



**HAL**  
open science

# Détection de l'effet de lentille gravitationnelle dans les données de Planck-HFI

Alexis Lavabre

► **To cite this version:**

Alexis Lavabre. Détection de l'effet de lentille gravitationnelle dans les données de Planck-HFI. Autre. Université Paris Sud - Paris XI, 2011. Français. NNT : 2011PA112060 . tel-00923133

**HAL Id: tel-00923133**

**<https://theses.hal.science/tel-00923133>**

Submitted on 2 Jan 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS-SUD 11

LABORATOIRE DE L'ACCÉLÉRATEUR LINÉAIRE

ÉCOLE DOCTORALE : ASTRONOMIE ET ASTROPHYSIQUE D'ÎLE-DE-FRANCE

**THÈSE DE DOCTORAT**

Soutenue le 25 mai 2011

par

**Alexis Lavabre**

---

**Détection de l'effet de lentille  
gravitationnelle dans les données  
de Planck-HFI**

---

**Directeur de thèse :** Stéphane PLASZCZYNSKI

**Composition du Jury :**

*Président du jury :* Achille STOCCHI

*Rapporteurs :* François-Xavier DÉSSERT

Paolo NATOLI

*Examineurs :* Julien LESGOURGUES

Jean-Luc STARCK

*Membres invité :* Jean-Loup PUGET

---

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>I Bases théoriques</b>	<b>3</b>
<b>1 Cosmologie</b>	<b>5</b>
1.1 Piliers du modèle . . . . .	5
1.1.1 Principe cosmologique . . . . .	5
1.1.2 Piliers observationnels du modèle du <i>Big Bang</i> . . . . .	7
1.2 Notions de Relativité Générale . . . . .	8
1.2.1 Géométrie de l'Univers . . . . .	8
1.2.2 Équation d'Einstein . . . . .	8
1.3 Modèle de Friedmann–Lemaître–Robertson–Walker . . . . .	9
1.3.1 Métrique de Robertson–Walker . . . . .	9
1.3.2 Équations de Friedmann–Lemaître . . . . .	10
1.4 Le <i>Big Bang</i> . . . . .	12
1.4.1 Évolution du contenu de l'Univers . . . . .	12
1.4.2 Limites du modèle classique du <i>Big Bang</i> . . . . .	13
1.4.3 Inflation . . . . .	14
<b>2 Fond micro-onde cosmologique</b>	<b>15</b>
2.1 Historique et présentation . . . . .	15
2.1.1 Historique . . . . .	15
2.1.2 Éléments de physique du CMB . . . . .	16
2.1.3 Spectre électromagnétique du CMB . . . . .	17
2.2 Anisotropies de température . . . . .	18
2.2.1 Dipole . . . . .	18
2.2.2 Anisotropies primaires . . . . .	19
2.2.3 Anisotropies secondaires . . . . .	19
2.3 Spectre de puissance angulaire des anisotropies de température . . . . .	21
2.3.1 Analyse multipolaire . . . . .	21
2.3.2 Propriétés des harmoniques sphériques . . . . .	21
2.3.3 Projection sur la sphère : statistique des $a_{lm}$ . . . . .	22
2.3.4 Erreur de mesure et variance cosmique . . . . .	22
2.3.5 Spectre de puissance . . . . .	23
2.4 Polarisation . . . . .	23
2.4.1 Polarisation d'une onde électromagnétique - Paramètres de Stokes . . . . .	23
2.4.2 Diffusion Thomson . . . . .	26
2.4.3 Polarisation du CMB . . . . .	26
2.4.4 Les champs E et B . . . . .	27
2.4.5 Spectres de puissance de polarisation du CMB . . . . .	29
2.5 Émissions d'avant-plan . . . . .	30
2.5.1 Émissions diffuses . . . . .	30
2.5.2 Émission des sources compactes . . . . .	31
2.5.3 Émissions polarisées . . . . .	32

<b>3</b>	<b>Effet de lentille gravitationnelle sur le CMB</b>	<b>33</b>
3.1	Présentation . . . . .	33
3.2	Cadre général . . . . .	34
3.3	Hypothèses . . . . .	34
3.3.1	Équation de lentille . . . . .	35
3.3.2	Potentiel de lentille . . . . .	36
3.3.3	Ordres de grandeur et validité des approximations . . . . .	37
3.4	Effet sur les observables du CMB . . . . .	38
3.4.1	Effet sur les cartes du CMB . . . . .	38
3.4.2	Effet sur le spectre de puissance des anisotropies de température . . . . .	39
3.4.3	Effet sur les spectres de polarisation du CMB . . . . .	41
<b>II</b>	<b>Compression des données de l'instrument HFI de Planck</b>	<b>43</b>
<b>4</b>	<b>L'expérience Planck</b>	<b>45</b>
4.1	Historique et présentation . . . . .	45
4.2	Objectifs scientifiques . . . . .	47
4.3	Description du satellite . . . . .	47
4.3.1	Gestion thermique . . . . .	48
4.3.2	Le système cryogénique . . . . .	49
4.3.3	Le système optique . . . . .	51
4.3.4	Le senseur stellaire . . . . .	51
4.3.5	Le SREM . . . . .	52
4.3.6	Communication avec la Terre . . . . .	53
4.3.7	Stratégie de pointage . . . . .	53
4.4	L'instrument basse fréquence (LFI) . . . . .	54
4.5	L'instrument haute fréquence (HFI) . . . . .	55
<b>5</b>	<b>Compression des données de l'instrument HFI à bord du satellite</b>	<b>61</b>
5.1	Position du problème . . . . .	61
5.1.1	Contraintes liées aux expériences spatiales . . . . .	61
5.1.2	Chaîne de transmission des données . . . . .	61
5.1.3	Transport des paquets . . . . .	62
5.1.4	Gestion du débit de données . . . . .	63
5.1.5	Compression des données . . . . .	63
5.2	Algorithme de compression . . . . .	64
5.2.1	Choix de l'algorithme . . . . .	64
5.2.2	Opérateur de quantification . . . . .	65
5.2.3	Théorèmes de quantification . . . . .	65
5.2.4	Application à un bruit blanc gaussien . . . . .	68
5.2.5	Application aux données de l'instrument HFI . . . . .	69
5.3	Compression des données de l'instrument HFI . . . . .	69
5.3.1	Signal brut de l'instrument HFI . . . . .	69
5.3.2	Application aux données de PLANCK . . . . .	71
5.3.3	Erreurs de compression . . . . .	75
5.4	Réglages des paramètres de compression . . . . .	77
5.4.1	Procédure de réglage des paramètres . . . . .	77
5.4.2	Réglages pendant la phase pré-observation . . . . .	78
5.4.3	Outils de surveillance quotidienne . . . . .	78
5.4.4	Réglages initiaux . . . . .	78
5.4.5	Réglages fins . . . . .	79
5.4.6	Suivi de l'évolution des paramètres de compression . . . . .	80
5.5	Suivi du débit . . . . .	84
5.5.1	Débit d'un signal purement gaussien . . . . .	84
5.5.2	Évolution réelle du débit . . . . .	85
5.5.3	Effet de l'émission de la poussière galactique . . . . .	85

5.5.4	Effet du transpondeur . . . . .	89
5.5.5	Effet du bruit télégraphique . . . . .	91
5.5.6	Événements particuliers . . . . .	92
5.6	Suivi des erreurs de compression . . . . .	92
5.6.1	Recherche de trous dans les données provoqués par la saturation du tampon mémoire . . . . .	92
5.6.2	Recherche des EOS . . . . .	93
5.6.3	Recherche des erreurs de compression . . . . .	93
5.7	Ajustement des paramètres de compression pour la seconde couverture du ciel . . . . .	99
5.8	Bilan . . . . .	102
<b>III</b>	<b>Estimation du spectre de déflexion dans les données de HFI</b>	<b>103</b>
<b>6</b>	<b>Simulation et analyse de signaux pour Planck</b>	<b>105</b>
6.1	Simulation des spectres de puissance . . . . .	105
6.1.1	Paramètres cosmologiques . . . . .	105
6.1.2	Simulation avec CAMB . . . . .	106
6.2	Simulation de cartes gaussiennes . . . . .	107
6.2.1	Simulation sur la sphère . . . . .	107
6.2.2	Simulation sur le plan . . . . .	109
6.3	Simulation de l'effet de lentille gravitationnelle . . . . .	110
6.3.1	Simulation sur la sphère . . . . .	110
6.3.2	Simulation sur le plan . . . . .	111
6.4	Simulation des effets instrumentaux . . . . .	113
6.4.1	Simulation de l'effet du lobe de l'instrument sur les cartes . . . . .	113
6.4.2	Simulation du bruit instrumental . . . . .	115
6.5	Estimation du spectre de puissance du CMB . . . . .	116
6.5.1	Estimation sur la sphère . . . . .	116
6.5.2	Estimation sur le plan . . . . .	117
6.6	Estimation du spectre de déflexion . . . . .	118
6.6.1	Présentation . . . . .	118
6.6.2	Dérivation dans le plan . . . . .	118
6.6.3	Bruits de l'estimateur . . . . .	121
6.6.4	Généralisation à la sphère . . . . .	122
<b>7</b>	<b>Développement d'une méthode d'analyse par patch pour Planck</b>	<b>123</b>
7.1	Outils de découpe d'une sphère en patch . . . . .	123
7.1.1	Problématique . . . . .	123
7.1.2	Pavage de la sphère . . . . .	123
7.1.3	Projection . . . . .	125
7.1.4	Estimation du signal sur une grille irrégulière . . . . .	126
7.1.5	Apodisation . . . . .	129
7.2	Outils d'inpainting local de la sphère . . . . .	130
7.2.1	Problématique . . . . .	130
7.2.2	Découpe des patches . . . . .	132
7.2.3	Inpainting local . . . . .	132
7.2.4	Reprojection . . . . .	134
7.3	Validation . . . . .	135
7.3.1	Algorithme d'approximation par série de Fourier . . . . .	135
7.3.2	Méthode d'estimation du signal projeté sur une grille régulière . . . . .	136
7.4	Optimisations . . . . .	137
7.4.1	Borne de l'intégrale pour l'estimateur du potentiel de lentille . . . . .	137
7.4.2	Taille des patches . . . . .	140
7.4.3	Taille de la grille de coefficients à estimer . . . . .	140
7.4.4	Choix de la fenêtre d'apodisation . . . . .	142
7.5	Traitement des erreurs . . . . .	143

7.5.1	Méthode Monte-Carlo . . . . .	143
7.5.2	Estimation des erreurs par Monte-Carlo . . . . .	143
7.5.3	Correction des erreurs systématiques par Monte-Carlo . . . . .	144
<b>8</b>	<b>Estimation de l'effet de lentille gravitationnelle dans les données de HFI</b>	<b>145</b>
8.1	Préparation des données . . . . .	145
8.1.1	Traitement des données ordonnées en temps . . . . .	145
8.1.2	Production des cartes . . . . .	146
8.1.3	Inpainting . . . . .	147
8.1.4	Séparation des composantes . . . . .	147
8.2	Jeux de données de départ pour l'analyse . . . . .	150
8.2.1	Données DX4 . . . . .	150
8.2.2	Carte des mesures combinées à 143GHz . . . . .	150
8.2.3	Carte des mesures combinées à 217GHz . . . . .	153
8.2.4	Carte issue de la séparation des composantes avec GMCA . . . . .	153
8.3	Masquage des données . . . . .	154
8.3.1	Masques de couverture . . . . .	154
8.3.2	Masques de sources compactes . . . . .	156
8.3.3	Masques galactiques . . . . .	156
8.3.4	Masques combinés . . . . .	159
8.4	Sélection des patches . . . . .	160
8.4.1	Sélection sur le nombre de pixels masqués . . . . .	160
8.4.2	Sélection sur des critères de gaussianité . . . . .	164
8.4.3	Sélection sur des critères d'isotropie . . . . .	165
8.5	Analyse préliminaire . . . . .	165
8.5.1	Découpe des patches . . . . .	165
8.5.2	Déconvolution des effets du lobe de l'instrument . . . . .	166
8.5.3	Estimation du spectre de puissance de température . . . . .	168
8.6	Estimation du spectre de la déflexion . . . . .	168
8.6.1	Estimation du potentiel de lentille . . . . .	168
8.6.2	Estimation du spectre de puissance angulaire de la déflexion . . . . .	170
8.6.3	Correction des biais de l'estimateur . . . . .	170
8.7	Correction des effets systématiques par Monte-Carlo . . . . .	172
8.7.1	Jeux de simulations . . . . .	173
8.7.2	Détermination de la correction . . . . .	174
8.7.3	Validation de la correction Monte-Carlo . . . . .	174
8.7.4	Correction du spectre de puissance de la déflexion sur les données . . . . .	174
	<b>Conclusions</b>	<b>179</b>
<b>A</b>	<b>Champ aléatoire en cosmologie</b>	<b>183</b>
<b>B</b>	<b>Approximation par séries de Fourier de fonctions irrégulièrement échantillonnées</b>	<b>185</b>
B.1	Présentation à une dimension . . . . .	185
B.1.1	Position du problème . . . . .	185
B.1.2	Formalisme . . . . .	185
B.1.3	Mise en œuvre . . . . .	186
B.2	Généralisation à deux dimensions . . . . .	186
B.2.1	Outils mathématiques . . . . .	186
B.2.2	Problème à deux dimensions . . . . .	187
B.2.3	Résolution . . . . .	188
B.2.4	Mise en œuvre . . . . .	189
<b>C</b>	<b>Paramètres de compression</b>	<b>191</b>
<b>D</b>	<b>Liste des acronymes</b>	<b>193</b>

# Table des figures

1.1	Relevé de galaxies 2dF . . . . .	6
1.2	Évolution de l’Univers . . . . .	12
2.1	Spectre électromagnétique du CMB . . . . .	17
2.2	Influence de l’effet Sunyaev-Zel’dovich sur le spectre de puissance du CMB . . . . .	20
2.3	Spectre de puissance de température du CMB . . . . .	24
2.4	Représentation schématique du spectre de puissance de température du CMB . . . . .	24
2.5	Diffusion Thomson . . . . .	26
2.6	Polarisation du rayonnement . . . . .	27
2.7	Génération de la polarisation par les perturbations scalaires . . . . .	28
2.8	Spectres de puissance du CMB . . . . .	30
2.9	Émissions des composantes en fonction de la longueur d’onde . . . . .	31
3.1	Approximation des lentilles minces . . . . .	35
3.2	Notations pour l’équation de la déflexion . . . . .	36
3.3	Spectres de puissance de la déflexion . . . . .	38
3.4	Effet de la déflexion sur le CMB . . . . .	40
3.5	Spectre de corrélation entre la température du CMB et le potentiel de lentille . . . . .	41
3.6	Impact de l’effet de lentille gravitationnelle sur le spectre de puissance de température . . . . .	42
4.1	Transmission de l’atmosphère en fonction de la longueur d’onde . . . . .	46
4.2	Vue d’artiste du satellite PLANCK . . . . .	48
4.3	Orbite du satellite au point de Lagrange L2 . . . . .	49
4.4	Schéma de la chaîne cryogénique de PLANCK . . . . .	50
4.5	Schéma du télescope de PLANCK . . . . .	51
4.6	Schéma du dispositif placé avant les bolomètres . . . . .	52
4.7	Schéma du plan focal de PLANCK . . . . .	54
4.8	Schéma des détecteurs de l’instrument LFI . . . . .	55
4.9	Schéma des détecteurs de l’instrument HFI . . . . .	56
4.10	Schéma de l’électronique de lecture et de polarisation des bolomètres . . . . .	58
5.1	Schéma de la chaîne de transmission des données . . . . .	62
5.2	Effet de l’opération de quantification sur un signal à valeurs réelles . . . . .	65
5.3	Erreur de quantification sur un signal à valeurs réelles . . . . .	66
5.4	Effet de la quantification sur la densité de probabilité du signal . . . . .	67
5.5	Densité de probabilité du bruit de quantification . . . . .	68
5.6	Allure des composantes du signal brut d’un bolomètre . . . . .	70
5.7	Période du signal brut sur-échantillonné . . . . .	71
5.8	Schéma des étapes de la compression des données . . . . .	72
5.9	Évolution du rapport $q/\sigma$ au cours du temps . . . . .	83
5.10	Corrélation entre le rapport $q/\sigma$ de thermomètres et le taux de rayons cosmiques . . . . .	84
5.11	Évolution du débit moyen par échantillon pour chaque voie . . . . .	88
5.12	Évolution du débit total . . . . .	89
5.13	Impact de l’émission de la galaxie sur le débit . . . . .	90
5.14	Effet sur le débit de la mise sous tension du transpondeur . . . . .	91



