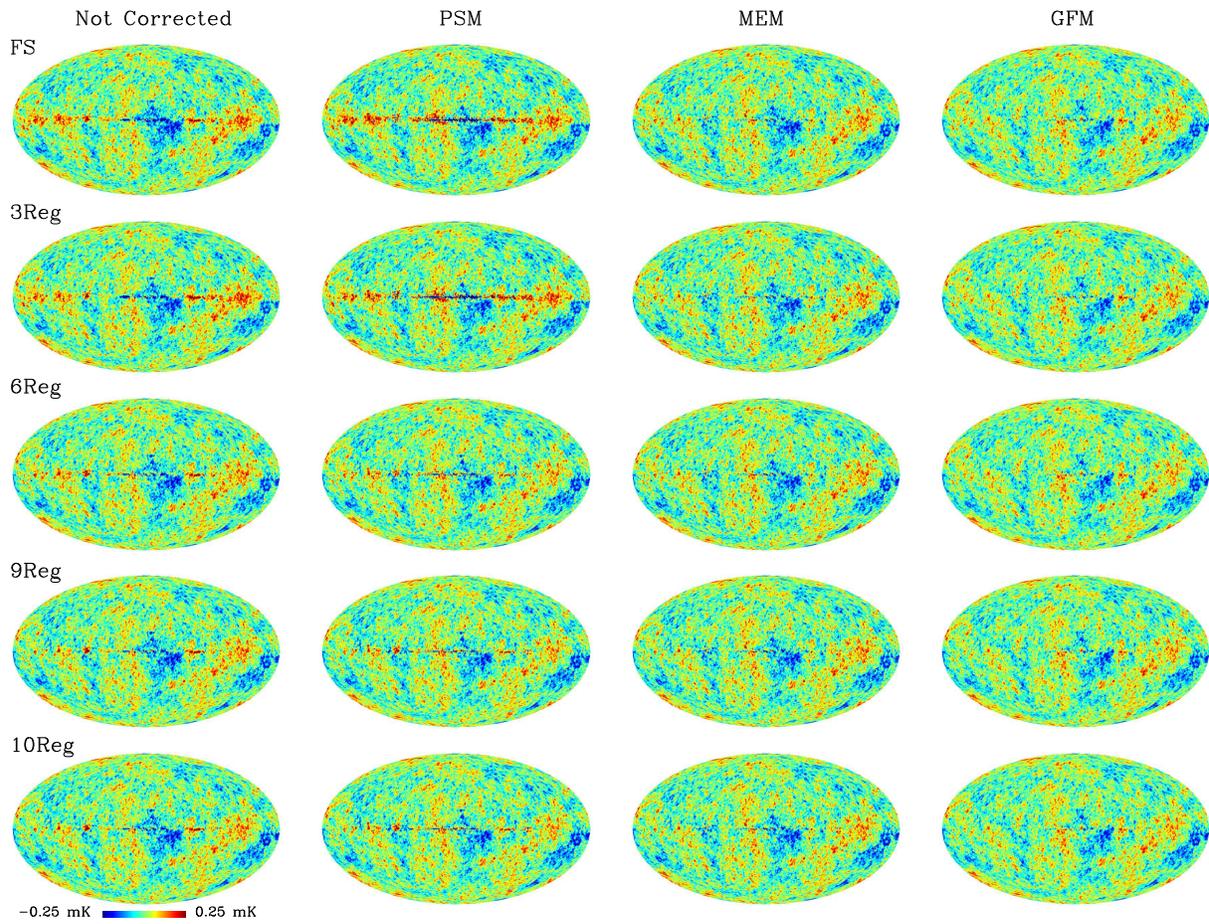


Figure 6.47: CMB maps derived with ILC and several partitions of the sky, after the bias has been removed. The latter has been computed by means of simulations with the PSM (second column), MEM maps (third column) and the GFM (fourth column). The non-corrected CMB maps are shown for comparison in the first column.

	No Correction	PSM	MEM	GFM
<i>FS</i>	5431.97	5520.34	4957.36	4550.14
<i>3Reg</i>	5335.31	5417.43	4918.87	4540.30
<i>6Reg</i>	6893.71	7101.39	8341.54	6370.29
<i>9Reg</i>	6154.47	–	–	–
<i>10Reg</i>	6160.00	–	–	–
	ILC			
	No Correction	PSM	MEM	GFM
<i>FS</i>	5390.41	6258.62	4924.48	4563.41
<i>3Reg</i>	5296.31	6232.33	4926.31	4537.68
<i>6Reg</i>	4990.35	5011.08	4906.86	4527.25
<i>9Reg</i>	5009.29	5066.70	4936.15	4531.85
<i>10Reg</i>	5021.44	5078.36	4948.17	4558.80

Table 6.6: Values of the variance (in  $\mu K^2$ ) of the merged CMB maps derived from the regional analysis performed with  $\text{ICA}$  and the ILC code. The maps are now corrected for the bias due to foreground residuals, evaluated according to the simulations. Different foreground models have been used: the PSM, the model derived merging the MEM maps and the maps produced by Ghosh et al. (2010).

trend than for the non-corrected maps (commented in Section 6.4.2). Among all the cases of analysis, the best estimation of the CMB is given by the six regions ILC analysis, applying the GFM bias correction. The last one seems to provide the best estimation of the bias, and therefore, to be the best model to be used.

Finally, as done in Section 6.3.2, we computed the local skewness and kurtosis of the bias cleaned ICA and ILC CMB estimations, depending on the correction applied. This allows us to complete the picture and to give us a direct way of judging the presence of residuals. The statistics maps are shown in Figures 6.48, 6.49 and 6.50. We present only the results derived with a pixel resolution of  $N_{side} = 8$ , although we checked that the residuals do not impact significantly the statistics maps with a higher resolution ( $N_{side} = 32$ ).

As expected, when the PSM foreground model is used to estimate the bias, both the skewness and kurtosis maps are affected by the residuals induced by the wrong bias estimation. Moreover, the skewness shows negative peaks along the Galactic plane, reflecting the over-subtraction of the foreground residuals: the contamination still present in the cleaned CMB maps indeed, have a negative sign.

When the MEM maps are used as tracers of the Galactic components, the skewness maps look uniform while the kurtosis ones show large value in the Galactic center. Only exception is the six regions analysis: in this case, both the statistics have large values along the Galactic plane, reflecting the properties of the map from which they are computed.

The situation improves remarkably when the GFM is used, demonstrating again the goodness of the model for the *WMAP* observations. The skewness and kurtosis maps are indeed uniform. Only few pixels randomly distributed show high statistic values. However, it is interesting to note that they do

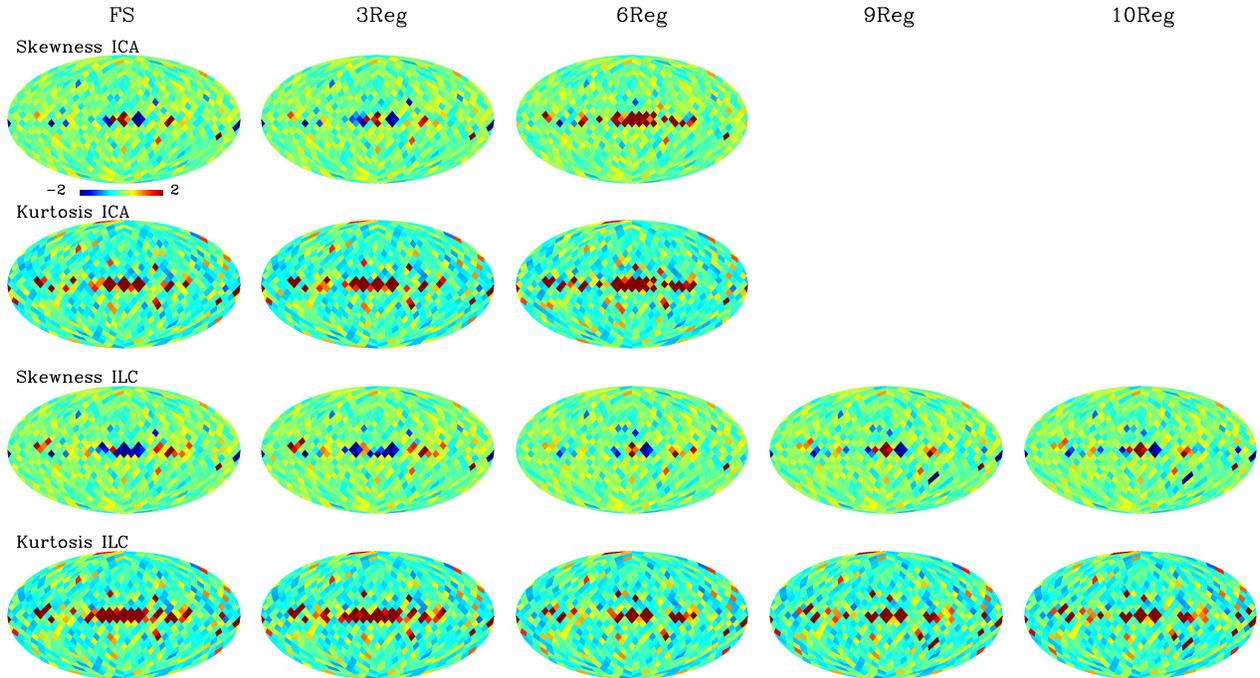


Figure 6.48: **PSM bias correction.** Maps of the local skewness and kurtosis computed from the bias cleaned CMB maps, derived with both the `ICA` and `ILC` codes, and for the different partitions of the sky. The bias correction is evaluated with simulations generated with the PSM. The maps have a resolution of  $N_{side} = 8$ .

not necessarily lie along the Galactic plane. Again, the six regions `6Reg` map returns contaminated statistics maps, as expected.

## 6.7 Discussion

In this chapter, we have presented a regional analysis performed on the *WMAP* five-year data, using both the `ICA` and `ILC` codes. The work aimed at studying the improvement of the performance of the algorithms depending on the partition of the sky, meaning the number of patches used, as well as the criterion to select them. Moreover, different statistics have been proposed as a figure of merit to evaluate the CMB estimation and the impact of possible non-Gaussianities.

We adopted different sets of regions, defined following both a priori criteria and knowledge derived from the full-sky analysis of the *WMAP* data with `ICA`.

We first studied the applicability of this kind of analysis with simulations. This was specifically required for `ICA` which has never been used on small regions of the sky. The `ILC` code, instead, has been originally implemented with a regional approach: therefore, simulations were useful to test the effect of different partitions of the sky with respect to the one proposed by the *WMAP* science team, as well as to compare the two methods.

For both the algorithm, the simulations were also necessary to estimate the bias introduced by

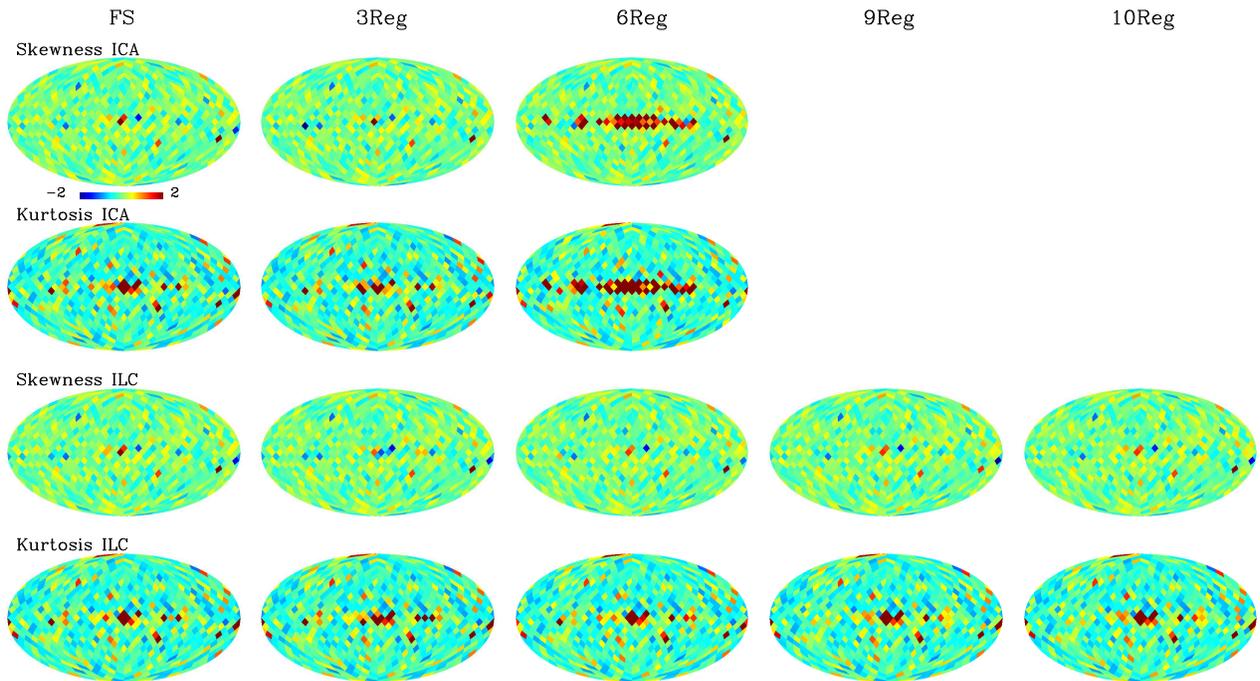


Figure 6.49: **MEM bias correction.** Maps of the local skewness and kurtosis computed from the bias cleaned CMB maps, derived with both the  $\text{MEM}$  and ILC coded, and for the different partitions of the sky. The bias correction is evaluated with simulations generated with the MEM maps. The maps have a resolution of  $N_{side} = 8$ .

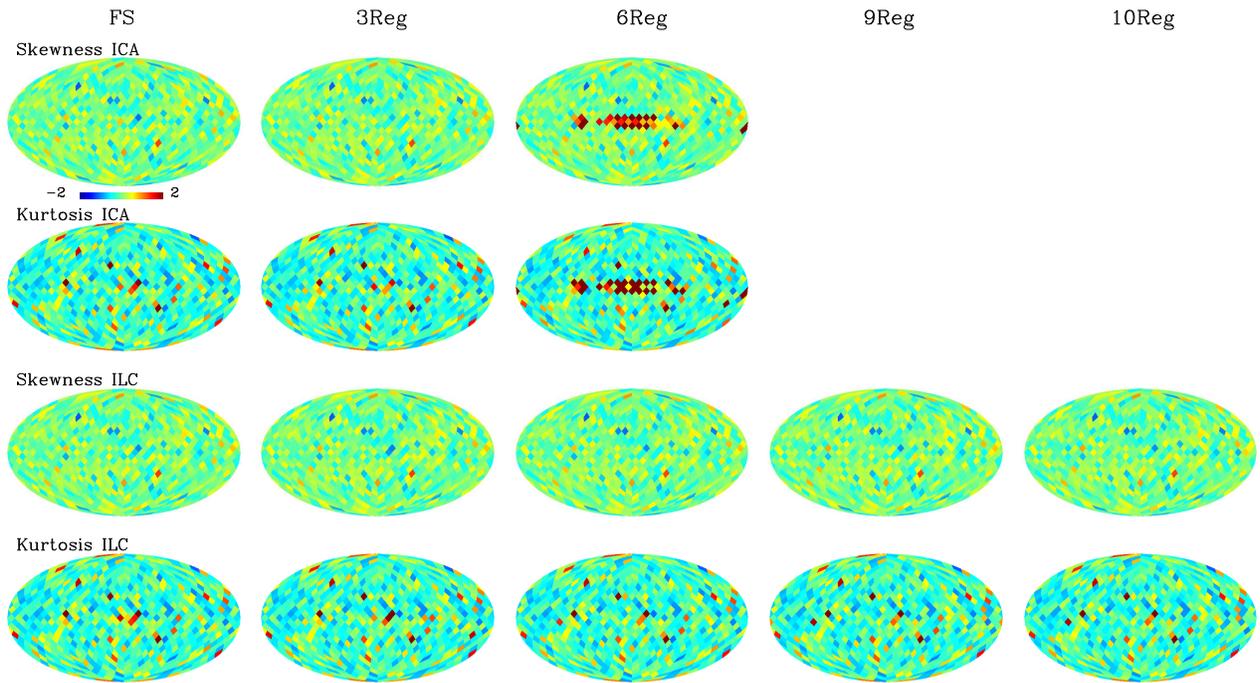


Figure 6.50: **GFM bias correction.** Maps of the local skewness and kurtosis computed from the bias cleaned CMB maps, derived with both the  $\text{ICA}$  and ILC coded, and for the different partitions of the sky. The bias correction is evaluated with simulations generated with the GFM. The maps have a resolution of  $N_{side} = 8$ .

foreground residuals in the CMB reconstruction, not perfectly subtracted by the internal combination of the data. The extent of this bias gives us an idea of the correction which, should be applied to the ILC or ICA CMB maps. However, this is just a rough estimation which does not necessarily match the real one, due to the possible discrepancy between real and simulated data. For this reason, we considered three different foreground models, as described by the PSM, the MEM maps and the maps produced by Ghosh et al. (2010). For all of them, the bias correction turned out to be properly computed and to improve the CMB estimation, when applied to the simulations results. However, this is not always the case for the CMB maps derived from the real data analysis. Specifically, the PSM gives a poor description of the *WMAP* data and, therefore, a wrong estimation of the bias of the CMB computed from real observations. Besides, we have shown that the MEM and GFM maps provide a better description of the data and therefore of the actual bias in the CMB estimation. This result supports the idea of preferring a foreground model directly derived from real observations, in order to have a reliable assessment of residuals in the CMB. On the other hand, there are some risks due to the fact that the method used to produce the model requires the ILC subtraction from the data, which are consequently biased.

The statistical properties of the recovered CMB maps seems to be in favor of a poor response of to the regional analysis, unlike the ILC one which improves. The lack of the convergence in some simulations, make this statement even stronger, specifically when a number of regions larger than six is adopted. Though, some instabilities are also connected to the geometrical extent of the patches of the sky, itself.

The same analysis has been repeated using the *WMAP* 5-year data and selected partitions of the sky among those adopted for the simulations. The reconstructed CMB maps are evaluated in terms of statistical properties, such as the power spectrum estimation, the variance, skewness and kurtosis computation, as well as phases analysis. The local evaluation of the kurtosis and the phase analysis of the CMB seem to be the best approaches to detect non-Gaussianities, such as residuals and artifacts induced by the analysis itself. Furthermore, the phases analysis can shed light on the morphology of the residuals.

All the results derived from simulations have been confirmed and strengthened, in particular the weak performance of compared to ILC. Therefore, the ICA approach is better advisable for a full-sky analysis, rather than on small patches of the sky, to have a good estimation of the CMB.

Given the different response of the codes to a regional analysis, we have dedicated particular attention to the ILC method, in order to study in more detail the dependence of the statistical properties of the CMB estimation on the number of regions. A systematic analysis shows that, as expected, the larger the number of regions, the better the CMB estimation. The latter is first qualitatively evaluated in terms of residuals visible in the CMB maps: these decrease in extension when increasing the number of regions used for the analysis. In addition, a quantitative evaluation is made studying the statistical properties of the same maps. Specifically, the variance of the CMB emission gets lower when enlarging the number of subdivisions of the sky. Furthermore, a local estimation of the skewness and kurtosis shows that, for angular scales larger than  $2^\circ$ , the lower the contaminants, the lower the deviations from Gaussianity. These basically reflect the spurious emission along the Galactic plane, which is not

significant and does not affect the CMB estimation on angular scales smaller than  $2^\circ$ .

Comparing the values of the variance derived from this systematic study with those derived from the other set of regions, it is possible to note that the trend with respect to the number of regions is not always respected. The ILC code is sensitive to the way the sky is subdivided: a physical motivated partition is better than a simple geometrical one.

All these results are interesting in order to find the best conditions to apply both the component separation methods, and derive the cleanest CMB maps. The regional analysis is a key improvement to CMB data analysis, considering a possible application to polarised data. Indeed, given the strongest contamination from foregrounds and the lack of knowledge concerning their properties in polarization, a better performance in the component separation is mandatory.



# Chapter 7

## Conclusions

The work of this thesis is a contribution to the problems of cleaning CMB observations of diffuse Galactic emission, and of shedding light on the nature and properties of such contaminants. Any advance in our knowledge of the physics of the foreground emissions gives a chance of enhancing our ability to cleaning CMB data. This is vital in order to use these observations for cosmological purposes.