5.15	Évolution de nombre d'échantillons perdus à cause des erreurs de compression . . . . .	96
5.16	Erreur de compression produite par l'impact d'un rayon cosmique . . . . .	98
5.17	Erreur de compression lors du balayage du centre galactique sur les voies à haute fréquence de l'instrument HFI . . . . .	99
5.18	Erreur de compression sur Jupiter pour la quatrième voie à 857GHz . . . . .	100
5.19	Effet du changement des paramètres de compression sur le débit . . . . .	101
6.1	Spectres de puissance angulaire théoriques des anisotropies de température et de la déflexion . . . . .	106
6.2	Pixelisation HEALPix . . . . .	108
6.3	Carte du ciel gaussienne. . . . .	109
6.4	Simulation d'une carte du CMB lentillée dans le plan . . . . .	112
6.5	Simulation des effets du lobe de l'optique et du bruit . . . . .	115
7.1	Pavage de la sphère et recouvrement . . . . .	124
7.2	Projection gnomonique . . . . .	126
7.3	Distorsion radiale . . . . .	127
7.4	Interpolation pour la projection . . . . .	128
7.5	Répartitions des pixels projetés . . . . .	129
7.6	Effet de la taille finie des patchs sur le spectre de puissance bi-dimensionnel . . . . .	130
7.7	Inpainting par patch d'une carte de CMB . . . . .	134
7.8	Comparaison dans l'espace de Fourier des estimations par FFT et par ajustement de séries de Fourier . . . . .	136
7.9	Comparaison dans l'espace réel des estimations par FFT et par ajustement de séries de Fourier . . . . .	137
7.10	Dépendance de la taille de la grille, de la méthode d'estimation par série de Fourier . . . . .	138
7.11	Comparaison des méthodes de découpe des patchs . . . . .	139
7.12	Comparaison des spectres de puissance de température pour les différentes méthodes de découpe . . . . .	139
7.13	Influence de la taille des patchs sur le spectre de la déflexion . . . . .	141
7.14	Comparaison des fonctions fenêtre de Kaiser pour $\alpha = 0.5$ et 1 et de la fonction « tophat » . . . . .	142
7.15	Influence du paramètre $\alpha$ de la fonction de Kaiser sur le spectre de puissance . . . . .	143
8.1	Nombre d'échantillons par pixel pour les cartes à 143GHz et 217GHz du jeu de données DX4 . . . . .	148
8.2	Exemple d'inpainting de la carte à 217GHz . . . . .	149
8.3	Cartes du jeu de données DX4, utilisées pour l'analyse . . . . .	151
8.4	Cartes du jeu de données DX4, utilisées pour l'analyse, masquées et inpaintées . . . . .	152
8.5	Masques de couverture . . . . .	155
8.6	Masques de sources compactes . . . . .	157
8.7	Masques galactiques . . . . .	158
8.8	Sélection des patchs . . . . .	161
8.9	Distributions des tests de gaussianité . . . . .	163
8.10	Exemple d'un patch découpé dans les données . . . . .	166
8.11	Profil des lobes dans l'espace de Fourier . . . . .	167
8.12	Spectre de puissance de température brut, par patch, estimé sur les données déconvoluées . . . . .	169
8.13	Estimation brute du spectre de puissance de la déflexion . . . . .	171
8.14	Biais de l'estimateur pour la carte à 217GHz . . . . .	172
8.15	Validation de la correction du spectre de la déflexion par Monte-Carlo . . . . .	175
8.16	Spectres de puissance de la déflexion des cartes combinées, corrigés par Monte-Carlo . . . . .	176
8.17	Spectre de puissance de la déflexion combiné . . . . .	177
8.18	Spectre de puissance de la déflexion de la carte GMCA, corrigé par Monte-Carlo . . . . .	178

# Liste des tableaux

5.1	Tableau récapitulatif du dictionnaire pour la compression. . . . .	74
6.1	Tableau récapitulatif de l'ellipticité des lobes . . . . .	114
8.1	Tableau récapitulatif du pourcentage de pixels masqués pour les différents masques .	154
8.2	Tableau récapitulatif de la taille des cercles masqués en fonction de la fréquence . . .	156
8.3	Tableau récapitulatif du nombre de patches restant après sélection . . . . .	160
8.4	Tableau récapitulatif de la taille du lobe de l'instrument en fonction de la fréquence	168
C.1	Tableau récapitulatif des paramètres de compression . . . . .	192



# Introduction

La cosmologie moderne a révolutionné notre vision de l'Univers. Ses bases théoriques et observationnelles sont maintenant suffisamment solides pour constituer un socle commun qui est de moins en moins remis en question. On est entré aujourd'hui dans l'ère de la cosmologie de précision et les mesures effectuées à partir des instruments modernes permettent de poser de fortes contraintes sur l'évolution cosmologique et de déterminer les paramètres des modèles avec une grande précision. Cependant de nombreuses questions restent toujours sans réponse, en particulier sur la nature et l'origine de l'énergie et de la matière noire, de l'inflation... La compréhension de ces phénomènes constitue un des enjeux majeurs de la cosmologie moderne.

Dans ce contexte, l'étude du fond diffus cosmologique (ou CMB pour le terme anglais *Cosmic Microwave Background*) est particulièrement intéressante. Prédit dès la fin des années 40 et découvert en 1965 par Penzias et Wilson, le CMB fait parti des piliers du modèle cosmologique du *Big Bang*. C'est une observable unique, provenant de l'époque de la recombinaison qui correspond à un moment où l'Univers était âgé de seulement 380 000 ans. Ce rayonnement fossile est la lumière la plus lointaine et par conséquent la plus vieille qu'on puisse observer dans l'Univers actuel et nous renseigne sur les conditions physiques de l'Univers primordial. Son degré élevé d'isotropie permet de conforter le principe cosmologique d'un Univers spatialement homogène et isotrope. De faibles écarts par rapport à son isotropie ont longtemps été recherchés. Le satellite COBE les a mis en évidence pour la première fois en 1992. Ses résultats ont marqué un tournant dans la cosmologie observationnelle. Par la suite de nombreuses expériences au sol, emportées par un ballon stratosphérique dans les hautes couches de l'atmosphère ou par un satellite dans l'espace, ont été menées afin de caractériser les propriétés de ces anisotropies. Les résultats les plus précis ont été obtenus par le satellite WMAP de la NASA.

Le satellite PLANCK est le satellite de troisième génération qui fait suite aux satellites COBE et WMAP et qui est dédié à l'étude du fond diffus cosmologique. Après plus de 15 ans de développement, par une collaboration mondiale, impliquant plus d'une quinzaine de laboratoires dans le monde, il a été lancé le 14 mai 2009 à 14 heures 12 minutes et 2 secondes. Les performances de ses instruments, en vol, se révélant d'un niveau remarquable, il devrait permettre de faire des mesures des anisotropies de température mais surtout de la polarisation du CMB, avec une sensibilité, une résolution angulaire et une couverture spectrale inégalées. C'est dans le cadre de la collaboration PLANCK, regroupant plus de 300 chercheurs et autant d'ingénieurs, d'horizons souvent différents, que j'ai effectué mon travail de thèse. Il s'est articulé autour de deux axes principaux : d'une part le développement d'outils d'analyse du ciel par morceaux que j'ai utilisé ensuite pour effectuer la première mesure de l'effet de lentille gravitationnelle sur le CMB, grâce aux données de l'instrument HFI et d'autre part l'optimisation et la surveillance régulière de la compression et du débit des données de l'instrument HFI depuis le début des observations. Ce programme de travail a été adapté au retard du lancement de PLANCK, par rapport à la date prévue en début de thèse.

La qualité des observations de PLANCK, donne accès pour la première fois aux effets de lentille gravitationnelle sur le CMB. Au cours de ma thèse, je me suis concentré sur le développement d'une méthode d'analyse originale, afin de mesurer le spectre de déflexion de la matière. L'approche que j'ai développée est basée sur une étude du ciel par morceaux. Elle a été motivée par la difficulté de traiter les contaminations galactiques dans une approche sur tout le ciel. Aucune mesure de l'effet de lentille gravitationnelle sur le CMB n'avait été faite auparavant et l'impact de la séparation de composantes n'avait pas été évalué. Par ailleurs, l'estimateur utilisé avait été développé dans un cas idéal et n'avait jamais été utilisé sur des vraies données, ce qui a nécessité des développements importants pour la caractérisation et la prise en compte des détails nécessaires au traitement des

données de PLANCK. En particulier, il a fallu trouver des solutions au problème du masquage des données, du bruit inhomogène des conditions aux bords. La solution retenue s'est appuyée sur une correction à base de simulations Monte-Carlo de cartes de température lentillées avec un bruit inhomogène et un lobe instrumental gaussien circulaire proches de ceux de PLANCK et a nécessité une attention particulière qui a nécessité de longs traitements.

L'équipe du LAL qui m'a accueilli pendant ma thèse a intégré le consortium PLANCK dès 1997. Elle a la responsabilité officielle de la production des cartes à partir des données ordonnées en temps ainsi que leur calibration pour l'instrument HFI. Elle a également participé au développement instrumental et a notamment développé et réalisé le DPU qui est l'ordinateur central de l'instrument HFI. C'est un élément essentiel qui est chargé de la gestion de l'électronique de lecture des bolomètres et de la cryogénie, de l'exécution des télécommandes envoyées par l'ordinateur de bord et du traitement des données de l'instrument HFI à bord du satellite, en particulier la compression des données brutes. Stéphane Plaszczyński et François Couchot ont développé l'algorithme de compression et ont la responsabilité de son bon fonctionnement pendant la mission. En parallèle de mon travail sur l'effet de lentille gravitationnelle, j'ai eu en charge l'optimisation des paramètres de cet algorithme, ainsi que la surveillance systématique des erreurs de compression et du débit des données, depuis le début des observations. J'ai entièrement développé les outils nécessaires à ces tâches et je les ai fournis à la communauté, afin qu'il soit possible de poursuivre cet effort après ma thèse. Ce travail s'inscrit dans la continuité des responsabilités de l'équipe du LAL et dans un effort plus large mené au sein de la collaboration qui consiste à s'assurer jour après jour du bon fonctionnement de l'ensemble de la chaîne de mesure depuis le détecteur jusqu'au centre de calcul.

L'exposé de mon travail, est organisé en trois parties principales. Une première partie introductive, suivie d'une partie sur la compression des données de l'instrument HFI puis une dernière partie consacrée aux développements de la méthode d'analyse de l'effet de lentille gravitationnelle originale que j'ai développée et son application aux données de l'instrument HFI de PLANCK. Les trois premiers chapitres introduisent respectivement le modèle cosmologique standard du *Big Bang*, les principales propriétés du rayonnement fossile et de l'effet de lentille gravitationnelle qu'il subit. L'instrument PLANCK et ses détecteurs sont présentés dans le chapitre 4. Le chapitre 5 détaille l'algorithme de compression des données de l'instrument HFI à bord du satellite et présente ma contribution au processus d'optimisation des paramètres de compression ainsi que le bilan du suivi régulier des erreurs de compression et du débit de données dont j'étais responsable et que j'ai réalisé depuis le début de la prise de données. Le chapitre 6 est consacré quant à lui, à l'ensemble des outils généraux que j'ai utilisés, adaptés ou développés afin de réaliser des simulations et pour estimer le spectre de puissance de température et de la déflexion sur des petites portions du ciel. Dans chaque cas l'approche par morceaux et celle sur la sphère sont présentées, en vue d'être appliquées à l'étude de la mesure de l'effet de lentille gravitationnelle par la méthode d'analyse du ciel par morceaux. Le chapitre 7 présente mon travail de développement et de caractérisation de la méthode d'analyse du ciel par morceaux. Enfin, l'étude de l'effet de lentille gravitationnelle que j'ai réalisé à partir des données de l'instrument HFI de PLANCK, conduisant à la première mesure du spectre de puissance de la déflexion est décrite dans le chapitre 8.

Première partie

**Bases théoriques**



# Chapitre 1

## Cosmologie

Les débuts de la cosmologie moderne remontent à la découverte par Hubble, en 1929, de la récession apparente des galaxies. La mise en évidence de ces effets dynamiques a profondément modifié la vision de l'Univers qu'on avait à cette époque et a conduit au développement du modèle du *Big Bang*. La dynamique globale de l'Univers est dominée par la gravitation car c'est la seule interaction à longue portée. En effet, les forces nucléaires faible et forte agissent uniquement à très faible distance et la force électromagnétique est également négligeable à grande échelle car l'Univers est globalement neutre ce qui produit un effet d'écrantage. Le cadre théorique du *Big Bang* s'appuie donc sur la Relativité Générale développée dans les années précédentes. Cette théorie a été construite à partir des bases de la relativité restreinte et généralise la dynamique de l'espace-temps en reliant le contenu en matière et en énergie de l'Univers à sa géométrie. C'est la théorie la plus générale de la gravitation.

Le rôle de la cosmologie est d'apporter une description physique de l'Univers. Cette discipline ne s'intéresse pas directement aux objets astrophysiques eux mêmes mais à la structure générale dans laquelle ils évoluent et aux lois qui régissent leur dynamique à grande échelle. Comparée à d'autres domaines scientifiques, la cosmologie possède certaines particularités qui ont influencé son développement et les démarches mises en oeuvre pour son développement. En effet, on ne peut observer qu'un seul Univers et on ne peut pas se déplacer sur des échelles cosmologiques, ce qui nous limite à un seul point de vue.

Les avancées continues dans les développements théoriques et la qualité des observations ont permis d'obtenir un modèle du *Big Bang*, capable de décrire les grandes lignes de l'évolution de l'Univers depuis ses premiers instants, jusqu'à nos jours. Il est à l'origine de deux prédictions majeures qui ont conduit à sa large adoption et à l'abandon des principaux modèles concurrents : le fond diffus cosmologique issu du découplage matière-rayonnement et l'abondance relative des éléments légers de l'Univers, produits au cours de la nucléosynthèse primordiale. Ce modèle s'appuie sur un petit nombre de paramètres dont la mesure constitue un des enjeux majeurs de la cosmologie observationnelle contemporaine. D'importantes questions restent toutefois sans réponse, en particulier, sur le processus d'inflation, la nature de l'énergie noire et de la matière noire.

### 1.1 Piliers du modèle

#### 1.1.1 Principe cosmologique

Dès le début de son développement, la cosmologie moderne s'est trouvée face à la complexité du problème à résoudre. Les bases théoriques de la Relativité Générale étaient difficilement applicables au modèle d'Univers le plus général. Les développements se sont alors appuyés sur des hypothèses de symétrie afin de produire des modèles simplifiés pour lesquels la résolution des équations était facilitée. C'est ainsi que le *principe cosmologique* a rapidement été introduit afin de simplifier les développements. Il suppose qu'à suffisamment grande échelle, l'Univers est spatialement homogène et isotrope. C'est-à-dire que ses propriétés sont identiques quel que soit le point de l'espace considéré et quelle que soit la direction d'observation. Il en résulte que les lois de la physique doivent être les mêmes en tout point de l'Univers.



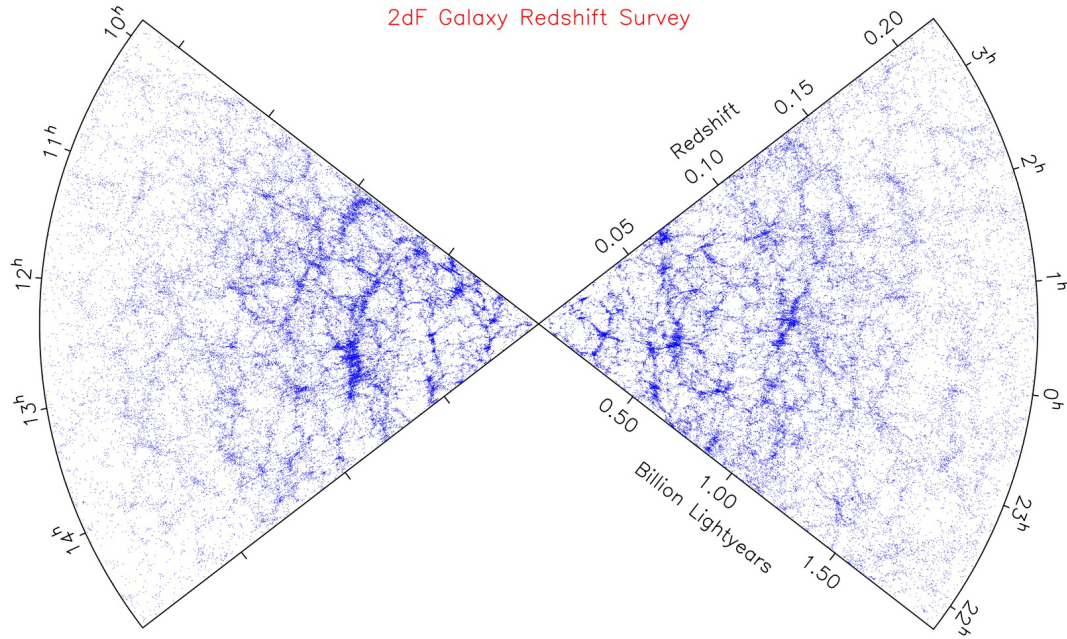


FIGURE 1.1 – **Relevé de galaxies 2dF.** La position des galaxies en fonction de leur décalage vers le rouge montre clairement une isotropie à grande échelle. À plus petite échelle, on voit apparaître des structures en forme de filaments. (voir Peacock et al. (2001))

L'hypothèse d'isotropie est aujourd'hui bien vérifiée par les observations mais ce n'était pas le cas au début du vingtième siècle lorsque le principe cosmologique a été introduit. En effet, il est bon de rappeler qu'à cette époque, la distribution de matière dans l'Univers était très mal connue. Ce n'est que dans les années 1920 qu'il a été montré qu'une partie des *nébuleuses* était en fait des objets extragalactiques. Auparavant, l'espace connu se résumait à l'environnement galactique qui nous apparaît fortement anisotrope. En plus de faciliter les calculs, le principe cosmologique a donc également été formulé par soucis d'esthétique. Il peut être vu comme une conséquence, ou une mise en pratique, du *principe de Copernic* qui suppose qu'il n'existe pas de points de vue privilégiés dans l'Univers.

Deux observables importantes renforcent la validité du principe cosmologique : la distribution spatiale à grande échelle des galaxies et le CMB. Les relevés de galaxies lointaines en fonction de leur décalage spectral vers le rouge, effectués par la collaboration 2dF par exemple, montrent qu'à grande échelle, la distribution des galaxies devient de plus en plus isotrope (voir figure 1.1). Les observations du CMB ont prouvé que sa température était remarquablement isotrope sur tout le ciel. Les premières variations relatives primordiales de température, par rapport à la valeur moyenne n'ont été découvertes qu'en 1992 et leur amplitude est d'environ  $5 \times 10^{-5}$ . Ces observations montrent que de façon statistique, la distribution de la matière à grande échelle est fortement isotrope. En utilisant le principe de Copernic qui suppose que notre position dans l'Univers n'a rien de particulier et que nous ne sommes pas privilégiés par rapport à un autre observateur éventuel, on peut en déduire que l'isotropie observée depuis la Terre implique l'homogénéité de l'Univers.