The analysis has been carried out by means of two techniques, namely  $\text{MCMC}$  and ILC. Their applicability and potential have also been investigated and compared in a systematic way, for the first time. This study has also allowed us to find the conditions that guarantee the greatest improvement in the estimation of the CMB emission, which is a significant step forward in understanding how to properly select component separation methods according to the context.

Interesting aspects related to the nature of the foreground emissions have been studied in great detail in this thesis, also providing independent confirmation of results already pointed out by previous works.

Nevertheless several questions remain open for further investigation, both in the direction of discovering the existence and the nature of unknown components, as well as of improving the techniques employed for such a study. In the following section, we would like to guide the reader through the main results of this thesis, highlighting the issues that are still unsolved, in the prospect of future observations.

### 7.1 Unsolved issues and prospects

#### 7.1.1 Is the anomalous dust component actually due to spinning dust?

The analysis described in Chapters 3 and 4 has detected the existence of a dust correlated emission component at frequencies lower than 60 GHz, the nature of which is still under debate. The most widely accepted interpretation, also supported by our results, advocates that this emission is related to PAH (Polycyclic aromatic hydrocarbon) spinning dust grains (Draine & Lazarian, 1998a): these grains undergo collisions with gas-phase species, thereby becoming charged and acquiring a significant permanent dipole moment. As a result of the frequent collisions, the spinning PAHs will produce

electric dipole emission from the decay of collisionally excited rotational levels.

However, there are other plausible scenarios that can not currently be ruled out, such as the hypothesis of emission from Very Large Grains (VLG) composed of fullerenes (see the full dust spectrum shown in Figure 1.11 in Chapter 1), or magnetic dipole emission from large grains subject to spontaneous magnetisation. Furthermore, within the frame of emission by spinning dust grains, recent studies have suggested that this anomalous emission could be produced by *Very Small Grains* (VSG). For this case, studying the emission from HII regions are decisive: since very small grains are not expected to be there, any detection of anomalous emission from those regions would exclude the hypothesis of VSGs and favour other interpretations. An extensive study of the anomalous emission in HII regions has been undertaken by Dickinson et al. (2006, 2007) who first observed the LPH96 region (G201.66+1.64) with the Cosmic Background Imager (CBI) at 31 GHz, and later the regions in the southern hemisphere. More recently, Dickinson et al. (2009) have also observed RCW175 at 31 GHz. In all these cases, they measured an excess of flux density which has been interpreted as the evidence for electric dipole emission from small spinning dust grains, rather than VSGs.

To better characterise the frequency spectrum of the anomalous component, and therefore to confirm or otherwise the different scenarios, a large set of new observations would certainly be useful. If the anomalous emission actually arises from spinning dust grains, we expect it to be correlated with the thermal emission from the same grains (although not perfectly since a range of grain sizes contributes). Ysard et al. (2010) recently found evidence of a correlation between the anomalous dust and 100  $\mu\text{m}$  IRAS observations, which becomes even stronger at 12  $\mu\text{m}$ . This result strongly supports the spinning dust hypothesis: in fact, current models predict that the spinning dust emission is dominated by small dust grains (PAHs) radiating at 12  $\mu\text{m}$ . The difficult removal of the zodiacal light that contaminates this emission is the major challenge for this kind of study, which nevertheless seems to be a promising approach for determining the nature of the anomalous component. The same correlation could be further exploited using high- and low-frequency Planck data. Extending from 30 to 900 GHz, the maps could be combined to better isolate the anomalous emission by removing thermal dust and free-free radiation.

However, Planck data alone will not be enough. In fact, the best way to establish the possible connection between the anomalous component and the spinning dust emission is to better characterise the spectrum of the latter. According to the models proposed by Draine & Lazarian (1998a), the spinning dust emission would have a peak roughly around 20 GHz, depending on the conditions in the environment. Therefore, observations at frequencies lower than 20 GHz would be useful to detect the turn-over in the spectrum and the peak itself.

Thus, the C-BASS project will be a good complement to the Planck observations: performing measurements at 5 GHz, it will provide a map of the sky both in intensity and polarisation, that will be useful as a new template for the synchrotron emission. In conjunction with full-sky surveys at other frequencies, it will then be possible to better disentangle the synchrotron contamination from the free-free and the anomalous dust emission. In addition, the QUIJOTE (Q U I JOint TEnerife) CMB experiment will provide one of the most sensitive 11 – 19 GHz measurements of both the synchrotron and the anomalous emissions on degree angular scales, although only on a limited region of the sky.

Such studies, however, will be challenging since they are complicated by several issues of both theoretical and practical natures. For example, to isolate the anomalous emission, we need to separate the synchrotron and free-free components, which is difficult due to the limited amount of observations currently available. Furthermore, the theoretical spinning dust models are not definitive: for example, recent refinements to the theory favour a shift of the peak emissivity by a factor between 2 and 4, and an increase of the peak frequency by factors from 1.4 to 1.8, depending on the component considered (Hoang et al., 2010). Predictions depend on our knowledge of the ISM and the dust grains that are currently studied in more detail using the Herschel, Spitzer and IRAS observations, complemented by a wide range of ancillary data (such as HI and CO emission): local variations of the dust spectrum need to be related to the gas physical conditions. The work proposed by Miville-Deschênes et al. (2001), Abergel et al. (2010) and Kirk et al. (2010) goes in this direction.

On the practical side, in Chapters 3 and 4, we have observed that a simple power law provides the best-fit model for the spectrum of the anomalous dust component, derived from our coefficients. Therefore, the hypothesis that the detected emissivity is the result of a mix of contributions from spinning dust grains in different environments is difficult to prove, since a single contribution is not identifiable. A spectroscopic analysis over a continuum range of frequencies would probably be the best solution to detect the detailed features of the spectrum.

### 7.1.2 Do we really understand the free-free emission and its spectrum?

Although the free-free emission is the foreground that seems to be better understood and characterised, our analysis showed that some aspects of it remain unclear.

#### Why is the electron temperature derived from the free-free to $H\alpha$ ratio lower than expected?

In Chapters 3 and 4, we noticed an inconsistency in the scaling between the  $H\alpha$  template and free-free emission at K and Ka bands. A value of 8000 K is conventionally adopted for the thermal electron temperature (see Section 1.4.2), but in fact, the K- and Ka-band results for the  $Kp2$ ,  $KQ85$  and  $KQ75$  sky coverage are in better agreement with temperatures in the range of 5000-6000 K. Also, the coefficients at higher frequencies are progressively consistent with a higher temperature value. The resultant spectrum, increasing with the frequency, is quite different from the theoretical expectations and we believe it might be interpreted as due to the contribution of another component spatially correlated with free-free emission. We did not exclude the possibility that it is an alternative signature of the bump claimed by Dobler et al. (2009) in the free-free spectrum, associated with spinning dust emission from the Warm Ionised Medium (WIM). Since other sources of emission could also be equally possible explanations, Gold et al. (2009) proposed an analytic form for the spinning dust spectrum to be included in the foreground model used to fit the *WMAP* data - this can actually represent a variety of physical sources of microwave emissions. It will be interesting in the future to fit the same model to our spectrum.

The coupling coefficients derived when a  $Kp0$  mask is applied are significantly different, having an amplitude that corresponds to an electron temperature below 4000 K. This latter behaviour was

also observed by Davies et al. (2006) for five regions specifically selected to be dominated by free-free emission. This is why we interpreted this result as possibly related to genuine temperature variations of the ionized gas at medium latitudes, due to a mixture of regions with different properties and therefore electron temperatures. However, other explanations are equally possible, such as problems with the emissivity conversion or with the  $H\alpha$  template.

### Does the bump in the free-free emission truly exist?

The anomalous spectral behaviour of the  $H\alpha$  correlation coefficients, has prompted us to pay particular attention to this question. Further motivation arose from the fact that Dobler et al. (2009) claimed to have detected a bump in the free-free spectrum, suggesting that it could be associated with correlated emission produced by spinning dust grains in the WIM.

We investigated that claim by repeating the foreground analysis presented in Chapters 3 and 4, following the approach suggested by Dobler et al. (2009): the *WMAP* data have been cleaned from the CMB emission, before applying (see Chapter 5).

Our simulation analysis seems to favour the presence of this feature in the free-free spectrum and, therefore, the existence of the  $H\alpha$ -correlated WIM emission. Moreover, it ruled out the hypothesis of the bump being a spurious feature induced by the ILC subtraction (as seen in Chapter 2, the ILC estimation is always affected by residuals due to the cross-correlation between the cosmological signal and the foreground components).

The same analysis applied to the *WMAP* observations yielded quite different results, together with a significant instability of in recovering the foreground components. Such a discrepancy between the real and simulated data analysis is partially explained by the adopted foreground model, which does not realistically trace the observed structures of the sky. A more complex picture of the Galactic emission, due to a strong mix among the components, possibly increases the difficulty in disentangling the components. It may produce a larger bias in the ICA and ILC CMB estimations, which therefore affects the computation of the coupling coefficients and the spectra of the single foregrounds.

These results highlight the necessity of new observations of the Interstellar Medium, to develop a clearer picture of the spinning dust emission and to more efficiently separate the different components at low frequencies. Thanks to the multi-frequency Planck observations the free-free emission will be more easily separated from other diffuse Galactic emission components (synchrotron and anomalous dust emission). In addition, the new generation of  $H\alpha$  surveys (e.g. WHAM South (Haffner et al., 2010)), will provide important new data relating to the ionised gas in the Galaxy. When compared to the dust distribution, they will allow us to estimate the dust absorption correction for the  $H\alpha$  emission, and therefore, the electron temperature of the ionised gas in various environments. This is necessary to derive the corresponding free-free brightness temperature. Along the Galactic plane, where the dust absorption makes it impossible to use the  $H\alpha$  emission, the electron temperature will be obtained by means of the RRLs from HII regions. Resulting from electrons and ions recombining in an ionised gas with subsequent cascade transitions to a ground state, the line emission can be used to derive

the electron temperature of the ISM and the correspondent free-free brightness temperature. This is very interesting in the context of future radio observations provided by experiments such as SKA and LOFAR.

New hints about the physical properties of the microwave sky will also come from the already quoted QUIJOTE experiment as well as QUIET (Q/U Imaging Experiment): focusing on specific regions of the sky, the spatial variation of the foreground properties will be possibly investigated in more detail.