Finalement, on peut noter qu'il existe une version plus forte du principe cosmologique appelé *principe cosmologique parfait*. En plus de l'homogénéité spatiale, cette hypothèse impose que l'Univers soit identique quelle que soit l'époque considérée. Il est à la base du développement du modèle d'Univers stationnaire notamment, par Fred Hoyle. Ce modèle nécessite, en outre, la création continue de matière pour compenser la dilution produite par l'expansion de l'Univers. Il a été abandonné dans les années 1960 devant la difficulté rencontrée pour expliquer les observations de l'abondance relative des éléments légers et du fond diffus cosmologique.

### 1.1.2 Piliers observationnels du modèle du *Big Bang*

Le modèle du *Big Bang* s'est construit petit à petit au cours des 100 dernières années. Il se base sur le principe cosmologique et repose sur trois observations fondamentales qui ont conduit à son développement et son acceptation.

- **L'expansion de l'Univers.** Cette observation est certainement celle qui a eu le plus gros impact en conduisant à une modification profonde de la vision de l'Univers qui était souvent considéré comme statique avant cette découverte. Il a été observé dès 1918 (Slipher (1918)) un décalage systématique vers le rouge des spectres des galaxies. Si  $\lambda_S$  désigne la longueur du rayonnement émis par une source et  $\lambda_O$  la longueur d'onde mesurée par un observateur. On définit le décalage vers le rouge  $z$  comme :

$$z = \frac{\lambda_O - \lambda_S}{\lambda_S} \quad (1.1)$$

On peut interpréter ce décalage vers le rouge (*redshift* en anglais) comme un effet Doppler, c'est-à-dire comme le changement de longueur d'onde d'un rayonnement induit par un mouvement relatif entre la source et l'observateur. Si la vitesse de la source par rapport à l'observateur est  $v$ , on a  $v = cz$ . Ainsi, la mesure du décalage vers le rouge des galaxies donne accès à leur vitesse radiale. Par ailleurs, il est possible de mesurer la distance des galaxies grâce à l'observation de céphéides. Ces étoiles variables sont très lumineuses et peuvent être observées dans des galaxies lointaines. Elles possèdent une relation période-luminosité absolue connue qui permet par la mesure de leur luminosité apparente et de leur période, d'estimer leur distance. En 1929, Hubble a montré que la vitesse d'éloignement des galaxies était proportionnelle à leur distance :

$$v = H_0 d \quad (1.2)$$

où  $H_0$  est la constante de Hubble. L'équation précédente est appelée la loi de Hubble. Si on applique le principe de Copernic, cette relation implique que toutes les galaxies s'éloignent les unes des autres, à une vitesse d'autant plus grande, que la distance qui les sépare est élevée, ce qui s'interprète comme une expansion globale de l'Univers.

- **La nucléosynthèse primordiale.** En tenant compte de la fraction d'éléments créés par les étoiles, les observations tendent à montrer que l'abondance des éléments légers, est relativement constante dans l'Univers (90% d'hydrogène, 10% d'hélium, 0.001% de deutérium et des traces de lithium). Par ailleurs, aucun mécanisme connu ne permet de former de deutérium dans les étoiles. Des mécanismes capables d'expliquer ces observations ont longtemps été recherchés mais aucun n'était capable de reproduire ces abondances relatives. Les avancées en physique nucléaire obtenues dans la première moitié du vingtième siècle ont alors conduit Gamov, Alpher et Herman à construire un modèle de nucléosynthèse primordiale. Ils avançaient alors, que dans le cadre d'un Univers en expansion issue d'une époque très dense et chaude, l'ensemble des éléments de l'Univers s'étaient formés dans les premières minutes, par fusion des éléments légers en éléments lourds. Aujourd'hui, les mesures précises des taux de réaction dans les accélérateurs de particules et le développement du modèle standard, permettent de modéliser plus précisément la première phase de l'Univers et de prédire les abondances relatives des différents éléments qui sont en bon accord qualitatif avec les observations. Ce résultat constitue un élément majeur en faveur de la théorie du *Big Bang*. On peut trouver une revue intéressante du sujet, abordant les aspects théoriques, la simulation et les observations, dans Smith et al. (1993).
- **Le fond diffus cosmologique.** Il a été prédit dans les années 1940 par Gamow, Alpher et Herman au cours de leurs travaux sur la nucléosynthèse primordiale et a été découvert fortuitement par Penzias et Wilson en 1965. La mesure de son spectre par le satellite COBE en 1992 a montré que son rayonnement est celui d'un corps noir parfait ce qui nécessite que l'Univers soit passé par une phase d'équilibre thermodynamique. La température du CMB mesurée à l'époque actuelle de 2.7K est fortement isotrope sur tout le ciel et s'explique par le décalage spectral lié à l'expansion de l'Univers. Ces observations remarquables, renforcent l'idée d'un univers statistiquement homogène et isotrope ainsi que les modèles d'Univers issus d'une phase chaude à l'équilibre thermodynamique donnant naissance à ce rayonnement. On reviendra sur ce sujet dans le chapitre 2.

## 1.2 Notions de Relativité Générale

La gravitation est la force qui domine la dynamique globale de l'Univers à grande échelle. La relativité générale, développée par Einstein en 1915, est la théorie de la gravitation la plus générale. Elle repose sur le *principe d'équivalence* et sur la relativité restreinte et donne une interprétation géométrique de la gravitation. Dans le cas particulier des faibles vitesses et des faibles champs elle se réduit à la théorie de la gravitation de Newton.

Comme en relativité restreinte, les événements de l'espace-temps sont repérés par des quadrivecteurs  $x^\mu = (x^0, x^1, x^2, x^3)$ . L'élément  $x^0 = ct$  est la coordonnée temporelle tandis que les éléments  $x^1, x^2$  et  $x^3$  sont les coordonnées spatiales.

### 1.2.1 Géométrie de l'Univers

Dans le cadre de la relativité générale, la géométrie de l'espace-temps est décrite par le *tenseur métrique*  $g_{\mu\nu}$ , souvent appelé également la métrique. Le tenseur métrique permet de généraliser aux espace-temps courbes, la définition de l'invariant relativiste  $ds$  qui permet de mesurer l'*intervalle* entre deux événements de l'espace-temps. Dans le cas général, on peut l'écrire sous la forme :

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (1.3)$$

On peut en déduire la trajectoire d'une particule libre, qui va suivre le chemin tel que l'intégrale sur ce chemin soit stationnaire :

$$\delta \int_{\text{chemin}} ds = 0 \quad (1.4)$$

Ce chemin est appelé une *géodésique*. Dans un espace euclidien, les géodésiques sont des lignes droites et une particule libre voyage donc en ligne droite. On peut montrer que pour une particule libre, les géodésiques obéissent à l'équation :

$$\frac{d^2 x^\alpha}{ds^2} + \Gamma_{\beta\gamma}^\alpha \frac{dx^\beta}{ds} \frac{dx^\gamma}{ds} = 0 \quad (1.5)$$

où  $\Gamma_{\beta\gamma}^\alpha$  est appelé symbole de Christoffel. Il dépend uniquement de la métrique et il est donné par :

$$\Gamma_{\beta\gamma}^\alpha = \frac{1}{2} g^{\alpha\beta} \left( \frac{\partial g_{\rho\beta}}{\partial x^\gamma} + \frac{\partial g_{\rho\gamma}}{\partial x^\beta} - \frac{\partial g_{\beta\gamma}}{\partial x^\rho} \right) \quad (1.6)$$

À partir du symbole de Christoffel on définit le tenseur de courbure de Riemann :

$$R^\sigma{}_{\mu\nu\kappa} = \frac{\partial \Gamma^\sigma{}_{\mu\nu}}{\partial x^\kappa} - \frac{\partial \Gamma^\sigma{}_{\mu\kappa}}{\partial x^\nu} + \Gamma^\lambda{}_{\mu\nu} \Gamma^\sigma{}_{\kappa\lambda} - \Gamma^\lambda{}_{\mu\kappa} \Gamma^\sigma{}_{\nu\lambda} \quad (1.7)$$

à partir duquel on peut définir le tenseur de Ricci :

$$R_{\alpha\beta} = R^\mu{}_{\alpha\mu\beta} \quad (1.8)$$

qui donne le scalaire de Ricci :

$$R = g^{\mu\nu} R_{\mu\nu} \quad (1.9)$$

Cette quantité s'interprète géométriquement comme la courbure locale de l'espace. Finalement, à partir des éléments précédents, on définit le tenseur d'Einstein par la relation suivante :

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \quad (1.10)$$

Ce dernier contient toute l'information géométrique de l'espace-temps.

### 1.2.2 Équation d'Einstein

L'équation d'Einstein, aussi appelée équation de champ d'Einstein, est l'équation fondamentale de la relativité générale. Elle décrit comment le contenu en matière et en énergie de l'Univers modifie la géométrie de l'espace temps. C'est une équation aux dérivées partielles qui gouverne la dynamique

gravitationnelle et assimile la courbure locale de l'espace, générée par la matière, à l'effet d'un champ gravitationnel.

Le contenu de l'univers est entièrement décrit par le tenseur énergie-impulsion  $T_{\mu\nu}$ . L'élément  $T_{00}$  correspond à la densité d'énergie et les autres termes diagonaux aux composantes de l'impulsion. La conservation de l'énergie impose  $T_{\mu;\nu}^{\nu} = 0$ , où le terme  $;\nu$  désigne la dérivée covariante par rapport à  $\nu$ . Einstein a cherché une généralisation de l'équation de Poisson de la physique Newtonienne :

$$\nabla^2\varphi = 4\pi G\rho \quad (1.11)$$

Il a donc cherché un tenseur construit uniquement à partir de la métrique qui permette de retrouver l'équation précédente dans la limite Newtonienne. Ce tenseur doit donc faire intervenir des dérivées d'ordre second et correspond au tenseur d'Einstein décrit dans la section précédente. Einstein a montré que ce tenseur était aussi à dérivée covariante nulle. Il en a déduit l'équation suivante :

$$G_{\mu\nu} = \frac{8\pi G}{c^2}T_{\mu\nu} \quad (1.12)$$

où  $G_{\mu\nu}$  est le tenseur d'Einstein,  $G$  la constante gravitationnelle et  $T_{\mu\nu}$  le tenseur énergie-impulsion. C'est cette équation qu'on appelle généralement l'équation d'Einstein. La constante  $\frac{8\pi G}{c^2}$  assure de retrouver l'équation de Poisson dans la limite des champs faibles. La forme la plus générale contient en plus une constante d'intégration et peut s'écrire :

$$G_{\mu\nu} = 8\pi GT_{\mu\nu} + \Lambda g_{\mu\nu} \quad (1.13)$$

Comme on l'a vu dans la section précédente, le tenseur  $G_{\mu\nu}$  ne dépend que de la métrique et de ses dérivées première et seconde. Ainsi, le membre de droite correspond au contenu physique de l'Univers tandis que le terme de gauche représente les aspects géométriques. Cette équation décrit donc comment la géométrie de l'espace-temps et son contenu en matière interagissent et correspondent donc à la loi de la gravitation dans le cadre de la relativité générale.

La constante  $\Lambda$  est appelée *constante cosmologique*. Elle a été introduite à l'origine par Einstein pour s'assurer qu'une solution d'univers statique existe. On trouve parfois le terme  $\Lambda g_{\mu\nu}$  dans le membre de gauche ou dans le membre de droite, suivant l'interprétation qu'on fait de la constante cosmologique en terme d'énergie ou de courbure.

Finalement, on peut écrire cette équation tensorielle sous la forme d'un jeu d'équations aux dérivées partielles, non linéaires. Le nombre d'équation se limite à 10 car les tenseurs  $T_{\mu\nu}$  et  $G_{\mu\nu}$  sont symétriques. Les solutions de l'équation d'Einstein sont les composantes du tenseur métrique. Connaissant le tenseur métrique, on peut alors calculer les trajectoires des particules ou du rayonnement par exemple étant donnée la géométrie de l'espace temps résultant. Elles suivent les géodésiques qui peuvent être calculées à partir de l'équation des géodésiques.

## 1.3 Modèle de Friedmann–Lemaître–Robertson–Walker

D'après le *principe cosmologique*, l'Univers est *spatialement* homogène et isotrope. Ainsi, pour un temps donné, les propriétés mesurées doivent être identiques dans toutes les positions de l'Univers et dans toutes les directions. Ces hypothèses sont bien vérifiées statistiquement par les observations, sur des échelles supérieures à 100Mpc. La vérification statistique n'implique pas que l'évolution de l'Univers soit totalement décrite par un modèle parfaitement isotrope mais pousse tout de même à utiliser ce choix comme première approximation et à traiter les différences comme des petites perturbations. C'est la démarche qui est généralement suivie en cosmologie.

Les travaux menés de façon indépendante par Friedmann et Lemaître sur les solutions des équations d'Einstein dans le cadre d'un Univers parfaitement homogène et ceux de Robertson et Walker sur la forme générale de la métrique dans ces mêmes conditions, constituent la base du formalisme nécessaire pour décrire un Univers sous l'hypothèse du principe cosmologique. On appelle généralement ce modèle, le modèle de Friedmann–Lemaître–Robertson–Walker, ou FLRW en abrégé.

### 1.3.1 Métrique de Robertson–Walker

L'homogénéité et l'isotropie spatiales impliquent que la courbure ne doit pas dépendre de l'espace mais rien ne l'empêche d'avoir une dépendance temporelle. Dans ces conditions, la métrique la plus

générale satisfaisant les symétries imposées, s'écrit en coordonnées sphériques :

$$ds^2 = (cdt)^2 - a(t)^2 \left[ \frac{dr^2}{1 - Kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right] \quad (1.14)$$

où  $t$  est le temps propre et  $c$  la vitesse de la lumière dans le vide. Le terme  $a(t)$  est appelé le *facteur d'échelle*. C'est la fonction qui contient l'ensemble des évolutions temporelles de la partie spatiale de la métrique. Les hypothèses d'homogénéité et d'isotropie spatiale n'imposent pas de contraintes particulières à cette fonction. À ce stade  $a(t)$  représente donc une fonction arbitraire du temps. Il faut résoudre l'équation de champ d'Einstein pour déterminer son évolution.

Le terme  $K$  est lié à la courbure de l'espace. S'il vaut 0, à  $t$  constant, la partie spatiale de la métrique s'identifie à celle d'un espace Euclidien classique avec la distance radiale donnée par  $a(t)r$ . On dit que l'Univers est plat, ou plus exactement *spatialement plat* car l'espace-temps, lui, reste un espace courbe.

Si  $K$  est différent de 0,  $a$  étant une fonction arbitraire, à ce stade, il est toujours possible de transformer la variable  $r$  par un simple facteur multiplicatif, qu'on incorpore à la fonction  $a$  afin d'obtenir une valeur pour  $K$  de  $\pm 1$ . Dans ce cas on dénote alors généralement  $K$  par la variable  $\kappa$ . On dit que l'Univers est fermé dans le cas où  $\kappa = 1$  et qu'il est ouvert si  $\kappa = -1$ . La variable  $r$  est sans unité et le facteur d'échelle a la dimension d'une distance.

### 1.3.2 Équations de Friedmann–Lemaître

Le principe cosmologique permet de déduire une forme générale de la métrique. Cependant, il ne donne aucune contrainte sur le paramètre d'échelle  $a(t)$ . Pour cela, il faut résoudre l'équation d'Einstein, ce qui nécessite de connaître la forme du tenseur énergie-impulsion. Sous les hypothèses du principe cosmologique et en supposant que le contenu de l'Univers peut être représenté par un fluide parfait de pression  $p$  et de densité d'énergie  $\rho$ , ce tenseur prend la forme simple suivante :

$$T_{\mu\nu} = (p + \rho c^2)u_\mu u_\nu - pg_{\mu\nu} \quad (1.15)$$

où,  $c$  est la vitesse de la lumière et  $u_i$  le quadrivecteur vitesse du fluide donné par :

$$u_\mu = g_{\mu\nu}u^\nu = g_{\mu\nu} \frac{dx^\nu}{ds} \quad (1.16)$$

Le tenseur énergie-impulsion étant de rang deux, on peut l'écrire sous forme matricielle :

$$T_{\mu\nu} = \begin{pmatrix} \rho(t) & 0 & 0 & 0 \\ 0 & -p(t) & 0 & 0 \\ 0 & 0 & -p(t) & 0 \\ 0 & 0 & 0 & -p(t) \end{pmatrix} \quad (1.17)$$

L'équation d'Einstein se réduit alors au jeu d'équations suivant :

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \quad (1.18)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \quad (1.19)$$

avec  $\dot{a} = \frac{da}{dt}$ . Ces équations sont appelées équations de Friedmann–Lemaître et décrivent l'expansion de l'espace dans le cadre de la relativité générale pour un modèle d'Univers spatialement homogène et isotrope dans lequel évoluent des fluides parfaits. Le rapport  $\frac{\dot{a}}{a}$  correspond au taux d'expansion de l'Univers :

$$H = \frac{\dot{a}}{a} \quad (1.20)$$

qui est une fonction du temps  $H = H(t)$ . Sa valeur actuelle est la constante de Hubble qui est généralement notée  $H_0$ . De façon plus générale, par convention, on trouvera toujours un indice 0 pour désigner la valeur des paramètres mesurés à l'époque actuelle.

En combinant les équations 1.18 et 1.19 on peut en déduire la loi de conservation de l'énergie dans un Univers en expansion :

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0 \quad (1.21)$$

On peut réécrire l'équation 1.18 sous la forme suivante :

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}(\rho + \rho_{courbure} + \rho_\Lambda) \quad (1.22)$$

avec

$$\rho_{courbure} = -\frac{3\kappa}{8\pi G a^2} \quad (1.23)$$

et

$$\rho_\Lambda = \frac{\Lambda}{8\pi G} \quad (1.24)$$

où  $\rho_\Lambda$  est la densité d'énergie associée à la constante cosmologique aussi appelée densité d'énergie du vide et  $\rho_{courbure}$  la densité d'énergie associée à la courbure.