### 7.1.3 What is the nature of the *WMAP* Haze?

In Chapters 3 and 4, we have used  $\text{fit}_{\text{synch}}$  to internally analyse the *WMAP* data, cleaned of the foreground emissions according to the coupling coefficients derived from a template fit with the same algorithm. This iterative analysis has shown the existence of a radiation component concentrated around the Galactic center and that is not correlated with any of the templates used to trace the Galactic components. A comparison between the analysis implemented using the Haslam map as synchrotron template, and that carried out using the K-Ka map has demonstrated that this emission is mainly connected to low frequencies observations: when using the second synchrotron template, the residual radiation appears to be lower in amplitude, a consequence of it being ended in the map itself.

Our results confirm those already found by Finkbeiner (2004a), who first detected this emission and named it as *WMAP* Haze. While further studies seem to have established the synchrotron nature of this emission, its origin is still under debate. Several scenarios have been suggested, including the hypothesis that it might be synchrotron emission from the by-products of dark matter annihilation.

The idea of dealing with new physics is certainly intriguing and has driven a deeper investigation of the process of dark matter annihilation. This has involved the study of other experiments. A preliminary analysis (Dobler et al., 2009) of the data produced by the Fermi Space Telescope has shown the existence of a  $\gamma$ -ray “haze” with a similar spatial morphology of the *WMAP* Haze. This would be generated by the same hard population of electrons responsible for the microwave signal, by means of inverse Compton interaction with the starlight. Like the *WMAP* Haze, the  $\gamma$ -ray haze could also be explained in terms of dark matter annihilation (Cholis et al., 2009).

On the other hand, the reliability of the models used to trace the different components of the sky is a crucial point for the detection of the haze. This apparent emission might be simply due to the lack of a reliable template for the synchrotron emission. Indeed, the Haslam map at 408 MHz, although widely used in the literature, is not representative of the actual contribution of synchrotron emission at the observing frequencies of *WMAP*, due to the frequency variations of the spectral index: using such a crude extrapolation of the morphology of synchrotron emission to frequencies larger than 20 GHz, can potentially introduce unphysical residuals (Mertsch & Sarkar, 2010). Indeed, Cumberbatch et al. (2009) have noted that the significance of the *WMAP* Haze is substantially reduced by allowing spatial variations in the frequency dependence of the synchrotron emission in the inner and outer parts of the Galaxy. In this sense, the K-Ka map is a better choice for a template: indeed, using this template, we have noticed a significant decrement in the amplitude of this residual, although some signatures of it

still remain.

A more recent study of the Fermi data proposed by Su et al. (2010) has demonstrated the existence of two large gamma-ray bubbles, extending 50 degrees above and below the Galactic center, with a width of about 40 degrees in longitude. The hard spectrum emission associated with these bubbles seems to be spatially correlated with the *WMAP* Haze, and more importantly, to correspond to features in the ROSAT X-ray maps at 1.5-2 keV. This result would actually confirm the Haze as a new emission and would associate it with a major episode of energy injection in the Galactic center, such as past accretion events onto the central massive black hole, or a nuclear starburst in the last  $\sim 10$  Myr. The hypothesis of dark matter annihilation or decay would consequently be ruled out.

As in the case of the free-free bump, further investigations with new observations are necessary to shed light on the problem. Fermi will continue to provide fundamental data enriching our knowledge of the ISM. C-BASS will complement the new Planck data as well as already existing observations, and help to fully define the spatial and frequency properties of the synchrotron emission.

#### 7.1.4 Will observations in polarisation help us to improve the component separation ?

Not all foreground components are expected to be polarised and not all at the same level. The two dominant components of diffuse polarised foreground emission in the 23-94 GHz range are synchrotron emission and thermal dust emission (see Page et al. (2007) and references therein). The free-free emission is unpolarised (unless it is scattered), and the spinning dust grains are expected to have polarisation fractions of 1% – 2% (Lazarian & Draine, 2000).

The different response to the magnetic field of our Galaxy implies different properties of the polarised foreground emissions. These are useful signatures for disentangling the different components in the sky. In fact, the relative contamination of the polarised CMB by polarised foregrounds is much stronger than in total intensity. Therefore, it becomes mandatory to characterize the properties of the Galactic emission components, before using CMB polarisation observations for cosmological purposes.

Observations in polarisation provide significant additional information to that obtained by studying total intensity alone. Several ground-based observatories are planning surveys to study polarised emission in our Galaxy (e.g. LOFAR and SKA). Planck will certainly improve the situation, regarding full-sky polarisation observations. Furthermore, there are several planned future experiments, which are custom designed to measure polarization. Apart from those already mentioned above, PILOT has particular relevance since it will provide critical information about the polarised dust emission: it will constrain the large scale geometry of the magnetic field in our Galaxy and allow us to study in detail the alignment properties of dust grains with respect to the magnetic field. In this field, the measurements of PILOT will be complementary to those of Planck at longer wavelengths and lower angular resolution.

#### 7.1.5 What is the optimal use of $\chi^2$ and ILC, and how could they be improved?

The region-based analysis implemented in Chapter 6 gave us the opportunity to compare and improve two different methods of component separation, namely  $\chi^2$  and ILC. Such a comparison has been carried out in a systematic way for the first time. In that analysis, we performed a component separation

of the *WMAP* 5yr-data by treating regions of the sky separately, taking into account the spatial variation of the properties of the Galactic radiation. While the first method turned out to be quite unstable if used on small patches of the sky, the second one improves with a regional approach, particularly in the case where the regions are defined according to the real variations of the sky.

Both of them, however, return CMB estimations biased by foreground residuals whose presence is due to the chance cross-correlation between the CMB and the foreground contaminants. Simulations, allowed us to estimate the size of such a bias, with the requirement of using a foreground model that is consistent with the real data. Among those available, the MEM maps as well as those proposed by Ghosh et al. (2010) seem to be good tracers of the foreground emissions in the *WMAP* data. The Planck Sky Model, on the contrary, is quite different from the real sky as seen by *WMAP*. The new Planck observations will allow a direct comparison with such a model, providing hints for future refinements.

One of the main issues of our regional analysis is how to best choose the partitioning of the sky in order to trace the actual spatial variations of the foregrounds. Apart from a simple geometrical approach to the problem (e.g. a simple parallel cut along the Galactic plane, see Chapter 6 for more details), we have suggested a new partition of the sky, the interesting and original aspect of it being that it is defined based on the foreground separation itself.

This result is not completely clearly related to the physical processes in the ISM, possibly reducing the efficiency of the regional analysis. In future, this subdivision can be refined if a better characterisation of the foreground mechanisms is achieved. In addition to studying the properties of the free-free emission as already mentioned, we also need to better characterise both the dust and synchrotron emissions. The Herschel experiment has already provided a large data set, suitable for defining the temperature and spectral index relation of the dust emission, although constrained to specific regions of the sky (see for example Kirk et al. (2010), Abergel et al. (2010) and Anderson et al. (2010)). The spatial variation of the spectral index of the synchrotron emission is another fundamental criterion to be explored when determining how to partition the sky. It might be achieved by using polarised observations such as those produced by Planck and C-BASS.

The versatility of also makes it suitable for foreground analysis by means of templates. However, the work performed in Chapters 3 and 4 has demonstrated that the best use of this algorithm is its application for the internal analysis of data pre-cleaned of foreground contamination. Indeed, such a combination, which is unique in the field of the component separation methods, provides us with a strong tool to discover potentially new foreground components. Although in our case we implemented both steps with the same method, the estimation of the foreground contamination could actually be better performed by using other component separation techniques. This combination is an interesting aspect to be investigated in future work, to achieve a better estimation of both the CMB and foreground emission.

A natural continuation of the regional study would be the extension of this kind of approach to the foreground analysis performed with . As in the case of the internal analysis, however, we have shown that this is not possible due to the instability of the algorithm when applied on a small sample of pixels. In this case, the instability becomes even stronger due to the mix among the foreground components.

A solution to overcome the limitations of the algorithm could be found by developing the approach we have adopted for studying the *WMAP* Haze. Although the Haze is synchrotron emission and should be included in the correspondent template, we described it with an independent map that has been fitted to the data together with the other templates. Thus we have derived the relative coefficients, whose frequency behaviour compared to that computed for the synchrotron emission, describes the spatial variation of the synchrotron radiation around the Galactic centre. In a similar way, we can describe each foreground component with a set of maps, each of them focused on a specific feature of the emission component considered. New multi-frequency observations, in addition to providing the necessary data set, might provide new hints on how to separate the features of the foreground components.

## Appendix A

# Monte Carlo simulations

*Note: the work presented in this appendix was done in collaboration with A.J. Banday and D.Maino and was published as part of Bottino et al. (2008). All the computations have been performed by M. Bottino.*

We have performed 1000 simulations of the microwave sky at each of the 5 *WMAP* frequencies, each containing a realisation of the CMB signal, the Galactic foreground emission, and instrumental noise appropriate to the specific channel. The study was undertaken at an effective resolution of  $1^\circ$ . Each CMB component corresponds to a Gaussian realisation of the theoretical best fit CMB angular power spectrum with no running spectral index as determined by the *WMAP* first-year analysis<sup>1</sup>. As verified by Maino et al. (2007) up to  $\ell \sim 400$  – the scales of interest for  $1^\circ$  smoothed data – the differences between the 3 year and 1 year best-fit models are less than 1.5% and therefore not expected to bias the results significantly. The white noise component is simulated for each of the 10 DAs in the usual manner. That is, for a given pixel a Gaussian random number is selected with zero mean and a variance corresponding to the ratio of the noise variance per observation for that channel and the number of observations for the pixel in question. The CMB component and noise realisations are then combined, and the maps smoothed. For those frequencies with multiple associated DAs, the band averaged sky maps are formed using simple averaging. For the foreground emission, we added to each simulated frequency map three Galactic templates as described in section 3.3 of chapter 1, appropriately scaled to the frequency in question. These maps were then used as input data for

As a convenient point of reference for our results, we have also included a simple template fitting analysis. In general, the correlation method can be extended to include various constraints on the data, e.g. fixed dust or free-free spectral indices. This is the case described by Hinshaw et al. (2007) in the analysis of the 3 year data of *WMAP* but we impose no such constraints here. The best-fit monopole is removed from each of the sky maps before computing the coefficients, to be consistent with the approach.