L'équation 1.22 permet de définir une densité critique qui annule la courbure ( $\kappa = 0$ ) :

$$\rho_c = \left(\frac{\dot{a}}{a}\right)^2 \frac{3}{8\pi G} \quad (1.25)$$

à partir de laquelle on définit des paramètres de densité pour chaque espèce  $i$ , présente dans l'Univers :

$$\Omega_i = \frac{\rho_i}{\rho_c} \quad (1.26)$$

avec

$$\Omega_{tot} = \sum_i \Omega_i \quad (1.27)$$

on en déduit trois régimes :

- si  $\Omega_{tot} > 1$  alors  $\kappa = 1$ . On a un Univers fermé.
- si  $\Omega_{tot} = 1$  alors  $\kappa = 0$ . On est alors dans le cas d'un Univers spatialement plat.
- si  $\Omega_{tot} < 1$  alors  $\kappa = -1$ . L'Univers est ouvert.

Les mesures récentes du CMB donne une valeur de  $\Omega_{tot}$  très proche de 1 avec une erreur de l'ordre du pourcent (voir Komatsu et al. (2011)).

Seules deux des trois équations 1.18, 1.19 et 1.21 sont indépendantes. Il n'est donc pas possible de déterminer l'évolution des trois paramètres inconnus  $\rho(t)$ ,  $p(t)$  et  $a(t)$ . Pour cela il faut en plus utiliser une équation d'état qui relie  $\rho(t)$  à  $p(t)$ . Dans le cas général  $\rho(t)$  et  $p(t)$  correspondent respectivement à la somme des densités d'énergie et de la pression des différents fluides de l'Univers. Lorsque la densité d'énergie d'un des fluides domine, on peut évaluer l'évolution de l'Univers dans une première approximation, en supposant que l'équation d'état est gouvernée par celle du fluide qui domine. On distingue alors plusieurs cas intéressants :

- **domination du rayonnement.** Les grandeurs relativistes, comme le rayonnement, possèdent une équation d'état de la forme  $p_R = \rho/3$ . En résolvant l'équation de conservation de l'énergie on trouve alors facilement que  $\rho_R \propto a^{-4}$ . L'équation 1.18 donne finalement  $a(t) \propto t^{1/2}$ .
- **domination de la matière non relativiste.** La matière non relativiste possède une pression négligeable. Dans ce cas, on trouve que  $\rho_{NR} \propto a^{-3}$ , ce qui donne  $a(t) \propto t^{2/3}$ .
- **domination de la constante cosmologique.** L'équation 1.18 se réduit à  $\left(\frac{\dot{a}}{a}\right)^2 = \frac{\Lambda}{3}$ . Après intégration on trouve  $a(t) \propto e^{\sqrt{\Lambda/3}t}$ , ce qui correspond à une croissance exponentielle de l'Univers. La densité d'énergie du vide donnée par l'équation 1.24 est constante. On peut alors déduire de l'équation de conservation de l'énergie (1.21), l'équation d'état pour un Univers dominé par la constante cosmologique :  $p = -\rho$ . Le vide possède donc une pression négative.

Ces résultats permettent d'ébaucher l'évolution globale du facteur d'échelle pour les différents régimes rencontrés au cours du temps. Les observations montrent que l'expansion semble s'accélérer ce qui pourrait indiquer qu'on entre dans une aire dominée par la constante cosmologique.

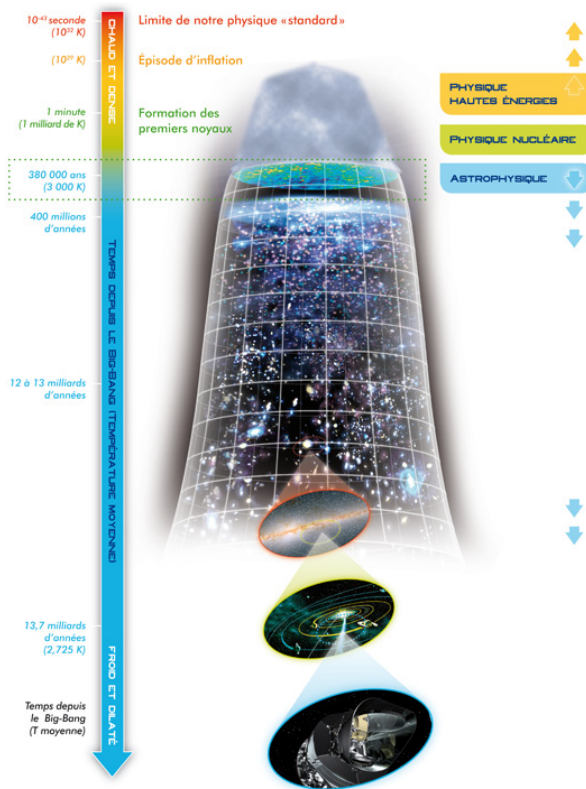


FIGURE 1.2 – **Évolution de l’Univers.** Principales étapes de l’évolution de l’Univers des premiers instants jusqu’à nos jours. (Crédit image : Canopée / Planck-HFI Consortium)

## 1.4 Le *Big Bang*

Les observations et les développements théoriques décrits précédemment ont permis de dégager un scénario d’évolution de l’Univers, auquel on donne généralement le nom de modèle du *Big Bang* chaud, ou modèle standard cosmologique. Dans ce modèle, l’univers est décrit comme un espace-temps à quatre dimensions en expansion, dans lequel évoluent des fluides parfaits. L’espace est supposé parfaitement homogène et isotrope et l’évolution de l’Univers est gouvernée par les équations de Friedmann–Lemaître.

### 1.4.1 Évolution du contenu de l’Univers

Dans le scénario du *Big Bang*, le contenu de l’Univers évolue selon les lois de la physique à partir d’une singularité initiale. Nos connaissances physiques commencent au temps de Planck  $t_p = \sqrt{\frac{\hbar G}{c^5}} \sim 10^{-44} s$ . Au-delà de cette limite les conditions de température et de densité sont si extrêmes que l’ensemble des interactions fondamentales doivent être unifiées pour décrire l’état de l’Univers. Aucune théorie n’existe à l’heure actuelle pour décrire la matière dans de telles conditions. La densité d’énergie étant si élevée que la Relativité Générale sort de son champ d’application. Nous donnons dans la suite les étapes essentielles de l’évolution de l’Univers depuis l’état initial supposé extrêmement dense et chaud. Ces étapes sont également résumées sur la figure 1.2

- **brisure électrofaible** ( $t \sim 10^{-12} s$ ,  $T \sim 10^{16} K$ ). Le mécanisme de Higgs permet de briser la symétrie  $SU(2) \times U(1)$ . La force faible et la force électromagnétiques se différencient. C’est la transition de phase électrofaible.
- **Transition quarks-hadrons** ( $t \sim 10^{-5} s$ ,  $T \sim 10^{13} K$ ). La température ne permet plus aux quarks de rester libres dans le plasma quark-gluons. Ils se combinent pour former les premiers hadrons (notamment les protons et les neutrons).
- **Nucléosynthèse primordiale** ( $t \sim 10^{-2} - 10^2 s$ ,  $T \sim 10^{11} - 10^9 K$ ). La température devient suffisamment basse pour permettre aux premiers noyaux atomiques de se former, en com-

mençant par les plus légers, le deutérium puis le tritium, l'hélium 3 et 4 et enfin le lithium 6 et 7. Les abondances relatives (90%H, 9%He...) prédites, sont en bon accord avec les observations et représentent un des plus grand succès de la théorie du *Big Bang*. La proportion des autres éléments formés est négligeable.

- **découplage des neutrinos.** À la même époque que la nucléosynthèse primordiale, les neutrinos se découplent du reste de la matière, ce qui produit un fond de neutrinos. Sa température actuelle est de 1.4K. (Mal)heureusement la section efficace des neutrinos cosmologiques avec la matière est trop faible et il n'existe, à l'heure actuelle, aucun détecteur suffisamment sensible pour être capable de détecter ces neutrinos.
- **égalité matière-rayonnement** ( $t \sim 55000$  ans,  $T \sim 65000$ K). Avant cette époque le contenu énergétique de l'Univers était dominé par le rayonnement. L'expansion de l'Univers suivait alors une loi en  $a(t) \propto t^{1/2}$ . L'égalité correspond au moment où la densité énergétique de la matière égale celle du rayonnement. À partir de cette époque le facteur d'échelle va varier comme  $a(t) \propto t^{2/3}$ .
- **découplage matière-rayonnement** ( $t \sim 380000$  ans,  $T \sim 3000$ K). Avant cette époque l'Univers était constitué d'un plasma de protons électrons et photons interagissant par diffusion Thomson. La section efficace d'interaction des photons avec un atome neutre étant très faible, lorsque la température est devenue suffisamment basse pour que les électrons puissent se lier aux protons et former ainsi les premiers atomes neutres, l'Univers est alors devenu rapidement transparent. Les photons ainsi émis constituent un rayonnement fossile directement observable à l'heure actuelle, appelé le fond diffus cosmologique (ou CMB en anglais). Cette époque est souvent appelée *recombinaison*. Ce nom peut cependant prêter à confusion car c'est en réalité la première fois dans l'histoire de l'Univers que les électrons et les protons se combinent. On reviendra plus en détail sur le CMB dans le chapitre 2.
- **formation des grandes structures** ( $t \sim 10^9$  ans,  $T \sim 15$ K). Les premières galaxies se forment par effondrement gravitationnelle à partir des surdensités primordiales.
- **aujourd'hui** ( $t \sim 13.7 \times 10^9$  ans). L'univers est inhomogène à petite échelle. On observe des structures comme les étoiles, les galaxies, les amas... Le fond diffus cosmologique a une température de 2.725 K.

### 1.4.2 Limites du modèle classique du *Big Bang*

Le modèle de FLRW est un modèle simplifié. Il est capable de prédire certaines observations mais n'est pas capable de répondre à toutes les questions. Nous donnons dans la suite quelques unes de ses limites les plus importantes :

- **Problème de l'horizon.** L'ensemble des points pouvant avoir envoyé un signal lumineux jusqu'à un point particulier  $O$  situé à l'origine d'un repère, dans un temps donné  $t$ , est contenu dans une sphère de rayon  $R_P(t)$  donné par :

$$R_P(t) = a(t) \int_0^t \frac{cdt'}{a(t')} \quad (1.28)$$

cette distance est appelée *horizon de la particule*. Les points étant situés au-delà de cet horizon n'ont pas pu être en contact causal avec l'origine du repère.

Les photons du CMB ont été émis à une époque appelée recombinaison. Depuis, ils se propagent quasiment librement. La période de la recombinaison est relativement courte et les photons du CMB observés à l'heure actuelle semblent donc provenir d'une même surface appelée surface de dernière diffusion (voir le chapitre 2 pour plus de détails). En intégrant l'équation précédente depuis l'origine de l'Univers jusqu'à l'époque de la recombinaison et en employant les relations donnant  $a(t)$  en fonction du temps aux différentes époques (voir section 1.3.2), on peut calculer la taille de la sphère définissant l'horizon des particules. Dans ce cas, le rayon de la sphère est généralement appelé le rayon de Hubble. À l'époque actuelle, la taille angulaire de l'horizon des particules au niveau de la surface de dernière diffusion est vue sous un angle d'environ  $1^\circ$  (voir par exemple Rich (2004)). Ainsi, deux points ayant été en contact causal aux premiers instants de l'Univers ne peuvent pas être séparés d'une distance angulaire de plus de  $1^\circ$  degré au niveau de la surface de dernière diffusion. Or, on observe que le CMB possède une température remarquablement isotrope sur l'ensemble du ciel, ce qui nécessite



que l'ensemble des points ait été en contact causal. Ce problème est connu sous le nom du *Problème de l'horizon*

- **Problème de la platitude.** On a vu dans la section 1.3.1 que la courbure de l'Univers est un paramètre du modèle. Les mesures montrent que sa valeur actuelle est très proche de 1. La densité énergétique de l'Univers est donc très proche de la densité critique ( $\Omega_T = 1$ ). On peut montrer à partir des équations de Friedmann–Lemaître, que si  $\Omega_T$  vaut 1 à un instant donné alors elle conserve cette valeur. Dans le cas contraire, elle ne peut que s'éloigner de cette valeur au cours du temps. L'incertitude sur la mesure de  $\Omega_T$  nous donne une limite de la déviation par rapport à la densité critique. On peut en déduire une limite sur la valeur initiale de  $\Omega_T$  nécessaire pour que sa valeur actuelle ne se soit pas trop éloignée de la densité critique afin d'être compatible avec les mesures. Les calculs montrent que cela nécessite  $|\Omega_T - 1| \sim 10^{-60}$ . Il n'est pas satisfaisant d'avoir un si faible intervalle de variation autorisé pour les conditions initiales. Ce problème est appelé le *problème de platitude*.
- **Problème de l'origine des fluctuations.** On observe à l'heure actuelle que l'Univers est inhomogène à petite échelle. Dans le cadre du modèle du *Big Bang*, on pense que ces inhomogénéités sont issues de la croissance de fluctuations primordiales, par effondrement gravitationnel, pendant l'expansion. Le CMB présente de faibles anisotropies qui résultent de fluctuations de densité au moment de la recombinaison. Dans le modèle classique du *Big Bang*, aucun mécanisme ne permet de générer ces fluctuations. Elles font donc partie des conditions initiales.

### 1.4.3 Inflation

Les problèmes présentés ci-dessus peuvent être résolus en introduisant une phase primordiale et transitoire d'expansion accélérée de l'Univers généralement appelée *inflation*. Par définition, pendant cette phase le facteur d'échelle  $a(t)$  vérifie :

$$\ddot{a}(t) > 0 \quad (1.29)$$

dans ces conditions, le produit  $a(t)H(t)$  est donc une fonction croissante. On peut réécrire l'équation 1.18 sous la forme :

$$\Omega_T - 1 = \frac{\kappa}{a^2 H^2} \quad (1.30)$$

Ainsi, la densité de l'Univers tend vers la densité critique. L'inflation apporte donc un mécanisme permettant de résoudre le problème de la platitude. Par ailleurs, l'expansion accélérée permet à l'Univers de croître à une vitesse apparente plus grande que la vitesse de la lumière ce qui permet à des zones en contact causal de sortir de l'horizon. On peut alors observer après la période d'inflation des régions du ciel qui sont hors de leur horizon respectifs mais qui ont été en contact causal dans la phase initiale de l'Univers, ce qui résout le problème de l'horizon et qui explique comment la température du CMB peut être corrélée sur tout le ciel. Enfin, l'inflation permet de porter à des échelles macroscopiques des fluctuations quantiques primordiales et fournit ainsi un mécanisme de génération des fluctuations de densité.

Dans le modèle le plus simple, l'inflation est provoquée par un champ scalaire  $\varphi$ . La densité d'énergie et la pression de ce champ sont données par :

$$\rho = \frac{1}{2}\dot{\varphi}^2 + V(\varphi) \quad (1.31)$$

et

$$p = \frac{1}{2}\dot{\varphi}^2 - V(\varphi) \quad (1.32)$$

où  $V(\varphi)$  est le potentiel. De nombreux modèles d'inflations ont été imaginés. Le choix du potentiel  $V(\varphi)$  conditionne leurs propriétés, notamment la durée et l'accélération de la phase d'expansion.

## Chapitre 2

# Fond micro-onde cosmologique

### 2.1 Historique et présentation

#### 2.1.1 Historique

L'existence d'un fond de photons d'origine cosmologique a été prédite vers la fin des années 1940 par Gamow et ses collaborateurs Alpher et Herman alors qu'ils travaillaient sur un modèle capable d'expliquer la formation des éléments dans l'Univers. Ils ont émis l'hypothèse que les éléments pouvaient être produits dans la phase primordiale chaude d'un Univers en expansion. Ils en ont déduit que si l'Univers était à l'équilibre thermodynamique à cette époque et qu'il le soit resté au cours de son évolution, il devait exister un rayonnement, vestige de la phase primordiale de l'Univers, produisant un fond observable à une température d'environ 5K de nos jours (Alpher & Herman (1948), Gamow (1948a), Gamow (1948b)).

Le fond diffus cosmologique, ou CMB (pour le terme anglais Cosmic Microwave Background) a par la suite été découvert, par hasard, dans les années 1960, par Penzias et Wilson, deux radio-ingénieurs travaillant pour les laboratoires Bell. Alors qu'ils essayaient de caractériser les différentes sources de bruit d'une antenne, ils ont mis en évidence un bruit dans le domaine radio, fortement isotrope sur tout le ciel et ne possédant pas de variations saisonnières, qu'ils n'arrivaient pas à expliquer par un effet instrumental ni aucune autre source connue. Ils ont publié leurs résultats en 1965 (Penzias & Wilson (1965)), en interprétant leurs observations comme un fond cosmologique décrit dans un article de Dicke, Peebles, Roll et Wilkinson publié parallèlement (Dicke et al. (1965)). Penzias et Willson ont reçu le prix Nobel de physique en 1978, pour leur découverte.

Dans les années 1970, les travaux pionniers de Peebles et Yu et de Sunyaev et Zel'dovich (Peebles & Yu (1970), Sunyaev & Zeldovich (1970)) ont montré pour la première fois, comment des fluctuations de densité de matière de l'Univers primordial pouvaient produire des anisotropies de température, d'amplitude relative de l'ordre de  $10^{-3}K$ . De telles fluctuations de température dans le rayonnement du CMB ont alors été activement recherchées mais n'ont pas été observées lorsque les mesures ont atteint une sensibilité suffisante. C'est cette constatation qui a conduit Peebles en 1982 à imaginer un modèle d'Univers dominé par une grande quantité de matière inconnue interagissant principalement avec la matière baryonique par le biais de l'interaction gravitationnelle (Peebles (1982)). Il montre que cette hypothèse conduit à une amplitude des fluctuations de température du CMB beaucoup plus faible que celle prédite précédemment. Afin d'expliquer la formation des structures à petite échelle cette matière doit être suffisamment froide. Cette hypothèse représente le fondement des modèles d'Univers de type CDM (pour Cold Dark Matter).

C'est le satellite COBE qui a mesuré pour la première fois les anisotropies de température du CMB avec une amplitude de l'ordre de  $10^{-5}K$ , grâce à son instrument DMR. Par ailleurs, à l'aide de son spectromètre FIRAS il a également confirmé avec une grande précision la forme de corps noir du spectre d'émission du CMB. La publication de ces résultats en 1992 (Smoot et al. (1992)), a marqué un tournant dans la cosmologie observationnelle et a été récompensée par le prix Nobel de physique en 2006. Depuis, un grand nombre d'expériences ont amélioré les mesures des anisotropies, dans le but notamment de déterminer leur spectre de puissance. On peut citer entre autres, des expériences au sol comme DASI, ACBAR, ACT ou SPT, des expériences embarquées sous des ballons stratosphériques comme Archéops, BOOMERanG ou MAXIMA ou à bord de satellite

comme WMAP et PLANCK tout récemment.

Le fond diffus cosmologique présente une importance particulière pour la cosmologie. Tout d'abord c'est le plus vieux rayonnement observable dans l'Univers actuel. En effet, avant l'époque de la recombinaison, le libre parcours moyen des photons était très faible et l'Univers était donc opaque. Par ailleurs, les propriétés statistiques des anisotropies du CMB dépendent fortement des conditions physiques de l'Univers primordial et fournissent ainsi des informations précieuses sur son évolution. La grande précision des mesures actuelles permet de poser de fortes contraintes sur les modèles d'Univers et donne également accès à la mesure des paramètres cosmologiques. Par ailleurs, le rayonnement du CMB est polarisé. La mesure des anisotropies de polarisation permet de construire deux observables supplémentaires qui devraient permettre de lever certaines dégénérescences entre les paramètres cosmologiques.

L'étude du CMB fournit des informations précieuses sur le contenu et la géométrie de l'Univers primordial et les conditions physiques qui y régnaient. Au cours de leur parcours, les photons du CMB voyagent pratiquement sans subir aucune perturbation. C'est la condition qui nous permet d'obtenir des informations sur les conditions primordiales. Cependant, ils subissent tout de même un certain nombre d'interactions qui perturbent leurs propriétés. Ces effets, dont l'origine est variée, sont souvent regroupés sous le terme d'anisotropies secondaires. Le rayonnement fossile permet ainsi également de sonder l'Univers dans des zones plus proches de nous par les traces laissées par ces interactions sur les photons du CMB au cours de leur trajet. Parmi celle-ci, on retiendra l'effet Sunyaev Zeldovich et bien entendu l'effet de lentille gravitationnelle sur lequel nous reviendront plus en détail dans le chapitre 3 et dont la mesure a motivé mon travail de thèse.

### 2.1.2 Éléments de physique du CMB

À partir de l'expansion de l'Univers observée par Hubble, on peut déduire que l'Univers devait être plus dense et plus chaud dans une époque plus lointaine. On s'intéresse ici à la phase ayant précédé l'émission du rayonnement du CMB. À cette époque, l'Univers était constitué d'un plasma totalement ionisé dont la dynamique était gouvernée par les interactions entre les protons, les électrons et les photons (les éléments plus lourds étant présents en quantité négligeable). Les photons avaient suffisamment d'énergie pour ioniser les atomes d'hydrogène. La réaction suivante :



dominait donc dans un état d'équilibre thermodynamique.

L'expansion de l'Univers provoque une dilution de son contenu, entraînant également une diminution de la température. Le rayonnement et la matière étant fortement couplés et à l'équilibre thermique, l'expansion de l'Univers entraîne une diminution de l'énergie des photons. Au fur et à mesure que le nombre de photons possédant suffisamment d'énergie pour ioniser les atomes d'hydrogène diminue, les premiers atomes d'hydrogènes stables commencent à se former jusqu'à obtenir un Univers entièrement neutre. Cette phase s'appelle la *recombinaison* et correspond à la transition de phase plasma-matière neutre. À partir de cette époque les photons voyagent librement dans l'Univers. On a un *découplage* entre la matière et le rayonnement.

Il faut une énergie de 13.6eV pour ioniser un atome d'hydrogène, ce qui correspond à une température d'environ 11000K. L'Univers étant très proche de l'équilibre thermique, la distribution des photons suit dans une très bonne approximation, une loi de corps noir de Planck. On peut en déduire que la température moyenne à l'époque de la recombinaison devrait donc être plutôt voisine de 3000K. En effet, étant donné le nombre très élevé de photons comparé au nombre de baryons, la distribution en énergie des photons à cette température, permet l'existence de suffisamment de photons d'énergie supérieure à 13.6eV.

Au cours de la recombinaison les interactions des photons se font principalement par diffusion Thomson (Voir section 2.4.2). La section efficace d'interaction des photons avec des atomes neutres étant très faible, lorsque les premiers atomes d'hydrogène se forment, le libre parcours moyen des photons augmente, jusqu'à devenir plus grand que le rayon de Hubble (rayon de la sphère d'Univers causalement accessible). Après la recombinaison, le libre parcours moyen ne peut qu'augmenter car la densité de protons et d'électrons diminue de façon globale à cause de la dilution provoquée par l'expansion de l'Univers. À partir de cette époque, les photons peuvent donc voyager librement dans l'Univers. Ce rayonnement est observable de nos jours et constitue le fond diffus cosmologique.

La période de la recombinaison est relativement courte. Si on considère qu'elle se termine quand l'ensemble des atomes se sont formés, on peut estimer la durée de cette phase grâce à l'étude de l'évolution de la fraction d'ionisation. Le court temps de cette phase entraîne que les photons observés ont environ tous été émis à la même époque et ont donc parcouru environ la même distance avant de nous parvenir. La surface de laquelle ils proviennent est donc très proche d'une sphère centrée sur l'observateur. Cette surface est souvent appelée surface de dernière diffusion car elle correspond à la position à laquelle les photons ont subi leur dernière diffusion dans l'Univers ionisé. Les observations montrent toutefois que l'Univers a connu une période dite de *réionisation* (voir par exemple Barkana & Loeb (2001)). On suppose que suite à leur effondrement gravitationnel et leur allumage, le rayonnement intense des premières étoiles massives a été capable d'ioniser l'Univers de façon globale. On peut donner des limites supérieures sur l'époque de la réionisation en étudiant le rayonnement émis par les quasars qui sont les objets connus les plus lointains. Cependant la détermination précise de cette époque constitue toujours un défi pour la cosmologie observationnelle.

### 2.1.3 Spectre électromagnétique du CMB

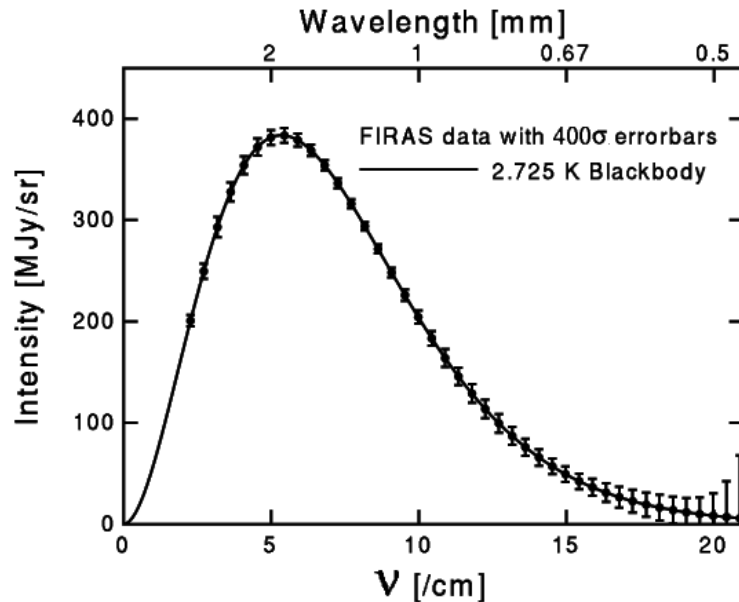


FIGURE 2.1 – **Spectre électromagnétique du CMB** . Les mesures de l'instrument FIRAS de COBE montrent que l'émission du CMB observée à l'époque actuelle est celle d'un corps noir presque parfait à 2.725K. On notera que les barres d'erreurs représentées sur la figure ci-dessus correspondent à la dispersion à 400 $\sigma$ , démontrant le très bon accord avec le spectre théorique d'un corps noir.

Depuis sa découverte en 1965, les observations successives du CMB ont montré que son spectre électromagnétique suivait une loi de corps noir de Planck presque parfaite. Sa luminance énergétique monochromatique est donnée par :

$$I(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1} \quad (2.2)$$

où  $c$  est la vitesse de la lumière,  $h$  la constante de Planck,  $k_B$  la constante de Boltzman et  $T$  la température du corps noir. À l'heure actuelle, la mesure la plus précise a été obtenue par l'instrument FIRAS du satellite COBE qui a effectué des mesures dans la gamme de longueur d'onde allant de la fraction de millimètre, au centimètre. La figure 2.1, présente cette mesure du spectre du rayonnement fossile et montre l'accord remarquable entre les observations et la courbe théorique. On peut montrer que l'expansion de l'Univers ne change pas la forme générale du spectre (voir par exemple Partridge (1995)). Le spectre mesuré par COBE donne donc accès au spectre au niveau de la surface de dernière diffusion avec un décalage vers le rouge à cause de l'expansion de l'Univers.

Le caractère de corps noir presque parfait du CMB nécessite que l'Univers soit passé par une phase d'équilibre thermique mais également que l'évolution de l'Univers depuis cette période jusqu'à aujourd'hui n'ait pas modifié sensiblement la forme du spectre. Le modèle du *Big Bang* satisfait ces deux conditions.

La forme du spectre permet de donner de fortes contraintes sur un éventuel relâchement tardif d'énergie. En effet tout mécanisme capable de produire des photons dans une époque suffisamment tardive pour qu'ils ne puissent pas se thermaliser, provoque des déviations par rapport à un spectre de corps noir. L'ensemble des mesures du spectre du CMB n'a pas permis de mettre en évidence de déviations significatives jusqu'à maintenant.

## 2.2 Anisotropies de température

Les anisotropies de température observées aujourd'hui ont plusieurs origines. On distingue en général les anisotropies primaires, produites par les conditions physiques qui régnaient dans l'Univers à l'époque de la recombinaison, des anisotropies secondaires qui ont pour origine les petites perturbations auxquelles ont été soumis les photons au cours de leur trajet de la surface de dernière diffusion jusqu'à l'observateur. Enfin, on observe également des anisotropies induites par l'effet Doppler dû au mouvement relatif de l'observateur par rapport au référentiel où le CMB est au repos.

### 2.2.1 Dipole

Après soustraction de la moyenne, une carte des anisotropies relatives de température du CMB montre une forte composante dipolaire, largement dominante par rapport à toutes les autres. Son amplitude est de l'ordre de  $10^{-3}$  par rapport à la température moyenne du CMB. Cette anisotropie est due à l'effet Doppler, produit par le mouvement relatif de l'observateur par rapport au référentiel du CMB au repos. On peut montrer que si l'observateur se déplace à une vitesse  $v$  par rapport au référentiel du CMB la température observée est donnée par (Peebles & Wilkinson (1968)) :

$$T(\theta) = T_0 \frac{(1 - \beta^2)^{1/2}}{(1 - \beta \cos \theta)} \quad (2.3)$$

où  $\beta = v/c$ ,  $c$  est la vitesse de la lumière,  $T_0$  est la température du CMB dans le référentiel où il est au repos et  $\theta$  est l'angle entre la direction d'observation et la direction du vecteur vitesse de l'observateur par rapport au référentiel du CMB. Si on développe l'équation précédente au premier ordre en  $\beta$  on obtient finalement :

$$T(\theta) = T_0(1 + \beta \cos \theta) \quad (2.4)$$

qui correspond à l'équation d'un dipôle. La température est donc modulée en fonction de la direction d'observation. Elle apparaît plus importante dans le sens du déplacement de l'observateur (où  $\cos \theta = 1$ ) et plus faible dans la direction opposée au mouvement (où  $\cos \theta = -1$ )

Les observations de COBE puis de WMAP ont permis de déterminer l'amplitude du dipôle :  $3.358 \pm 0.017 \times 10^{-3} K$  ainsi que sa direction, donnée en coordonnées galactiques par  $(l, b) = (263.86^\circ \pm 0.03^\circ, 48.24^\circ \pm 0.10^\circ)$  (voir Hinshaw et al. (2007)). Ces résultats indiquent qu'on se déplace à une vitesse d'environ 370km/s par rapport au CMB. Cette vitesse comprend la vitesse du Soleil dans la Galaxie qui est de l'ordre de 220km/s dans une direction environ opposée à celle du dipôle, la vitesse de la Galaxie dans le groupe local et la vitesse de l'amas local par rapport au CMB (voir par exemple Blanchard (1991)). Sur une échelle de temps humaine, les variations de vitesse dues à notre mouvement dans la Galaxie et au mouvement de notre Galaxie dans l'amas local peuvent être considérées comme constantes. Cependant, notre mouvement de révolution autour du soleil avec une vitesse de l'ordre de 30km/s, entraîne une variation de notre vitesse par rapport au CMB, ce qui produit une modulation du dipôle mesurable. La dynamique du système solaire étant bien connue il est possible de soustraire cette composante de façon précise. On distingue souvent cette composante qu'on appelle dipôle orbital du dipôle global auquel on se réfère sous le nom de dipôle solaire ou dipôle cosmologique.

Si on pousse plus loin le développement de l'équation 2.3, on peut voir que l'effet cinématique ne se limite pas à un dipôle. Cependant, l'amplitude des termes d'ordre supérieur en  $\beta$  devient

vite négligeable car  $\beta$  est de l'ordre de  $10^{-3}$ . On notera tout de même que le terme d'ordre 2 est responsable d'une anisotropie quadrupolaire d'amplitude comparable à celle des anisotropies du CMB.

Lorsqu'on s'intéresse à l'étude des anisotropies du CMB, on travaille en général à partir d'une carte à laquelle on a soustrait le dipôle, ce qui donne accès aux anisotropies du CMB lui-même qui ont une amplitude moyenne de l'ordre de  $10^{-5}$ . Il peut en principe exister un terme dipolaire dans les anisotropies propres au CMB mais il n'est pas possible de le distinguer du dipôle cosmologique.

## 2.2.2 Anisotropies primaires

Les anisotropies primaires correspondent aux anisotropies de température qu'on pourrait observer au niveau de la surface de dernière diffusion. Ces anisotropies gardent donc la trace des conditions physiques qui régnaient dans l'Univers à l'époque de la recombinaison et sont le vecteur principal d'information sur l'époque primordiale. Plusieurs mécanismes sont en concurrence au moment de la recombinaison. Les variations de densité, issues des fluctuations primordiales, grandissent sous l'effet de la gravitation entraînant une attraction locale de la matière et une augmentation de la température ce qui fait augmenter la pression de radiation. Lorsque la pression est suffisamment importante elle s'oppose totalement à l'effondrement. La conjonction de ces deux mécanismes conduit alors à la formation d'oscillations de densité qu'on appelle généralement les oscillations acoustiques. Trois mécanismes principaux, liés aux fluctuations présentes dans l'Univers au moment du découplage des photons du CMB, sont à l'origine des anisotropies primaires du CMB :

- **Perturbation de densité.** Les fluctuations quantiques primordiales sont dilatées jusqu'à des échelles macroscopiques au cours de la période d'inflation. Ces fluctuations de densité intrinsèques produisent des variations locales de température. Les photons du CMB conservent la température qu'ils avaient au niveau de la surface de dernière diffusion, ce qui produit des anisotropies dans leur température apparente sur le ciel. Dans le cas de fluctuations adiabatiques et gaussiennes, les variations sont données par :

$$\frac{\Delta T(\hat{\mathbf{n}})}{T_{CMB}} = \frac{1}{3} \frac{\delta\rho}{\rho}(\hat{\mathbf{n}}) \quad (2.5)$$

où,  $\hat{\mathbf{n}}$  est un vecteur unitaire sur la sphère indiquant la direction d'observation considérée.

- **Effet Sachs-Wolfe.** Lorsqu'un photon entre dans un puits de potentiel, il gagne de l'énergie (on a un décalage vers le bleu du spectre), lorsqu'il en ressort il reperd cette énergie (décalage vers le rouge du spectre). Les variations de densité au niveau de la surface de dernière diffusion conduisent donc à des variations de température du CMB suivant que les photons ont gagné ou perdu de l'énergie dans les puits de potentiel. Cet effet est appelé effet Sachs-Wolfe.
- **Effet Doppler.** Le plasma est animé localement d'une vitesse d'ensemble qui varie en fonction de la position considérée, à cause des oscillations acoustiques présentes dans l'Univers à cette époque. Au niveau de la surface de dernière diffusion, les photons présentent donc une vitesse relative par rapport à l'observateur qui dépend de la direction d'observation. Au moment du découplage, cette empreinte est figée et contribue donc aux variations de température observées sur le ciel, par effet Doppler. On a :

$$\frac{\Delta T(\hat{\mathbf{n}})}{T_{CMB}} \propto -\mathbf{v}(\hat{\mathbf{n}}) \cdot \hat{\mathbf{n}} \quad (2.6)$$

où, comme précédemment,  $\hat{\mathbf{n}}$  est un vecteur unitaire sur la sphère indiquant la direction d'observation considérée. Les zones de sur-densité apparaissent donc plus chaudes car le fluide se dirige vers l'observateur tandis que dans les zones de sous-densité on a l'effet inverse, le fluide semble s'éloigner et la zone apparaît plus froide.

## 2.2.3 Anisotropies secondaires

Les anisotropies secondaires sont produites par les interactions subies par les photons au cours de leur trajet depuis la surface de dernière diffusion jusqu'à l'observateur (voir Aghanim et al. (2008)). On peut distinguer deux types d'effets : ceux produits par interaction gravitationnelle comme l'effet Sach-Wolf intégré ou l'effet de lentille gravitationnelle et ceux produits par l'interaction des photons avec un milieu ionisé comme l'effet Sunyaev-Zeldovich.