Using these simulations, we can determine the suitability of the non-quadratic functions adopted in the analysis for template fitting. Figure A.1 summarises the situation when using the Haslam

---

<sup>1</sup>[http://lambda.gsfc.nasa.gov/data/map/powspec/wmap\\_lcdm\\_pl\\_model\\_yr1\\_v1.txt](http://lambda.gsfc.nasa.gov/data/map/powspec/wmap_lcdm_pl_model_yr1_v1.txt)

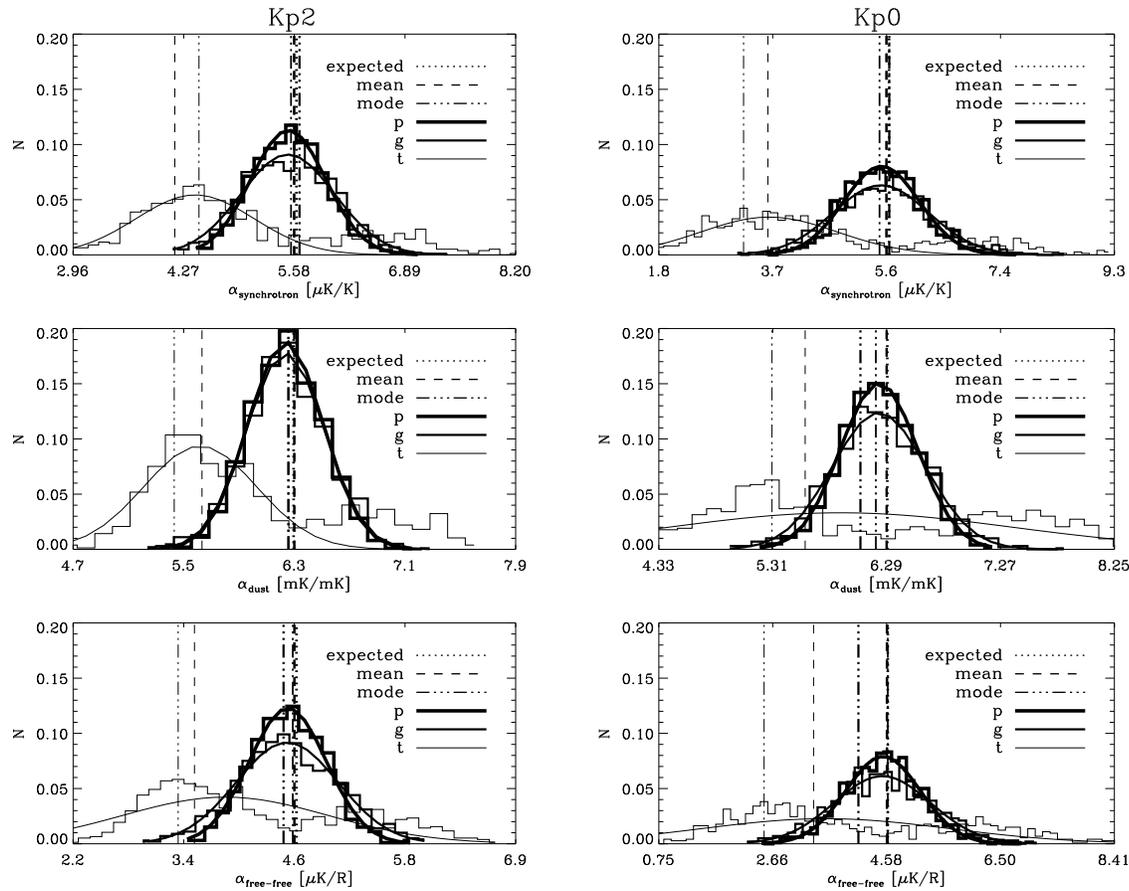


Figure A.1: Coupling coefficients distribution determined from simulations of *WMAP* K-band data with the  $p$ -,  $g$ - and  $t$ -functions. The mean (long-dashes and dashed line), the mode (dash-dot and dash-dot-dot line) and the input values (dotted line) are also shown, together with the best-fit Gaussian profiles.

408 MHz sky map as a tracer of the synchrotron emission. The templates were scaled using the weights from Table 3 of Bennett et al. (2003), where for the Q-, V- and W- bands we have considered the average values.

The distribution of coefficients determined by  $t$  when utilising the  $t$  function is clearly unsatisfactory – the distribution is not Gaussian and has broad asymmetric tails, and is strongly biased away from the input values. In the case of the Kp0 mask, the distribution also shows evidence of bimodality, and for many simulations the  $t$  analysis does not even converge. We conclude that, when using the  $t$  function  $t$  is unable to distinguish Galactic emission in the *WMAP* data as described by template maps, and we do not include results based on this non-quadratic function for the analysis of the real data.

In contrast, the behaviour of the  $p$  and  $g$  functions for determining the coupling coefficients seems to be suitable for our purposes. Although we show results only for the K-band simulations, we find that at all frequencies the scaling factors are very well described by Gaussian distributions – in fact the corresponding measures of the skewness and kurtosis values are small in all cases. The mean values

for the simulated distributions of coefficients are very similar to the inputs, and there is no evidence of significant bias. This is true also for the mode if we consider  $p$ , but with  $g$  the mode is slightly smaller than the input value for the dust and free-free emission: indeed, it could explain partially the difference between the coefficients obtained in the real analysis. However, the main difference between the two functions is that the distribution determined with  $g$  is broader than with the  $p$  function: corresponding to a larger statistical uncertainty for analyses made with the former.

It is likely that some of the performance features of the different non-linear functions are related to the statistical features of the physical emission as traced by the template sky maps. Not surprisingly, the results indicate that the template fits using  $g$  are sensitive to the extent of the mask applied to the data. For a narrower Galactic exclusion region, the template fits have smaller errors. This is certainly in part due to simple ‘sample variance’ type arguments, but may also be connected with the degree to which bright foreground emission near to the applied cut can stabilise the fits. Variations in the fitted amplitudes of the foreground components against the observed data as a function of mask may also reveal genuine changes in the physical properties and spectral behaviour of the foregrounds with latitude.

In Figure A.2 we have plotted the fitted Q-band amplitudes for one template against another for the various template permutations in order to study the significance of cross-talk between the maps. There is an apparent anti-correlation between the synchrotron and the dust coefficients (as specified by the linear correlation coefficient  $r$ ), but no clear evidence for such behaviour between synchrotron and free-free or free-free and dust. This behaviour is essentially identical for all of the *WMAP* frequencies. Some care should then be exercised in interpreting the coefficients for the former two components.

We have repeated the simulations and analysis using the *WMAP* K-Ka map as the synchrotron template, in this case only for the Q-, V- and W-bands. The templates were scaled using the coefficients from Hinshaw et al. (2007), but adopting the mean of the coefficients over the differencing assemblies. The distributions of the coefficients are again symmetric and well fitted by Gaussians as in the case of the Haslam template. Furthermore, the results are consistent with the previous ones concerning the dependence of the coefficients on the  $g$  functions. The analogous scatter plots between pairs of coefficients are shown in Figure A.3. This time, the synchrotron coefficients are anti-correlated with the free-free results. Indeed, this should be expected since the template must contain both synchrotron and free-free emission. However, the synchrotron coefficients are more strongly anti-correlated with the dust values, which is easily interpreted as being due to the presence of anomalous dust correlated emission in the K-Ka template. This factor is not taken into account by Hinshaw et al. (2007) in their constrained template analysis.

Finally, we performed the same statistical study with the simple  $\chi^2$  minimisation in order to compare the performances of the two methods. The results depend on the template considered. For the synchrotron and free-free emission, the uncertainties are larger than the  $g$  values obtained with either the  $p$ - or  $g$ -function, substantially so in the former case. For the dust emission they are essentially equivalent. It is therefore clear that  $g$  is a reliable method to perform template fitting, in some cases outperforming the conventional  $\chi^2$  minimisation. This is particularly true for  $p$ , and we could conclude that it is the most stable non-linear function, although this is not a definitive argument.

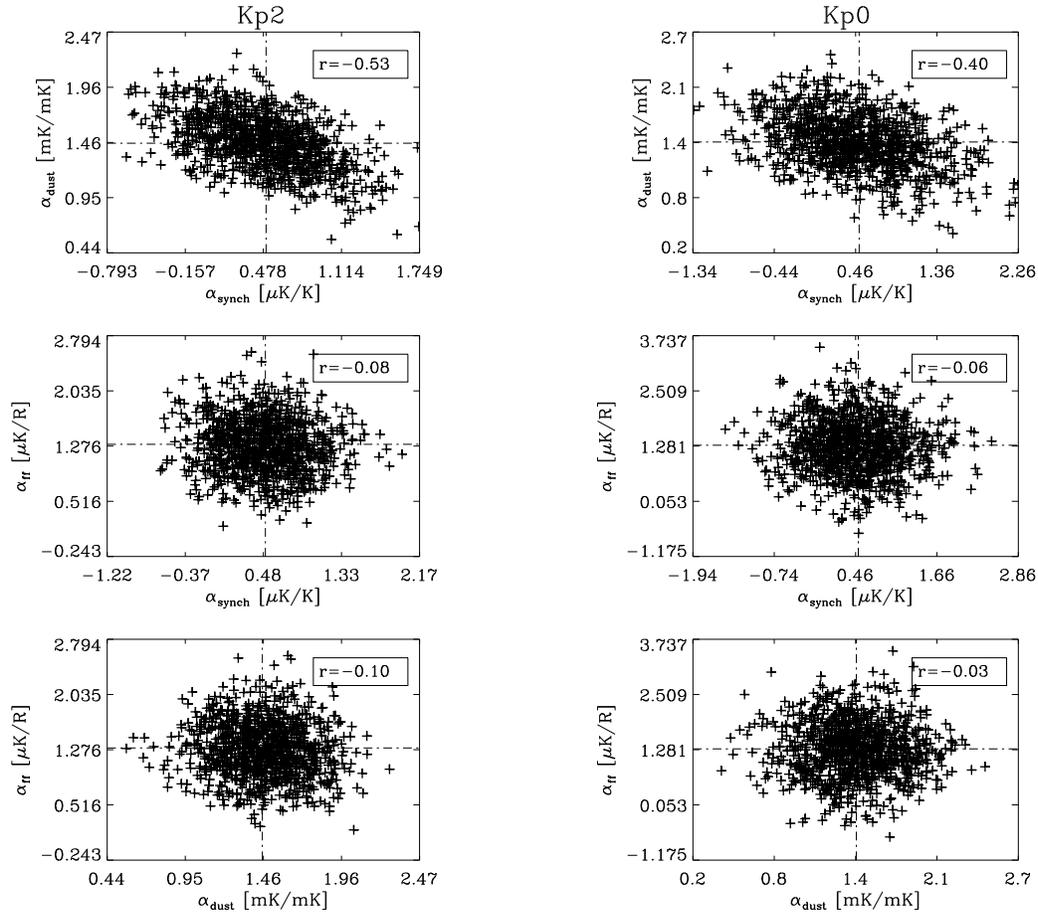


Figure A.2: Scatter plots of the coupling coefficients obtained by  $\chi^2$  with the  $p$ -function using the simulations performed using the Haslam map as synchrotron template. We show as an example the results in the Q-band. There is very little change in the correlation properties as a function of frequency.