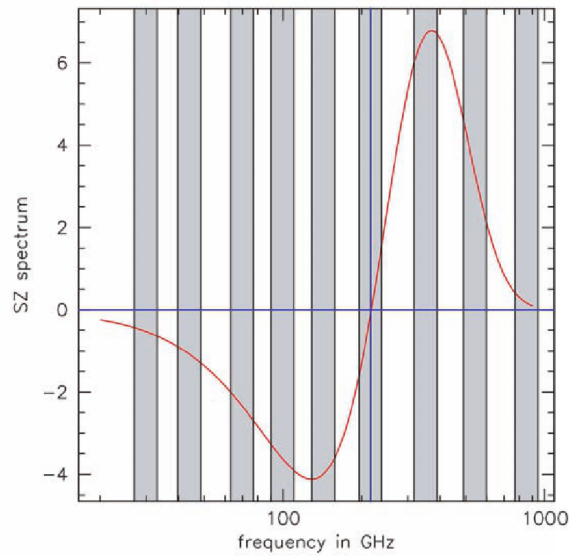


FIGURE 2.2 – **Influence de l'effet Sunyaev-Zel'dovich sur le spectre de puissance du CMB.** Les photons du CMB peuvent être diffusés par les particules chargées du gaz chaud présent dans les amas de galaxies. Ils gagnent alors de l'énergie, ce qui va modifier localement le spectre électromagnétique du CMB comparé au spectre de puissance du CMB moyen, en introduisant un déficit de photons à basse énergie et une augmentation à plus haute énergie. C'est cette différence qui est tracée sur la figure ci-dessus en rouge, avec les bandes de fréquences d'observation des instruments de PLANCK (en gris). Les unités sont arbitraires. (Tiré de Planck Science Team (2005))

- **Effet Sachs-Wolfe intégré.** C'est l'effet causé par effet Sachs-Wolfe par les variations du champ de potentiel gravitationnel le long de la ligne de visée. Si le puits de potentiel se creuse pendant la traversée du photon, l'énergie perdue à la sortie est plus importante que l'énergie qu'il a gagné en entrant dans le puits de potentiel. Le bilan correspond donc à une perte d'énergie.
- **Effet de lentille gravitationnelle.** Les géodésiques des photons sont perturbées par la présence de matière, notamment les grands amas de galaxies. Les photons sont donc déviés de leur trajectoire en fonction de la distribution de matière qu'ils rencontrent au cours de leur trajet. Ces petites déflexions modifient alors l'image apparente de la source et les propriétés statistiques du signal. Le chapitre 3 présente plus en détail cet effet.
- **Effet Sunyaev-Zel'dovich.** Au cours de leur parcours, les photons du CMB vont subir des interactions avec les électrons libres présents dans le gaz chaud des amas de galaxies. Cet effet porte le nom d'effet Sunyaev-Zel'dovich (ou SZ) du nom des auteurs qui l'ont décrit pour la première fois (Zeldovich & Sunyaev (1969)). L'interaction se fait par diffusion Compton inverse ce qui dans une première approximation provoque simplement une augmentation de l'énergie des photons. On observe donc un décalage global du spectre électromagnétique des photons ayant subi une interaction ce qui produit un déficit de photons à basse énergie et une augmentation à plus haute énergie par rapport au spectre moyen du CMB (voir figure 2.2). Cette empreinte caractéristique dans le spectre du CMB permet de détecter les amas de galaxies. La répartition des fréquences du satellite PLANCK a été choisie afin d'optimiser cette détection (voir figure 2.2). On distingue en général l'effet SZ thermique dû à la composante de la vitesse des électrons induite par l'agitation thermique et l'effet SZ cinématique provoqué par la vitesse globale de l'amas de galaxie. L'effet SZ thermique est dominant (voir par exemple Aghanim et al. (2008))

## 2.3 Spectre de puissance angulaire des anisotropies de température

Dans cette section, nous nous intéressons à l'étude des propriétés statistique du CMB.

### 2.3.1 Analyse multipolaire

Comme on l'a vu dans la section 2.1.3, l'émission du CMB est une émission de corps noir presque parfaite. Sa température<sup>1</sup> est quasiment isotrope sur tout le ciel. Seules de faibles anisotropies sont observées. Leur l'amplitude relative par rapport à la température moyenne est de l'ordre de  $10^{-5}$ . On s'intéresse donc en général à ces variations relatives de la température<sup>2</sup> du CMB :  $\Delta T(\theta, \phi)/T_0 = (T(\theta, \phi) - T_0)/T_0$ , où  $T_0$  est la température moyenne sur tout le ciel ( $T_0 = \int_{4\pi} d\Omega T(\theta, \phi)$ ),  $\theta$  et  $\phi$  sont les coordonnées sphériques. Ces variations de température ne dépendent que de la direction d'observation sur le ciel, par conséquent, on peut les décomposer sur la base des *harmoniques sphériques*,  $Y_{\ell m}$  :

$$\frac{\Delta T(\theta, \phi)}{T_0} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{m=+\ell} a_{\ell m} Y_{\ell m}(\theta, \phi) \quad (2.7)$$

Ces fonctions,  $Y_{\ell m}$ , forment une base orthonormée sur la sphère, les  $a_{\ell m}$  sont donc donnés par :

$$a_{\ell m} = \int d\Omega Y_{\ell m}^*(\theta, \phi) \Delta T(\theta, \phi)/T_0 \quad (2.8)$$

De plus comme  $\Delta T(\theta, \phi)/T_0$  prend des valeurs réelles, les  $a_{\ell m}$  vérifient la condition  $a_{\ell m}^* = a_{\ell -m}$ . L'équation 2.8 donne des  $a_{\ell m}$  sans unité. On trouve souvent une définition similaire sans la division par  $T_0$  qui donne alors des  $a_{\ell m}$  dont les unités sont celles d'une température. D'après l'équation 2.7, on voit qu'il y a  $2\ell + 1$  coefficients  $a_{\ell m}$ , pour chaque valeur de  $\ell$  donnée.

Pour une quantité représentant une déviation par rapport à la moyenne, le terme  $a_{00}$  est nul car la fonction  $Y_{00}$  est une constante. Il est alors courant de faire démarrer la somme de l'équation 2.7 à 1, dans ce cas. Par ailleurs, comme on l'a vu dans la section 2.2.1, le terme dipolaire ( $\ell = 1$ ) est dominé par l'effet Doppler lié à notre mouvement par rapport au référentiel du CMB supposé au repos. Il n'est pas possible de distinguer le dipôle d'origine cosmologique présent dans les fluctuations à grande échelle du dipôle produit par effet Doppler. Ainsi, en général, on soustrait le terme dipolaire des cartes et on fait démarrer la somme de l'équation 2.7 à 2.

### 2.3.2 Propriétés des harmoniques sphériques

On liste ici, certaines propriétés pratiques des harmoniques sphériques qui sont souvent employées. Comme les harmoniques sphériques forment une base orthonormée, on peut écrire :

$$\int d\Omega Y_{\ell m}(\theta, \phi) Y_{\ell' m'}^*(\theta, \phi) = \delta_{\ell\ell'} \delta_{mm'} \quad (2.9)$$

En sommant sur tous les  $m$  correspondant à un même multipole  $\ell$ , on trouve la relation suivante :

$$\sum_m |Y_{\ell m}(\theta, \phi)|^2 = \frac{2\ell + 1}{4\pi} \quad (2.10)$$

Enfin, on retiendra l'extension d'une onde plane en fonction des harmoniques sphériques qui s'écrit :

$$e^{i\mathbf{k}\cdot\mathbf{x}} = 4\pi \sum_{\ell m} i^\ell j_\ell(kx) Y_{\ell m}(\hat{\mathbf{x}}) Y_{\ell m}^*(\hat{\mathbf{k}}) \quad (2.11)$$

où, on a introduit les fonctions de Bessel sphériques,  $j_\ell$  et  $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$  et  $\hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|$  sont des vecteurs unitaires sur la sphère pointant dans la même direction que  $\mathbf{x}$  et  $\mathbf{k}$  respectivement.

1. En toute rigueur, dans le cas général parler de la température d'un rayonnement n'a pas vraiment de sens. Pour le CMB cependant, comme on l'a vu, les observations montrent que son émission suit une loi de corps noir parfait dont le spectre d'émission ne dépend que de sa température. Définir la température de corps noir équivalent du CMB possède donc un sens précis. Dans la suite on parlera donc indifféremment du rayonnement du CMB, de sa température de corps noir équivalente et même plus souvent de sa température tout court.

2. Dans la suite on parlera indifféremment des anisotropies de température relative, de variations de température ou tout simplement de la température du cmb.



### 2.3.3 Projection sur la sphère : statistique des $a_{\ell m}$

Les détecteurs de PLANCK observent le rayonnement provenant de la surface de dernière diffusion (avec un décalage vers le rouge du spectre dû à l'expansion). Cette surface est considérée comme une sphère (qu'on peut choisir normalisée). On a donc effectué une projection pour passer des fluctuations aléatoires tri-dimensionnelles à un champ sur la sphère. Les fluctuations sont décrites dans l'espace de Fourier à trois dimensions comme des variables aléatoires décorréliées (voir l'annexe A). Un modèle de découplage instantané (voir par exemple Bartlett (1999)) nous indique que l'opération de projection est une transformation linéaire. Ainsi les coefficients  $a_{\ell m}$  de la décomposition harmonique des anisotropies de température  $\Delta T(\hat{\mathbf{n}})$  sur la sphère restent des variables aléatoires décorréliées.

La mesure du CMB, correspond à la mesure d'une réalisation particulière d'un processus aléatoire. On s'intéresse alors uniquement à ses propriétés statistiques. Comme on l'a vu dans la section 2.2.2, les anisotropies de température du CMB sont couplées aux fluctuations de densité de la matière, générées par les perturbations primordiales qui sont supposées gaussiennes. Le CMB donne donc accès aux fluctuations de l'Univers au moment de la recombinaison dans un régime où l'évolution de l'Univers était encore largement dominée par une évolution linéaire. Les anisotropies de température du CMB au niveau de la surface de dernière diffusion sont donc également gaussiennes. La description d'une telle distribution est plus compliquée dans l'espace direct car elle est définie par une matrice de covariance complète. Dans l'espace des harmoniques sphériques, en revanche, les  $a_{\ell m}$  sont des variables aléatoires décorréliées suivant chacune une distribution gaussienne.

Comme on s'intéresse aux variations par rapport à la moyenne du CMB on a :

$$\langle a_{\ell m} \rangle = 0 \quad (2.12)$$

Par ailleurs, comme les  $a_{\ell m}$  sont des variables aléatoires décorréliées on a :

$$\langle a_{\ell m} a_{\ell' m'}^* \rangle = C_\ell \delta_{\ell \ell'} \delta_{m m'} \quad (2.13)$$

Cette relation définit le *spectre de puissance angulaire* théorique  $C_\ell$  qui donne donc la variance des  $a_{\ell m}$  pour un  $\ell$  donné. Pour une variable aléatoire gaussienne tous les moments d'ordre supérieur sont nuls. Le spectre de puissance angulaire contient donc toute l'information. Les modèles d'Univers s'attachent à prédire les spectres théoriques. La comparaison des observations aux modèles permet alors de déterminer les paramètres cosmologiques.

La mesure que l'on peut faire à partir des observations du CMB correspond à la mesure d'une réalisation particulière d'un tirage gaussien. Pour en déduire les propriétés statistiques, il faudrait en principe moyenner les observations sur plusieurs réalisations de l'Univers ou, sous l'hypothèse d'ergodicité moyenner les observations obtenues depuis différents lieux de l'espace. Cependant, on n'a accès qu'à un seul Univers et on ne peut réaliser les observations que depuis un seul point de vue. On n'a donc accès qu'à une seule réalisation du CMB. L'isotropie de l'Univers permet toutefois de construire un estimateur pour le spectre de puissance à partir de la mesure des  $a_{\ell m}$  de la réalisation particulière à laquelle on a accès. En effet, dans ces conditions les valeurs moyennes ne dépendent que de  $\ell$ . Ainsi on peut estimer le spectre de puissance angulaire du CMB, à partir des  $a_{\ell m}$  issus des observations, de la façon suivante :

$$\hat{C}_\ell = \frac{1}{2\ell + 1} \sum_m |a_{\ell m}|^2 \quad (2.14)$$

### 2.3.4 Erreur de mesure et variance cosmique

Les  $a_{\ell m}$  étant des variables gaussiennes indépendantes, d'après l'équation 2.14, pour une réalisation, chaque  $\hat{C}_\ell$  estimé suit une loi de  $\chi^2$  à  $2\ell + 1$  degrés de liberté. Moyenné sur un ensemble d'Univers l'espérance des  $\hat{C}_\ell$  est égale au spectre de puissance théorique  $C_\ell$  :

$$\langle \hat{C}_\ell \rangle = C_\ell \quad (2.15)$$

Pour une réalisation donnée, la variance de chaque  $\hat{C}_\ell$  est alors donnée par la variance d'une loi de  $\chi^2$  à  $2\ell + 1$  degrés de liberté et d'espérance  $C_\ell$ , soit :

$$\sqrt{\text{Var}(\hat{C}_\ell)} = \frac{2C_\ell}{\sqrt{2\ell + 1}} \quad (2.16)$$

La statistique des  $C_\ell$  n'est pas gaussienne. Cependant, d'après le théorème central limite, la loi de  $\chi^2$  tend vers une distribution normale quand le nombre de degrés de liberté tend vers l'infini. C'est un résultat important qui donne du sens aux barres d'erreurs tracées sur les points expérimentaux quand on présente les  $C_\ell$ . En effet, lorsqu'on associe des erreurs aux mesures, on fait généralement l'hypothèse implicite que ces mesures sont distribuées de façon gaussienne. L'erreur à  $1\sigma$  peut alors être interprétée comme un intervalle à 65% de niveau de confiance. Toutefois, à bas  $\ell$ , la distribution du  $\chi^2$  est fortement asymétrique et n'est pas bien approchée par une gaussienne. L'estimation des erreurs à bas  $\ell$  doit donc faire l'objet de développements spécifiques.

La variance donnée par l'équation 2.16 est la variance intrinsèque liée à toutes mesures possibles des anisotropies de température. Elle est souvent appelée *variance cosmique* et traduit une spécificité de la cosmologie, domaine dans lequel l'observateur n'a accès qu'à une réalisation du processus physique qu'il étudie.

La variance cosmique limite la précision avec laquelle on peut comparer une observation, au spectre théorique prédit par un modèle. Par exemple, PLANCK réalisera la mesure ultime du spectre de puissance angulaire des anisotropies de température du CMB, jusqu'à  $\ell \sim 2000$ , ce qui signifie qu'il réalisera une mesure limitée par la variance cosmique. Dans ce cas, l'erreur de mesure statistique attendue pour un  $\ell$  mesuré, sera inférieure à la variance cosmique pour ce  $\ell$ . Obtenir une mesure plus précise ne permettra donc pas de réaliser une meilleure comparaison au modèle étant donné qu'on attend une variation intrinsèque du spectre de puissance par rapport à la valeur théorique, pour la réalisation particulière de notre Univers, plus grande que l'erreur de la mesure. Autrement dit, peu importe la précision de la mesure du CMB et la précision sur la détermination des  $a_{\ell m}$  et du spectre de puissance qui en résulte, on aura toujours une erreur intrinsèque du fait qu'on a accès à une seule réalisation particulière du processus aléatoire. Il existe ainsi toujours une gamme de modèles compatibles avec les mesures.

### 2.3.5 Spectre de puissance

Le spectre de puissance des fluctuations de température du CMB a été mesuré avec une grande précision par l'expérience WMAP. On peut voir ces résultats sur la figure 2.3. On distingue généralement trois zones particulières. À bas  $\ell$ , le spectre est relativement plat, cette région est appelée le plateau Sachs-Wolfe et correspond à la contribution des fluctuations du potentiel gravitationnel qui sont supposées être invariantes d'échelle. Elle contribue donc à la même puissance, quelle que soit l'échelle, ce qui produit un spectre plat. La seconde région contient une série de pics et de creux qui sont produits par les oscillations acoustiques en jeu au moment de la recombinaison. Finalement, on constate une diminution rapide de la puissance à plus haut  $\ell$ , provoquée par le moyennage des fluctuations à petite échelle, sur la ligne de visée dans l'épaisseur de la « surface » de dernière diffusion d'une part et à l'amortissement de Silk (ou amortissement diffusif) d'autre part, provoqué par la diffusion des photons (voir Lineweaver (1997) et la figure 2.4).

## 2.4 Polarisation

Avant la recombinaison, les photons du CMB subissent de nombreuses interactions avec les électrons libres du plasma par diffusion Thomson. La présence d'anisotropies de température au moment du découplage, permet à ce mécanisme de polariser partiellement le rayonnement du CMB. L'étude de la polarisation du CMB apporte deux nouvelles observables qui devraient permettre de fixer de nouvelles contraintes sur les modèles d'Univers et l'estimation des paramètres cosmologiques. Leurs mesures est cependant très difficile car le niveau du signal est plus d'un ordre de grandeur en dessous de celui des anisotropies de température.

### 2.4.1 Polarisation d'une onde électromagnétique - Paramètres de Stokes

Nous introduisons dans cette section, le formalisme nécessaire à l'étude de la polarisation d'une onde électromagnétique. On considère une onde électromagnétique, de pulsation  $\omega$  se propageant selon l'axe  $z$  d'un repère orthonormé  $Oxyz$ . Par définition, le champ électrique de l'onde,  $\mathbf{E}$ , est toujours perpendiculaire à la direction de propagation. On peut donc suivre son évolution au cours

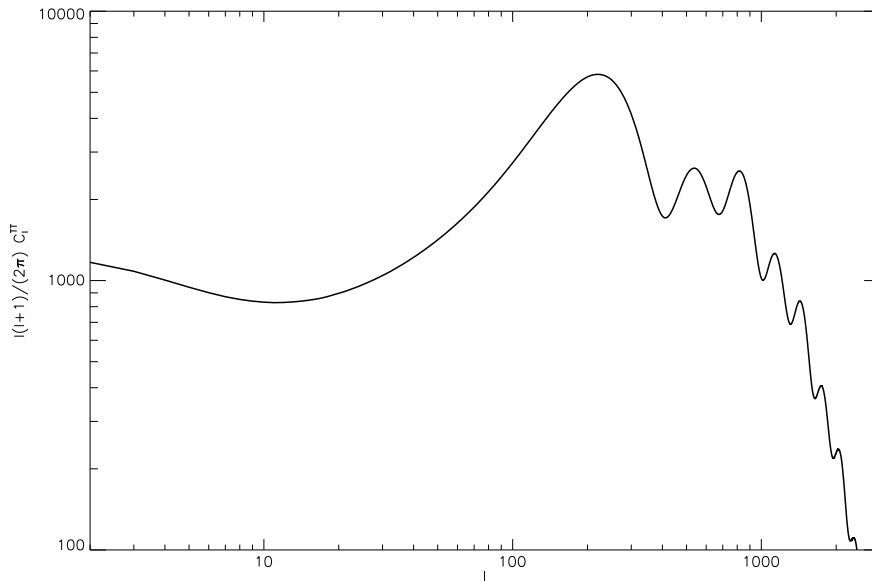


FIGURE 2.3 – **Spectre de puissance de température du CMB.** Le calcul a été effectué avec le programme CAMB à partir des paramètres cosmologiques déterminés par WMAP dans un modèle  $\Lambda$ CDM. Les unités sont des  $\mu K^2$

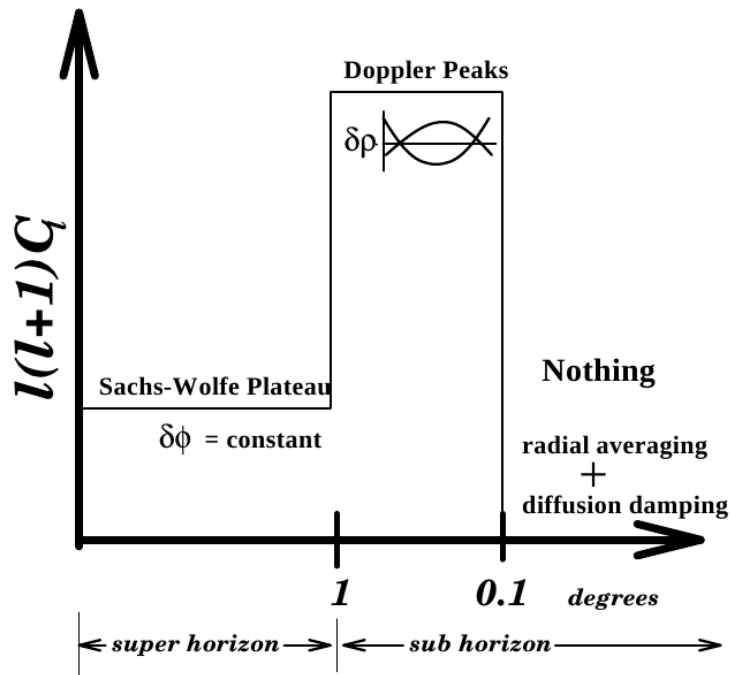


FIGURE 2.4 – **Représentation schématique du spectre de puissance de température du CMB**. On distingue trois zones particulières : le plateau Sachs-Wolfe à grande échelle angulaire suivi des oscillations acoustiques qui dominent jusqu'à une échelle angulaire de  $0.1^\circ$ , puis l'amortissement diffusif. D'après Lineweaver (1997).

du temps, en le projetant dans le plan  $(Oxy)$ . Dans le cas général, on obtient les équations :

$$E_x(t) = A_x(t)e^{i(\omega t)} \quad (2.17)$$

$$E_y(t) = A_y(t)e^{i(\omega t + \phi)} \quad (2.18)$$

La polarisation de l'onde est déterminée par la direction de son champ électrique. Lorsque  $A_x$ ,  $A_y$  et  $\phi$  évoluent de façon aléatoire au cours du temps, le champ électrique n'a pas de direction privilégiée. On dit alors que l'onde n'est pas polarisée. Une onde non polarisée peut être représentée comme la superposition d'ondes possédant chacune une polarisation indépendante de celles des autres et aléatoire. Finalement, lorsque les composantes du champ électrique sont corrélées, on dit que l'onde est polarisée.

Pour une onde polarisée, dans le cas le plus général où  $\phi$  est quelconque et  $A_x$  est différent de  $A_y$ , l'extrémité de la projection du vecteur champ électrique en fonction du temps, décrit une ellipse dans le plan  $(Oxy)$ . C'est pourquoi on parle de polarisation *elliptique*<sup>3</sup>. On distingue également deux cas particuliers :

- lorsque  $A_x = A_y$ , on a une polarisation dite *circulaire*.
- lorsque  $\phi = 0$  on a une polarisation *linéaire*.

Pour caractériser la polarisation, il est donc nécessaire de déterminer  $A_x$ ,  $A_y$  et  $\phi$ . Il est commode d'utiliser pour cela, les paramètres de Stokes qui ont l'avantage de pouvoir être mesurés directement de manière assez simple. Ils sont donnés par :

$$I = \langle |E_x|^2 \rangle + \langle |E_y|^2 \rangle = I_x + I_y \quad (2.19)$$

$$Q = \langle |E_x|^2 \rangle - \langle |E_y|^2 \rangle = I_x - I_y \quad (2.20)$$

$$U = \langle |E_x^{45}|^2 \rangle - \langle |E_y^{45}|^2 \rangle = I_x^{45} - I_y^{45} \quad (2.21)$$

$$V = 2A_x A_y \sin \phi \quad (2.22)$$

L'exposant 45 signifie qu'on se place dans une base orientée à 45° autour de l'axe  $z$ , par rapport à la base  $(x, y, z)$ . Le premier paramètre,  $I$ , est l'intensité totale, c'est-à-dire la somme des intensités dans les deux directions  $x$  et  $y$ . Le deuxième paramètre,  $Q$ , est la différence entre l'intensité transmise par un polariseur linéaire orienté selon l'axe  $x$  et un autre orienté selon l'axe  $y$ . Le troisième paramètre,  $U$ , est la différence entre l'intensité transmise par un polariseur linéaire orienté à 45° de l'axe  $x$  et un autre orienté à -45°. Enfin le dernier paramètre,  $V$ , correspond à la différence entre l'intensité transmise par un polariseur circulaire droit et un polariseur circulaire gauche.

Il est donc possible de mesurer les trois premiers paramètres à l'aide de deux polariseurs linéaires convenablement orientés. Il faut garder à l'esprit toutefois que la mesure de  $Q$  et  $U$  dépend du choix du repère. Le dernier paramètre qui donne accès au signe de la phase, est difficilement accessible aux mesures, en général, notamment car le temps d'intégration des détecteurs est trop élevé devant la fréquence de l'onde.

À partir des paramètres de Stokes, on définit le degré de polarisation  $p$  par :

$$p = \frac{\sqrt{Q^2 + U^2 + V^2}}{I} \quad (2.23)$$

Il permet de déterminer la proportion polarisée d'un rayonnement. Pour une onde non polarisée, le champ électrique varie aléatoirement et donc  $Q$ ,  $U$  et  $V$  sont nuls. Dans ce cas, le degré de polarisation vaut zéro. Au contraire pour une onde *totalelement polarisée*,  $E_x$ ,  $E_y$  et  $\phi$  sont constants donc :

$$\begin{aligned} Q^2 + U^2 + V^2 &= [A_x^2 - A_y^2]^2 + [2A_x A_y \cos \phi]^2 + [2A_x A_y \sin \phi]^2 \\ &= (A_x^2)^2 + (A_y^2)^2 - 2A_x^2 A_y^2 + 4A_x^2 A_y^2 (\cos^2 \phi + \sin^2 \phi) \\ &= (A_x^2)^2 + (A_y^2)^2 + 2A_x^2 A_y^2 \\ &= (A_x^2 + A_y^2)^2 \\ &= I^2 \end{aligned}$$

et  $p$  vaut 1. On peut également définir la fraction de polarisation linéaire. Elle est donnée par :

$$f = \frac{\sqrt{Q^2 + U^2}}{I} \quad (2.24)$$

Finalement l'intensité de polarisation linéaire est donnée par :

$$P = \sqrt{Q^2 + U^2} \quad (2.25)$$

<sup>3</sup>. La polarisation elliptique peut toujours s'écrire comme la somme d'une polarisation linéaire et d'une polarisation circulaire

### 2.4.2 Diffusion Thomson

La diffusion Thomson est la limite à basse énergie de la diffusion Compton, qui décrit la diffusion élastique d'un rayonnement électromagnétique par une particule chargée. Cette description est valable dans la limite où l'énergie des photons est petite devant l'énergie de masse de la particule ( $h\nu \ll mc^2$ , où  $m$  est la masse de la particule,  $c$  la vitesse de la lumière dans le vide,  $\nu$  la fréquence de l'onde et  $h$  la constante de Planck) de telle sorte qu'on peut considérer que la trajectoire de la particule diffusante n'est pas affectée.

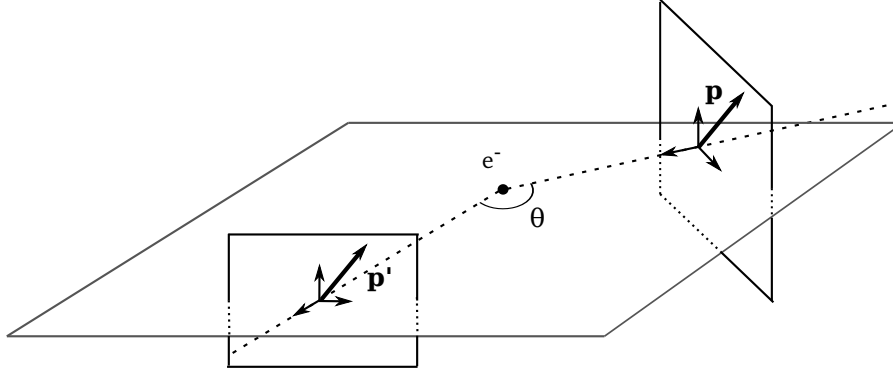


FIGURE 2.5 – **Diffusion Thomson.** La diffusion Thomson correspond à la diffusion élastique d'un photon par une particule chargée. La section efficace dépend de l'angle entre la direction de polarisation de l'onde incidente  $\hat{\mathbf{p}} = \frac{\mathbf{p}}{|\mathbf{p}|}$  et diffusée  $\hat{\mathbf{p}}' = \frac{\mathbf{p}'}{|\mathbf{p}'|}$ .

La section efficace d'interaction est indépendante de la longueur d'onde, en revanche elle dépend de la direction de polarisation. Elle est donnée par :

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{4\pi\epsilon_0 mc^2}\right)^2 |\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}'|^2 \quad (2.26)$$

où  $\hat{\mathbf{p}}$  et  $\hat{\mathbf{p}}'$  représentent les vecteurs unitaires indiquant la direction de polarisation des photons incidents et diffusés respectivement (voir la figure 2.5) et  $d\Omega$  correspond à l'angle solide infinitésimal.

Cette expression montre que la section efficace est modulée par l'angle entre la direction de polarisation de l'onde incidente et de l'onde diffusée. Pour une direction d'observation donnée, les composantes de la polarisation pour lesquelles la section efficace est la plus grande, vont donc être privilégiées. Un calcul simple montre que la composante contenue dans le plan d'observation est intégralement transmise tandis que la composante perpendiculaire à ce plan s'annule. Lorsque le rayonnement incident est non polarisé et isotrope, cet effet conduit juste à une modification éventuelle de l'amplitude car l'onde résultante, dans une direction donnée, est issue de la superposition d'ondes polarisées dans toutes les directions de façon isotrope et possède donc une polarisation moyenne nulle. Finalement, lorsque le rayonnement incident, même non polarisé, présente une anisotropie quadrupolaire, la composante de la polarisation diffusée va dépendre de la direction du rayonnement incident, ce qui conduit à un rayonnement diffusé partiellement polarisé, comme on peut le voir sur la figure 2.6.

### 2.4.3 Polarisation du CMB

Au moment de la recombinaison, l'Univers est un plasma totalement ionisé d'électrons, de protons et de photons. Ces derniers sont en interaction permanente avec les baryons. La température à cette époque est de quelques milliers de degrés. Les particules sont non relativistes et les interactions avec les photons se font par diffusion Thomson. Comme on a vu dans l'équation 2.26, la section efficace d'interaction est proportionnelle à l'inverse de la masse au carré de la particule avec laquelle les photons interagissent. Le rapport entre la masse du proton et de l'électron étant de l'ordre de  $10^3$ , d'après la formule 2.26, ce sont donc les diffusions par les électrons qui dominent la dynamique.

La température au niveau de la surface de dernière diffusion n'est pas parfaitement isotrope. Les anisotropies produisent alors une polarisation partielle du rayonnement par diffusion des photons

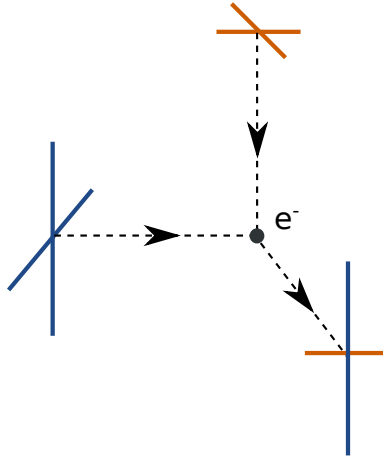


FIGURE 2.6 – **Polarisation du rayonnement.** En présence d’anisotropies dans le rayonnement, l’intensité du rayonnement diffusé par diffusion Thomson va dépendre de la direction. Sur la figure ci-dessus on constate que les intensités incidentes dans les deux directions perpendiculaires transmises ne sont pas identiques, ce qui produit un rayonnement émergent polarisé pour l’observateur.

sur les particules dans les zones de sur-densité ou de sous-densité comme expliqué dans la section 2.4.2 et sur la figure 2.7. La diffusion Thomson ne modifie pas la phase et ne peut donc pas générer de polarisation circulaire. Aucun autre mécanisme connu ne permet de polariser le rayonnement du CMB. Par conséquent, on suppose en général qu’il ne possède pas de polarisation circulaire.

On peut montrer que pour le CMB seules les anisotropies quadrupolaires peuvent polariser le rayonnement (voir par exemple Kosowsky (1996)). Les sources d’anisotropies quadrupolaires peuvent être de trois nature :

- **Les perturbations scalaires.** Ce sont les perturbations créées par les sur-densités ou sous-densités au moment de la recombinaison. Les électrons sont attirés par les sur-densités sous l’effet de la gravité. Plus ils s’approchent du centre, plus ils gagnent de l’énergie potentielle et plus leur vitesse augmente jusqu’à la limite où la pression de radiation devient plus importante que l’attraction gravitationnelle. Un électron situé dans une région proche d’une sur-densité, suffisamment loin de la zone où la pression domine, subit une accélération par effet gravitationnel dans la direction radiale à la sur-densité. Dans le référentiel de l’électron tombant vers la sur-densité, la densité augmente tangentiellement au mouvement. Dans la direction radiale en revanche, les électrons semblent s’éloigner car leur accélération est d’autant plus grande qu’ils se rapprochent du centre. Ainsi, la densité dans la direction radiale paraît plus faible tandis que celle dans la direction tangentielle paraît plus élevée. Les sur-densités produisent donc des anisotropies quadrupolaires générant une polarisation du rayonnement du CMB (voir figure 2.7). On a l’effet inverse autour des sous-densités ce qui produit une anisotropie quadrupolaire et une figure de polarisation inversée (voir également la figure 2.7).
- **Les perturbations vectorielles.** Elles correspondent à des mouvements de vortex du fluide et sont prédites dans le cadre d’Univers présentant des défauts topologiques qui sont aujourd’hui largement infirmés par les observations. Les univers possédant une phase d’inflation ne prédisent pas de perturbations vectorielles.
- **Les perturbations tensorielles.** Elles sont produites par le passage d’ondes gravitationnelles primordiales. Ces ondes sont prédites par les modèles d’inflation mais n’ont pas encore été mises en évidence. Elles possèdent un intérêt particulier car elles permettent de sonder les premiers instants de l’Univers. Leur mise en évidence est très difficile et reste un des objectifs importants de la cosmologie observationnelle des prochaines années.