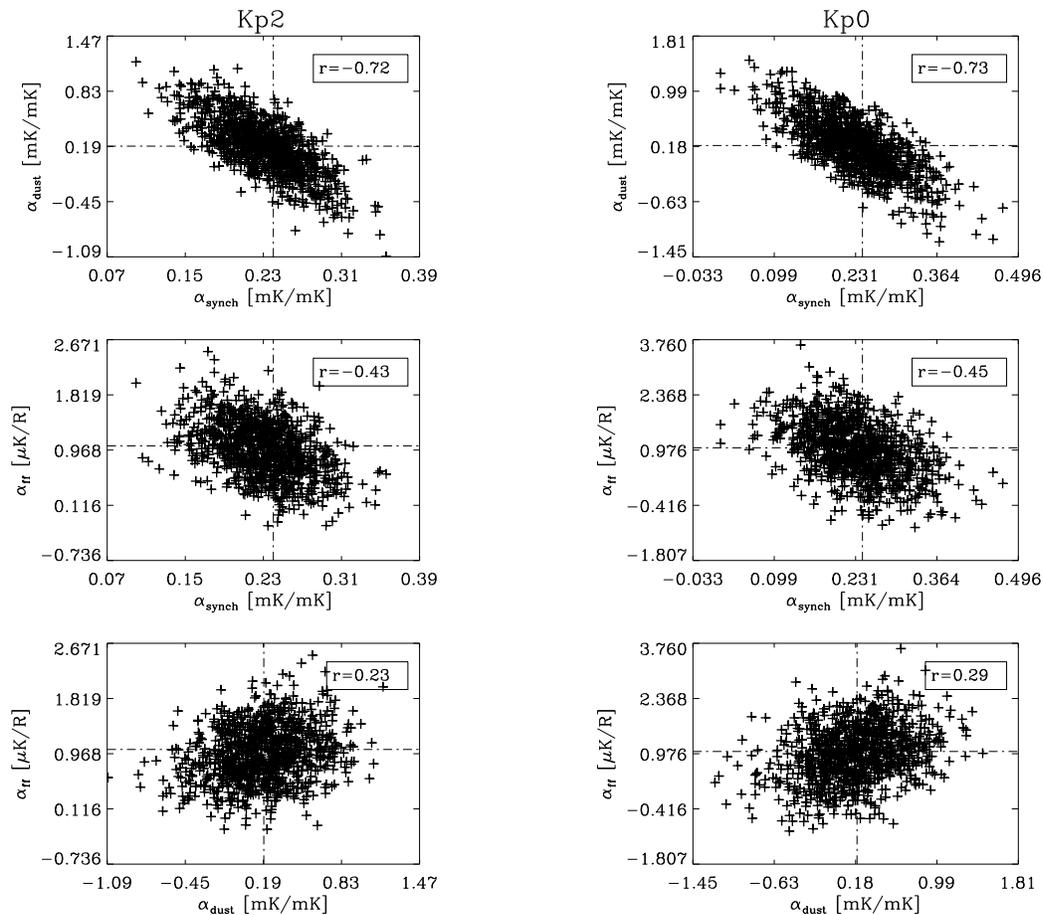


Figure A.3: Scatter plots of the coupling coefficients obtained by  $\dots$  with the  $p$ -function using the simulations performed with the K-Ka map as synchrotron template. We show as an example the results in the Q-band. There is very little change in the correlation properties as a function of frequency.

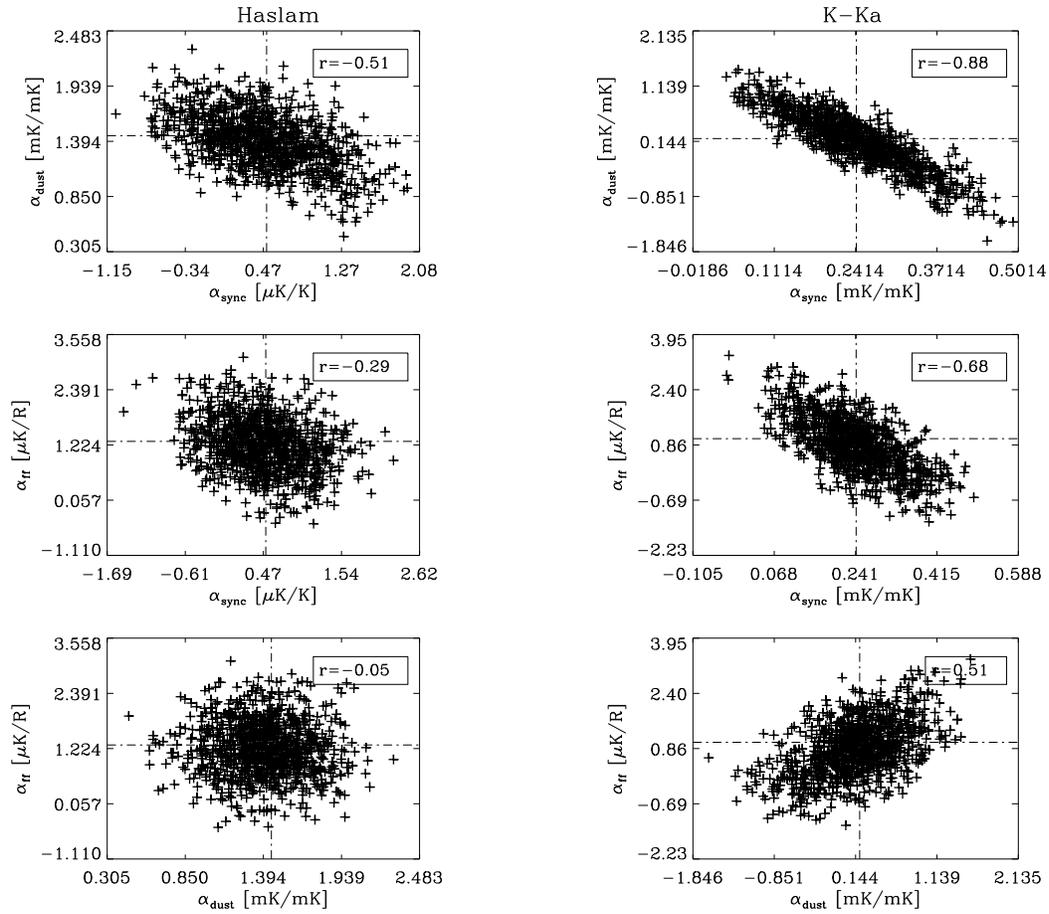


Figure A.4: Scatter plots of the scaling factors obtained via the simple  $\chi^2$  method on  $Kp2$  sky coverage with simulations performed using either the Haslam map (left column) or the K-Ka map (right column) as the synchrotron template. We show as an example the results in the Q-band. There is very little change in the correlation properties as a function of frequency.

---

Figure A.4 shows scatter plots between pairs of recovered Q-band coefficients when using either the Haslam or K-Ka maps as synchrotron templates. It is interesting to note that there is a hint of anti-correlation between the synchrotron and free-free coefficients using the Haslam template, but that this becomes significant when the K-Ka template is employed. Moreover, a significant correlation between free-free and dust is seen in the latter case. The simple  $\chi^2$  method appears to demonstrate more cross-talk between the coefficients of the different components than seen for the  $\chi^2$  analysis. This may help to understand some of the differences seen in the results obtained with the observed data.



# Acknowledgements

First of all, I would like to thank Tony Banday for being my supervisor, even from far away, for having made a lot of efforts trying to teach me something and for having ‘standed’ my English skills.

I would also like to thank Davide Maino for the collaboration in the work of this thesis and generally, for the unconditional support since I started my degree thesis. Together with him, I would also like to mention Prof. Marco Bersanelli and his research group in Milan, with whom this adventure has started (in particular Benedetta and Simona).

I am grateful to Prof. Dr. Simon White for useful suggestions during this work and for having always attentively followed my progress.

The accomplishment of this result would not have been possible without the support of all my friends. In different ways, they have all accompanied me and always suggested the right approach to the difficulties. I am sure I will forget someone, although I will try my best not to do it!

Starting from those who spent more time with me during this PhD, I would like to thank my office mates that have made my daily life at the institute pleasant and fun: in particular, Luca (also because of his help in programming and fixing all my computer related issues, but overall because he made me laugh a lot), and Margot (for having shared with me chat and...chocolate!). Always at MPA, I would like to thank André and Mona (among several reasons, also because they have been my greatest scientific support), Martin Reinecke (for his help with compiling issues), Stuart (also for having corrected part of my thesis), Caroline and Silvia, but also not strictly MPA people such as Giulia&Daria (my super lunch mates, with whom I have shared joys and difficulties) and the members of the ‘Thursday lunch group’ (Giovanni, Giulia, Daria, Luca, Marco, Christoph, Margherita, Alessio and Alessandro). Outside MPA, my greatest source of support was Anna, Ilaria&Alberto (and Tommasino!), the Tardini family, but also the Cristoforetti family, Luca&Cri, Nina, Laura, Marghe, Lele&Eli, Chicco, Luisa&Dario, and still all my friends from Milan or Genoa, and now spread all over the world, but not less close to me. The list would be very long (probably too much!) to write all the names, but I have to mention at least Marta&Ben, Laura and her wonderful family, Guido, Martí, Dani, the Curto family, Franceschina, Anna, Czesto, Marta L., Paola and Ilaria.

Finally, I would like to specially thank all the members of my family (old and new!), because without them I would have never reached this point. In particular, I am grateful to my husband, to whom I have dedicated this thesis, since I do not find other words to say thanks!

Paola



# Bibliography

- Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, *Physical Review Letters*, 104, 101101
- Abergel, A., Arab, H., Compiègne, M., et al. 2010, *A&A*, 518, L96+
- Ali-Haïmoud, Y., Hirata, C. M., & Dickinson, C. 2009, *MNRAS*, 395, 1055
- Alpher, R. A. & Herman, R. C. 1948, *Physical Review*, 74, 1737
- Anderson, L. D., Zavagno, A., Rodón, J. A., et al. 2010, *A&A*, 518, L99+
- Baldi, P., Kerkyacharian, G., Marinucci, D., & Picard, D. 2007, *ArXiv e-prints*
- Baldi, P., Kerkyacharian, G., Marinucci, D., & Picard, D. 2009, *Annals of Statistics*, 37, 1150
- Banday, A. J., Dickinson, C., Davies, R. D., Davis, R. J., & Górski, K. M. 2003, *MNRAS*, 345, 897
- Barkana, R. & Loeb, A. 2001, *Phys. Rep.*, 349, 125
- Barkats, D., Bischoff, C., Farese, P., et al. 2005, *ApJL*, 619, L127
- Baumann, D., Jackson, M. G., Adshead, P., et al. 2009, in *American Institute of Physics Conference Series*, Vol. 1141, American Institute of Physics Conference Series, ed. S. Dodelson, D. Baumann, A. Cooray, J. Dunkley, A. Fraisse, M. G. Jackson, A. Kogut, L. Krauss, M. Zaldarriaga, & K. Smith , 10–120
- Bennett, C. L., Banday, A. J., Gorski, K. M., et al. 1996, *ApJL*, 464, L1+
- Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2010, *ArXiv e-prints*
- Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2003, *ApJS*, 148, 97
- Bernardeau, F. 1998, *A&A*, 338, 375
- Bernui, A. 2008, *Phys. Rev. D*, 78, 063531
- Bielewicz, P., Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2005, *ApJ*, 635, 750
- Blanchard, A. & Schneider, J. 1987, *A&A*, 184, 1
- Bonaldi, A., Ricciardi, S., Leach, S., et al. 2007, *MNRAS*, 382, 1791
- Bond, J. R. 1996, in *Cosmology and Large Scale Structure*, ed. R. Schaeffer, J. Silk, M. Spiro, & J. Zinn-Justin, 469–+
- Born, M. & Wolf, E. 1980, *Principles of Optics Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, ed. Born, M. & Wolf, E.

- Bottino, M., Banday, A. J., & Maino, D. 2008, MNRAS, 389, 1190
- Bottino, M., Banday, A. J., & Maino, D. 2010, MNRAS, 402, 207
- Cabella, P. & Kamionkowski, M. 2004, ArXiv Astrophysics e-prints
- Cardoso, J. & Souloumiac, A. 1993, IEE-Proceedings-F, 140, 362
- Casandjian, J., Grenier, I., & for the Fermi Large Area Telescope Collaboration. 2009, ArXiv e-prints
- Casassus, S., Cabrera, G. F., Förster, F., et al. 2006, ApJ, 639, 951
- Chandrasekhar, S. 1960, Radiative transfer, ed. Chandrasekhar, S.
- Chiang, L., Naselsky, P. D., & Coles, P. 2009, ApJ, 694, 339
- Chiang, L., Naselsky, P. D., Verkhodanov, O. V., & Way, M. J. 2003, ApJL, 590, L65
- Cholis, I., Dobler, G., Finkbeiner, D. P., Goodenough, L., & Weiner, N. 2009, Phys. Rev. D, 80, 123518
- Cleary, K. A., Taylor, A. C., Waldram, E., et al. 2005, MNRAS, 360, 340
- Colombo, L. P. L. & Pierpaoli, E. 2010, MNRAS, 1116
- Copi, C. J., Huterer, D., Schwarz, D. J., & Starkman, G. D. 2006, MNRAS, 367, 79
- Copi, C. J., Huterer, D., & Starkman, G. D. 2004, Phys. Rev. D, 70, 043515
- Cruz, M., Tucci, M., Martínez-González, E., & Vielva, P. 2006, MNRAS, 369, 57
- Cruz, M., Turok, N., Vielva, P., Martínez-González, E., & Hobson, M. 2007, Science, 318, 1612
- Cumberbatch, D. T., Zuntz, J., Kamfjord Eriksen, H. K., & Silk, J. 2009, ArXiv e-prints
- Davies, R. D., Dickinson, C., Banday, A. J., et al. 2006, MNRAS, 370, 1125
- Davies, R. D., Watson, R. A., & Gutierrez, C. M. 1996, MNRAS, 278, 925
- Dawson, K. S., Holzapfel, W. L., Carlstrom, J. E., et al. 2002, ApJ, 581, 86
- de Oliveira-Costa, A., Kogut, A., Devlin, M. J., et al. 1997, ApJL, 482, L17+
- de Oliveira-Costa, A. & Tegmark, M. 2006, Phys. Rev. D, 74, 023005
- de Oliveira-Costa, A., Tegmark, M., Page, L. A., & Boughn, S. P. 1998, ApJL, 509, L9
- de Oliveira-Costa, A., Tegmark, M., Zaldarriaga, M., & Hamilton, A. 2004, Phys. Rev. D, 69, 063516
- de Zotti, G., Ricci, R., Mesa, D., et al. 2005, A&A, 431, 893
- Delabrouille, J. & Cardoso, J. 2007, ArXiv Astrophysics e-prints
- Delabrouille, J., Cardoso, J., Le Jeune, M., et al. 2009, A&A, 493, 835
- Delabrouille, J., Cardoso, J., & Patanchon, G. 2003, MNRAS, 346, 1089
- Dennison, B., Simonetti, J. H., & Topasna, G. A. 1998, Publications of the Astronomical Society of Australia, 15, 147

- Desert, F., Boulanger, F., & Puget, J. L. 1990, *A&A*, 237, 215
- Dicke, R. H., Peebles, P. J. E., Roll, P. G., & Wilkinson, D. T. 1965, *ApJ*, 142, 414
- Dickinson, C., Casassus, S., Pineda, J. L., et al. 2006, *ApJL*, 643, L111
- Dickinson, C., Davies, R. D., Allison, J. R., et al. 2009, *ApJ*, 690, 1585
- Dickinson, C., Davies, R. D., Bronfman, L., et al. 2007, *MNRAS*, 379, 297
- Dickinson, C., Davies, R. D., & Davis, R. J. 2003, *MNRAS*, 341, 369
- Dobler, G. & Finkbeiner, D. P. 2008a, *ApJ*, 680, 1222
- Dobler, G. & Finkbeiner, D. P. 2008b, *ApJ*, 680, 1235
- Dobler, G., Finkbeiner, D. P., Cholis, I., Slatyer, T. R., & Weiner, N. 2009, *ArXiv e-prints*
- Draine, B. T. & Lazarian, A. 1998a, *ApJL*, 494, L19+
- Draine, B. T. & Lazarian, A. 1998b, *ApJ*, 508, 157
- Dvali, G., Gruzinov, A., & Zaldarriaga, M. 2004, *Phys. Rev. D*, 69, 083505
- Dwek, E., Arendt, R. G., Fixsen, D. J., et al. 1997, *ApJ*, 475, 565
- Enqvist, K. & Sloth, M. S. 2002, *Nuclear Physics B*, 626, 395
- Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2004, *ApJ*, 612, 633
- Finkbeiner, D. P. 2003, *ApJS*, 146, 407
- Finkbeiner, D. P. 2004a, *ApJ*, 614, 186
- Finkbeiner, D. P. 2004b, *ArXiv Astrophysics e-prints*
- Finkbeiner, D. P., Davis, M., & Schlegel, D. J. 1999, *ApJ*, 524, 867
- Fixsen, D. J. & Mather, J. C. 2002, *ApJ*, 581, 817
- Freeman, P. E., Genovese, C. R., Miller, C. J., Nichol, R. C., & Wasserman, L. 2006, *ApJ*, 638, 1
- Gaustad, J. E., McCullough, P. R., Rosing, W., & Van Buren, D. 2001, *Publ. Astr. Soc. Pac.*, 113, 1326
- Ghosh, T., Delabrouille, J., Remazeilles, M., Cardoso, J., & Souradeep, T. 2010, *ArXiv e-prints*
- Gold, B., Bennett, C. L., Hill, R. S., et al. 2009, *ApJS*, 180, 265
- Gold, B., Odegard, N., Weiland, J. L., et al. 2010, *ArXiv e-prints*
- Haffner, L. M., Reynolds, R. J., Madsen, G. J., et al. 2010, *ArXiv e-prints*
- Haffner, L. M., Reynolds, R. J., Tufte, S. L., et al. 2003, *ApJS*, 149, 405
- Hansen, F. K., Banday, A. J., Eriksen, H. K., Górski, K. M., & Lilje, P. B. 2006, *ApJ*, 648, 784
- Hansen, F. K., Cabella, P., Marinucci, D., & Vittorio, N. 2004, *ApJL*, 607, L67

- Harrison, E. R. 1970, *Phys. Rev. D*, 1, 2726
- Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, *A&A Supp.*, 47, 1
- Hinshaw, G., Nolta, M. R., Bennett, C. L., et al. 2007, *ApJS*, 170, 288
- Hinshaw, G., Weiland, J. L., Hill, R. S., et al. 2009, *ApJS*, 180, 225
- Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, *ApJ*, 567, 2
- Hoang, T., Draine, B. T., & Lazarian, A. 2010, *ApJ*, 715, 1462
- Holder, G. P., McCarthy, I. G., & Babul, A. 2007, *MNRAS*, 382, 1697
- Hooper, D., Finkbeiner, D. P., & Dobler, G. 2007, *Phys. Rev. D*, 76, 083012
- Hooper, D., Zaharijas, G., Finkbeiner, D. P., & Dobler, G. 2008, *Phys. Rev. D*, 77, 043511
- Hu, W. 2003, *Annals of Physics*, 303, 203
- Hu, W. & White, M. 1997, *New Astronomy*, 2, 323
- Hubble, E. 1929, *Proceedings of the National Academy of Science*, 15, 168
- Huffenberger, K. M., Eriksen, H. K., & Hansen, F. K. 2006, *ApJL*, 651, L81
- Huffenberger, K. M., Eriksen, H. K., Hansen, F. K., Banday, A. J., & Górski, K. M. 2008, *ApJ*, 688, 1
- Hyvarinen, A. 1999, *IEEE Signal Processing Letters*, 6, 145
- Hyvarinen, A. & Oja, E. 2000, *Neural Networks*, 13, 411
- Iglesias-Groth, S. 2006, *MNRAS*, 368, 1925
- Jaffe, T. R. 2006, *Morphological Studies of the CMB: Non-standard Models and Foregrounds*
- Jonas, J. L., Baart, E. E., & Nicolson, G. D. 1998, *MNRAS*, 297, 977
- Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997a, *Physical Review Letters*, 78, 2058
- Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997b, *Phys. Rev. D*, 55, 7368
- Kaplinghat, M., Phalen, D. J., & Zurek, K. M. 2009, *Journal of Cosmology and Astro-Particle Physics*, 12, 10
- Kim, J., Naselsky, P., & Christensen, P. R. 2008, *Phys. Rev. D*, 77, 103002
- Kirk, J. M., Polehampton, E., Anderson, L. D., et al. 2010, *A&A*, 518, L82+
- Kogut, A. 1999, in *Astronomical Society of the Pacific Conference Series*, Vol. 181, *Microwave Foregrounds*, ed. A. de Oliveira-Costa & M. Tegmark, 91–+
- Kogut, A., Banday, A. J., Bennett, C. L., et al. 1996a, *ApJ*, 460, 1
- Kogut, A., Banday, A. J., Bennett, C. L., et al. 1996b, *ApJL*, 464, L5+
- Kogut, A., Lineweaver, C., Smoot, G. F., et al. 1993, *ApJ*, 419, 1

- Komatsu, E., Kogut, A., Nolta, M. R., et al. 2003, *ApJS*, 148, 119
- Kosowsky, A. 1996, *Annals of Physics*, 246, 49
- Kovac, J. M., Leitch, E. M., Pryke, C., et al. 2002, *Nature*, 420, 772
- Kuo, C. L., Ade, P. A. R., Bock, J. J., et al. 2004, *ApJ*, 600, 32
- Land, K. & Magueijo, J. 2005, *Physical Review Letters*, 95, 071301
- Land, K. & Magueijo, J. 2007, *MNRAS*, 378, 153
- Lawson, K. D., Mayer, C. J., Osborne, J. L., & Parkinson, M. L. 1987, *MNRAS*, 225, 307
- Lazarian, A. & Draine, B. T. 2000, *ApJL*, 536, L15
- Leach, S. M., Cardoso, J., Baccigalupi, C., et al. 2008, *A&A*, 491, 597
- Leitch, E. M., Kovac, J. M., Halverson, N. W., et al. 2005, *ApJ*, 624, 10
- Leitch, E. M., Kovac, J. M., Pryke, C., et al. 2002, *Nature*, 420, 763
- Leitch, E. M., Readhead, A. C. S., Pearson, T. J., & Myers, S. T. 1997, *ApJL*, 486, L23+
- Linden, T. & Profumo, S. 2010, *ApJL*, 714, L228
- Lorimer, D. R., Faulkner, A. J., Lyne, A. G., et al. 2006, *MNRAS*, 372, 777
- Lyth, D. H. & Wands, D. 2002, *Physics Letters B*, 524, 5
- Maino, D., Donzelli, S., Banday, A. J., Stivoli, F., & Baccigalupi, C. 2007, *MNRAS*, 374, 1207
- Maino, D., Farusi, A., Baccigalupi, C., et al. 2002, *MNRAS*, 334, 53
- Mason, B. S., Pearson, T. J., Readhead, A. C. S., et al. 2003, *ApJ*, 591, 540
- McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2005, *MNRAS*, 359, 1583
- McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2006, *MNRAS*, 371, L50
- McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2008, *MNRAS*, 388, 659
- McQuinn, M. & Zaldarriaga, M. 2010, *ArXiv e-prints*
- Mertsch, P. & Sarkar, S. 2010, *ArXiv e-prints*
- Miville-Deschênes, M., Abergel, A., & Boulanger, F. 2001, in *ESA Special Publication*, Vol. 460, *The Promise of the Herschel Space Observatory*, ed. G. L. Pilbratt, J. Cernicharo, A. M. Heras, T. Prusti, & R. Harris, 471–+
- Miville-Deschênes, M., Lagache, G., Boulanger, F., & Puget, J. 2007, *A&A*, 469, 595
- Miville-Deschênes, M., Ysard, N., Lavabre, A., et al. 2008, *A&A*, 490, 1093
- Mollerach, S. 1990, *Phys. Rev. D*, 42, 313
- Monteserín, C., Barreiro, R. B., Vielva, P., et al. 2008, *MNRAS*, 387, 209

- Montroy, T. E., Ade, P. A. R., Bock, J. J., et al. 2006, *ApJ*, 647, 813
- Moroi, T. & Takahashi, T. 2001, *Physics Letters B*, 522, 215
- Narkowich, F., Petrushev, P., & Ward, J. 2006, *SIAM J. Math Anal.*, 38, 530
- Nolta, M. R., Dunkley, J., Hill, R. S., et al. 2009, *ApJS*, 180, 296
- Page, L., Hinshaw, G., Komatsu, E., et al. 2007, *ApJS*, 170, 335
- Paladini, R., Burigana, C., Davies, D. R., et al. 2002, *VizieR Online Data Catalog*, 339, 70213
- Paladini, R., Davies, R. D., & De Zotti, G. 2004, *MNRAS*, 347, 237
- Park, C. 2004, *MNRAS*, 349, 313
- Park, C., Park, C., & Gott, III, J. R. 2007, *ApJ*, 660, 959
- Patanchon, G., Cardoso, J., Delabrouille, J., & Vielva, P. 2005, *MNRAS*, 364, 1185
- Piacentini, F., Ade, P. A. R., Bock, J. J., et al. 2006, *ApJ*, 647, 833
- Planck User Manual. 2008
- Platania, P., Bensadoun, M., Bersanelli, M., et al. 1998, *ApJ*, 505, 473
- Platania, P., Burigana, C., Maino, D., et al. 2003, *A&A*, 410, 847
- Pollack, J. B., Hollenbach, D., Beckwith, S., et al. 1994, *ApJ*, 421, 615
- Readhead, A. C. S., Myers, S. T., Pearson, T. J., et al. 2004, *Science*, 306, 836
- Rees, M. J. & Sciama, D. W. 1968, *Nature*, 217, 511
- Reich, P. & Reich, W. 1986, *A&A Supp.*, 63, 205
- Reich, P. & Reich, W. 1988, *A&A*, 196, 211
- Reich, P., Testori, J. C., & Reich, W. 2001, *A&A*, 376, 861
- Ricci, R., Sadler, E. M., Ekers, R. D., et al. 2004, *MNRAS*, 354, 305
- Rubiño-Martín, J. A., Rebolo, R., & Mediavilla, E. 2010, *The Cosmic Microwave Background. From quantum fluctuations to the Present Universe*, ed. Rubiño-Martín, J. A. and Rebolo, R. and Mediavilla, E.
- Rybicki, G. B. & Lightman, A. P. 1986, *Radiative Processes in Astrophysics*, ed. Rybicki, G. B. & Lightman, A. P.
- Sachs, R. K. & Wolfe, A. M. 1967, *ApJ*, 147, 73
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525
- Seljak, U. 1996, *ApJ*, 463, 1
- Seljak, U. & Zaldarriaga, M. 1997, *Physical Review Letters*, 78, 2054
- Shannon, C. E. & Weaver, W. 1949, *The mathematical theory of communication*, ed. Shannon, C. E. & Weaver, W.

- Silk, J. 1968, *ApJ*, 151, 459
- Smoot, G. F. 1997, in *Microwave Background Anisotropies*, 3–14
- Smoot, G. F., Bennett, C. L., Kogut, A., et al. 1992, *ApJL*, 396, L1
- Spergel, D. N., Bean, R., Doré, O., et al. 2007, *ApJS*, 170, 377
- Starobinskij, A. A. 1993, *Soviet Journal of Experimental and Theoretical Physics Letters*, 57, 622
- Su, M., Slatyer, T. R., & Finkbeiner, D. P. 2010, *ArXiv e-prints*
- Sunyaev, R. A. & Zeldovich, Y. B. 1972, *Comments on Astrophysics and Space Physics*, 4, 173
- Sutherland, R. S. 1998, *MNRAS*, 300, 321
- Tegmark, M., de Oliveira-Costa, A., & Hamilton, A. J. 2003, *Phys. Rev. D*, 68, 123523
- Tegmark, M. & Efstathiou, G. 1996, *MNRAS*, 281, 1297
- Toffolatti, L., Argüeso Gomez, F., de Zotti, G., et al. 1998, *MNRAS*, 297, 117
- Toffolatti, L., Negrello, M., González-Nuevo, J., et al. 2005, *A&A*, 438, 475
- Tucci, M., Martínez-González, E., Vielva, P., & Delabrouille, J. 2005, *MNRAS*, 360, 935
- Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., & Cayón, L. 2004, *ApJ*, 609, 22
- Vio, R. & Andreani, P. 2008, *ArXiv e-prints*
- Waldram, E. M., Pooley, G. G., Grainge, K. J. B., et al. 2003, *MNRAS*, 342, 915
- Watson, R. A., Rebolo, R., Rubiño-Martín, J. A., et al. 2005, *ApJL*, 624, L89
- Wright, E. L., Bennett, C. L., Gorski, K., Hinshaw, G., & Smoot, G. F. 1996, *ApJL*, 464, L21+
- Wright, E. L., Chen, X., Odegard, N., et al. 2009, *ApJS*, 180, 283
- Ysard, N., Miville-Deschênes, M. A., & Verstraete, L. 2010, *A&A*, 509, L1+
- Zaldarriaga, M. 1997, *Phys. Rev. D*, 55, 1822
- Zaldarriaga, M. & Seljak, U. 1999, *Phys. Rev. D*, 59, 123507
- Zeldovich, Y. B. 1972, *MNRAS*, 160, 1P
- Zhang, J., Bi, X., Liu, J., et al. 2009, *Phys. Rev. D*, 80, 023007



## Refereed publications

- *New insights into foreground analysis of the WMAP five-year data using*  
**Bottino M.**, Banday A.J., Maino D.  
2010, MNRAS 402, 207B
- *Foreground analysis of the WMAP three-year data with*  
**Bottino M.**, Banday A.J., Maino D.  
2008, MNRAS 389, 1190B