#### 2.4.4 Les champs E et B

Les équations 2.19 définissent les paramètres de Stokes et permettent de caractériser l’intensité et la polarisation d’un rayonnement électromagnétique dans une direction particulière. Dans le cadre

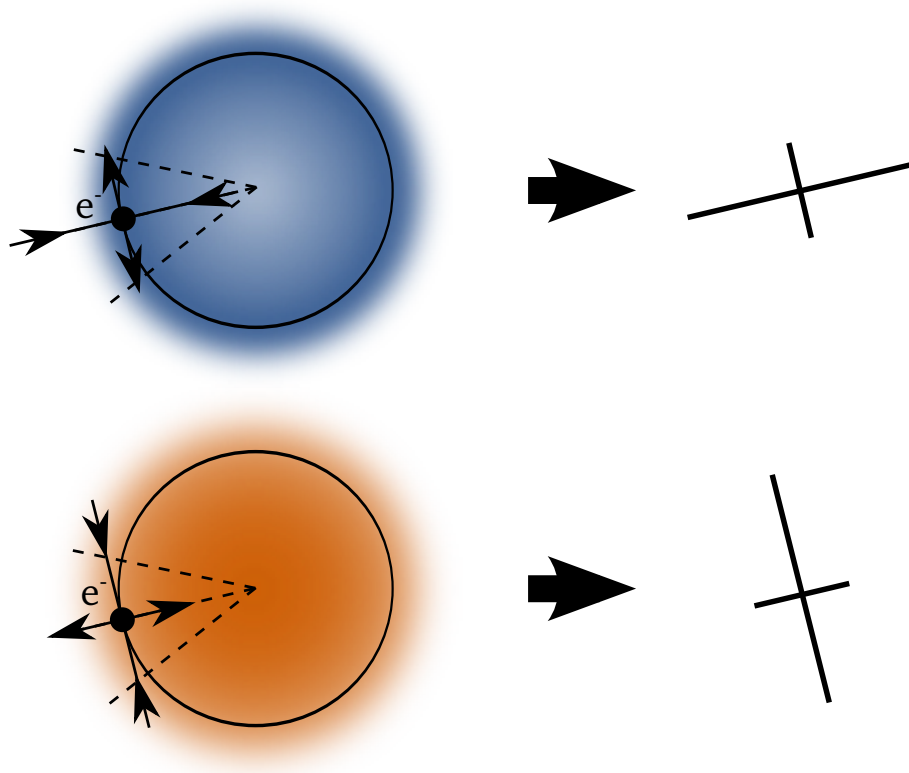


FIGURE 2.7 – **Génération de la polarisation par les perturbations scalaires.** À proximité des sur-densités ou des sous-densités, le plasma est animé d'un mouvement qui produit des anisotropies quadrupolaires. Dans ces régions le CMB apparaît donc polarisé linéairement dans une direction radiale (autour des sous-densités) ou tangentielle (autour des sur-densités).

de l'étude du CMB, il est difficile d'utiliser les observables  $Q$  et  $U$  pour étudier les propriétés statistiques de la polarisation car elles dépendent complètement du choix du repère dans lequel elles sont mesurées. Elles ne sont pas invariantes par rotation autour de la ligne de visée.

La solution consiste à construire deux nouvelles observables scalaires qui ne dépendent plus du choix du repère, par une transformation linéaire et inversible de  $Q$  et  $U$ . On commence par introduire les quantités de spin 2 :  $(Q \pm iU)(\hat{\mathbf{n}})$  (voir Zaldarriaga & Seljak (1997)). On peut les décomposer sur la base des harmoniques sphériques spinnées d'ordre 2 (Zaldarriaga & Seljak (1997)) :

$$(Q + iU)(\hat{\mathbf{n}}) = \sum_{\ell m} a_{2,\ell m 2} Y_{\ell m}(\hat{\mathbf{n}}) \quad (2.27)$$

$$(Q - iU)(\hat{\mathbf{n}}) = \sum_{\ell m} a_{-2,\ell m -2} Y_{\ell m}(\hat{\mathbf{n}}) \quad (2.28)$$

On définit alors les champs E et B, dans l'espace des harmoniques sphériques :

$$a_{\ell m}^E = \frac{-(a_{2,\ell m} + a_{-2,\ell m})}{2} \quad (2.29)$$

$$a_{\ell m}^B = \frac{i(a_{2,\ell m} - a_{-2,\ell m})}{2} \quad (2.30)$$

Le nom E et B viennent de l'analogie avec la parité des champs électrique et magnétique. En effet après un changement de parité du repère local, E est invariant tandis que B change de signe. Une détection des modes B met en évidence la présence de perturbations tensorielles (associées aux ondes gravitationnelles primordiales).

### 2.4.5 Spectres de puissance de polarisation du CMB

De la même manière que pour les anisotropies de température, on peut décomposer les champs scalaires E et B sur la base des harmoniques sphériques.

$$E(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{m=+\ell} a_{\ell m}^E Y_{\ell m}(\theta, \phi) \quad (2.31)$$

et

$$B(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{m=+\ell} a_{\ell m}^B Y_{\ell m}(\theta, \phi) \quad (2.32)$$

On définit alors les spectres de puissance angulaire par :

$$\langle a_{\ell m}^E a_{\ell m}^{E*} \rangle = C_{\ell}^{EE} \quad (2.33)$$

et

$$\langle a_{\ell m}^B a_{\ell m}^{B*} \rangle = C_{\ell}^{BB} \quad (2.34)$$

Finalement, on peut définir les spectres croisés entre les observables T, E et B de la façon suivante :

$$\langle a_{\ell m}^X a_{\ell m}^{Y*} \rangle = C_{\ell}^{XY} \quad (2.35)$$

où  $X$  et  $Y$  peuvent prendre n'importe quelle valeur T, E ou B (si  $X = Y$ , on retrouve les définitions précédentes). On peut montrer que comme T et B d'une part et E et B d'autre part ont des parités opposées, leurs spectres de corrélation sont nuls (Zaldarriaga & Seljak (1997)). La figure 2.8 présente les spectres de puissance du CMB correspondant aux prédictions attendues pour un modèle de type  $\Lambda$ CDM. Ils ont été calculés avec CAMB, à partir des paramètres cosmologiques issus de sept années d'observation du satellite WMAP.



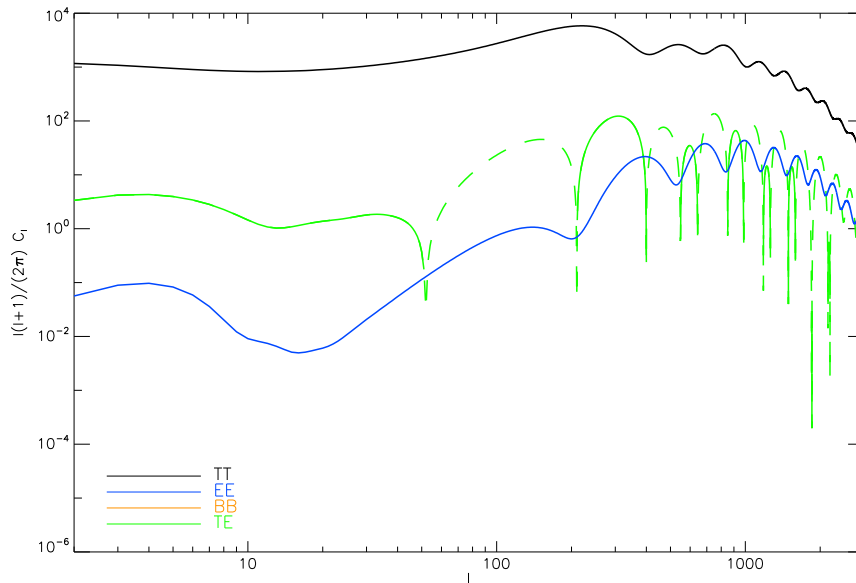


FIGURE 2.8 – **Spectres de puissance du CMB.** Le calcul a été effectué avec CAMB à partir des paramètres cosmologiques déterminés par WMAP intégrant 7 ans de données pour un modèle  $\lambda$ CDM. Les unités sont des  $\mu K^2$

## 2.5 Émissions d'avant-plan

De nombreuses composantes astrophysiques émettent aux mêmes longueurs d'onde que le CMB. On distingue souvent les différentes émissions par leur aspect diffus, ponctuel ou quasi-ponctuel ainsi que par leur origine galactique ou extragalactique. Le signal mesuré à une fréquence donnée, correspond donc à la superposition de l'ensemble des composantes émettant à cette fréquence. Lorsqu'on s'intéresse à l'étude du CMB ces avant-plans constituent un bruit et il est nécessaire de séparer ou soustraire leur contribution. Entre 217GHz et 857GHz, en particulier, les observations de PLANCK seront les premières à couvrir l'ensemble du ciel. On distingue les émissions d'avant-plan des anisotropies secondaires car ce ne sont pas des modifications du CMB lui-même mais bien des émissions distinctes.

### 2.5.1 Émissions diffuses

Les émissions diffuses sont principalement d'origine galactique à l'exception de l'émission du fond diffus infrarouge. Ce rayonnement a été détecté pour la première fois en 1996 (Puget et al. (1996)) et résulte de l'émission cumulée des galaxies infrarouges le long de la ligne de visée. Les autres sources diffuses sont toutes d'origine galactique. Sur la figure 2.9 on peut voir l'amplitude de l'émission des principales sources galactiques diffuses, comparée à celle du CMB, en fonction de la fréquence. On trouve :

- **l'émission de la poussière galactique.** Elle domine à haute fréquence. Elle est produite par le rayonnement des grains de poussière. Le milieu interstellaire est essentiellement constitué de gaz très dilué (principalement de l'hydrogène neutre) ainsi que de grains de poussière d'un diamètre inférieur au micromètre (environ 1% de la masse du gaz). Les grains apparaissent sous différentes tailles et compositions et possèdent donc des propriétés physiques différentes. Ils absorbent en général les rayonnements UV et visible, sont chauffés, puis réémettent dans une gamme de longueur d'onde allant de l'infrarouge lointain au submillimétrique. On peut distinguer trois populations de grains (Desert et al. (1990)), chacune contribuant à l'émission dans un domaine spectral particulier. L'émission dans des longueurs d'onde en dessous de  $10\mu m$  est dominée par celle des grains de petite taille (de l'ordre du nanomètre). Ils sont principalement constitués de chaînes carbonnées sous forme de PAH (Poly Aromatic Hydrocarbons). Pour les longueurs d'onde en dessus de  $100\mu m$  l'émission est engendrée par des gros

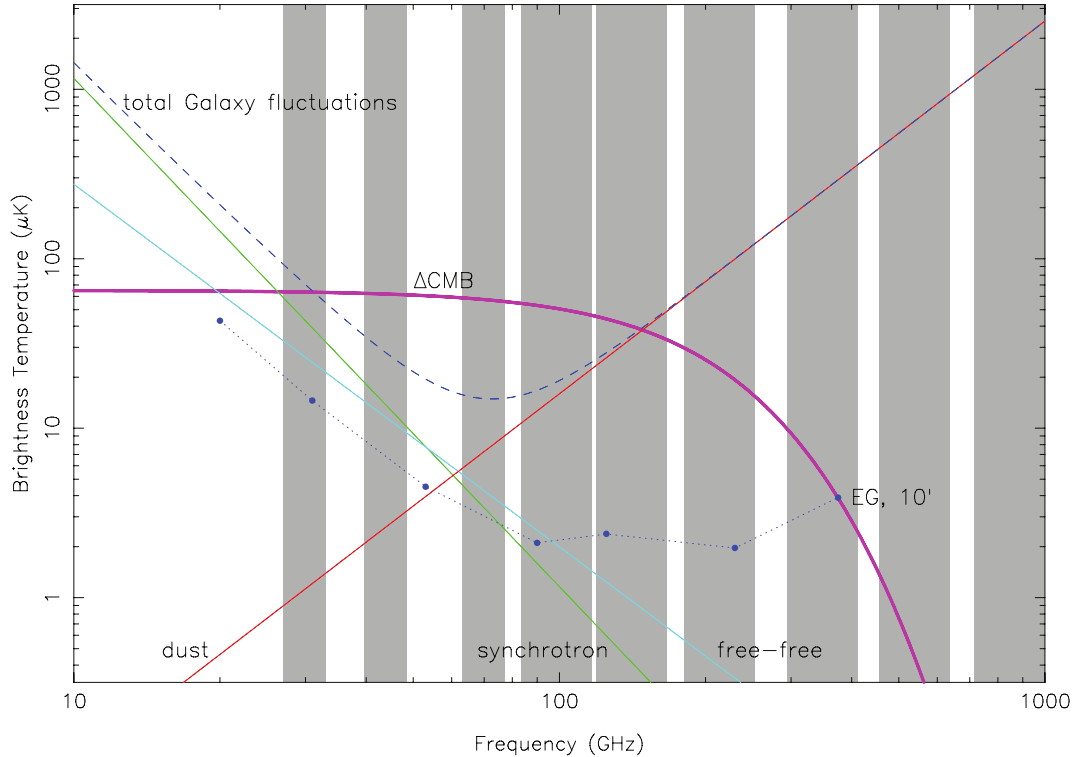


FIGURE 2.9 – **Émissions des composantes en fonction de la longueur d'onde.** Les bandes de fréquence d'observation des instruments de PLANCK sont représentées en gris. On voit que l'émission du CMB domine autour de 100GHz. (Tiré de Planck Science Team (2005))

grains, ayant un diamètre de l'ordre de  $0.1\mu m$ . Ce sont des grains à l'équilibre thermique avec le rayonnement provenant des étoiles environnantes. La troisième catégorie de grains, les grains intermédiaires ne sont pas à l'équilibre thermique car leur petite taille ne leur permet pas d'absorber suffisamment de photons. De plus ils possèdent une petite capacité thermique et sont alors fortement chauffés lors de l'absorption d'un photon. Ils subissent donc de grandes variations de température.

- **le rayonnement synchrotron galactique.** Dans la galaxie, les électrons peuvent être accélérés à des vitesses relativistes. Ils vont suivre des trajectoires en spirale autour des lignes de champ magnétique galactique et produisent un rayonnement par effet synchrotron. L'intensité de ce rayonnement suit une loi de puissance ( $I \sim \nu^\alpha$ , où  $\nu$  est la fréquence et  $\alpha$  vaut -2.5). Il domine donc l'émission à basse fréquence comme on peut le voir sur la figure 2.9.
- **le rayonnement de freinage** (on trouve aussi le terme anglais free-free ou allemand Bremsstrahlung). Il est émis par les électrons du milieu interstellaire, freinés par le champ électrique produit par les ions des régions fortement ionisées (essentiellement les régions riches en hydrogène ionisé (Régions H II)). Le spectre d'émission de ce rayonnement est également une loi de puissance, de plus faible pente que le rayonnement synchrotron.
- **l'émission zodiacale.** Elle correspond à l'émission de la poussière du système solaire. La distribution des grains est confinée principalement dans le plan de l'écliptique. Cette émission est donc maximale aux faibles latitudes écliptiques.

La longueur du trajet parcouru dans la galaxie varie en fonction de la ligne de visée. Dans la direction du plan galactique on traverse une large partie de la galaxie, en revanche, perpendiculairement à ce plan, la distance parcourue dans la galaxie est plus faible. L'intensité des émissions galactiques est donc en général plus importante à faible latitude galactique.

### 2.5.2 Émission des sources compactes

L'émission des sources compactes est l'émission de sources ponctuelles ou quasi-ponctuelles et correspond au rayonnement émis par des corps suffisamment éloignés pour avoir une taille angulaire

faible comparée à la taille du lobe de l'instrument. Aux longueurs d'onde de PLANCK, on distingue des sources d'origine extragalactique, galactique et du système solaire. Les sources extragalactiques sont principalement des galaxies radios (dont l'émission est dominée par l'émission synchrotron) et des galaxies infrarouges mais on trouve également d'autres sources comme les quasars, les blazars... Les sources galactiques correspondent principalement à des restes de supernovae, des étoiles de faible masse en formation ou des nuages moléculaires froids. Enfin, dans le système solaire les planètes, les astéroïdes ou les comètes sont susceptibles d'être observés comme des objets quasi-ponctuels. Un catalogue de sources compactes a été récemment publié par la collaboration PLANCK (voir Planck Collaboration et al. (2011)).

### 2.5.3 Émissions polarisées

En plus des anisotropies de température, le satellite PLANCK va faire la mesure de la polarisation du CMB. Les émissions polarisées d'avant-plan doivent donc être également caractérisées afin de pouvoir soustraire leur contribution. Certaines sources compactes sont polarisées, notamment des sources radios ou des restes de supernovae (comme la nébuleuse de crabe). Les émissions diffuses produisent également un signal polarisé. Certaines populations de grains de la poussière galactique ont une forme allongée et possède un moment magnétique. Les modèles prédisent qu'en présence d'un champ magnétique, ces grains vont se mettre en rotation autour de leur petit axe le long des lignes de champs magnétiques. Ils vont émettre préférentiellement dans la direction de leur grand axe. L'émission ainsi obtenue apparaît donc polarisée dans une direction perpendiculaire au champ magnétique (voir par exemple Ponthieu et al. (2005)). Le rayonnement synchrotron présente lui aussi une polarisation intrinsèque. Le degré de polarisation de l'émission synchrotron galactique peut atteindre plusieurs pourcents selon les régions du ciel considérées. Enfin, le rayonnement de freinage n'est lui pas polarisé mais il est émis dans des régions fortement ionisées. Comme pour le CMB, les photons peuvent y subir des diffusions Thomson. Ainsi en présence d'anisotropies, le rayonnement de freinage peut être partiellement polarisé.

## Chapitre 3

# Effet de lentille gravitationnelle sur le CMB

L'effet de lentille gravitationnelle correspond à la déviation des trajectoires des rayons lumineux, causée par la présence de masse. Dans ce chapitre nous nous intéressons plus particulièrement à l'effet sur les photons du CMB. Dans la première section nous introduisons les équations de base qui régissent cet effet. Dans la section suivante nous présentons le potentiel de lentille qui intègre les effets de lentille gravitationnelle sur la ligne de visée. Les conséquences sur les observables du CMB sont finalement données dans la troisième section.

### 3.1 Présentation

Les rayons lumineux se propagent le long des géodésiques de l'espace-temps. Leur trajectoire est donc modifiée lorsqu'ils voyagent à travers un champ gravitationnel inhomogène. Cet effet est décrit avec une grande précision par la théorie de la relativité générale d'Einstein. La mesure de la déviation par le Soleil, des rayons lumineux provenant d'étoiles de l'amas des Hyades, lors d'une éclipse totale de soleil en 1919, a conduit à une large médiatisation de la théorie de la relativité générale. Mais ce n'est que dans la seconde moitié du vingtième siècle que la précision des mesures a permis de réellement confirmer les prédictions de cette théorie. L'effet de lentille gravitationnelle ne dépendant que de la perturbation de la métrique par le potentiel gravitationnel, c'est un outil important pour l'astrophysique car il donne accès aux masses des objets indépendamment de la nature de la matière qui les compose et de leur émission électromagnétique.

L'image de galaxies lointaines vue à travers un système de lentille gravitationnelle peut apparaître fortement déformée, ou encore, lorsque la déflexion est suffisamment forte, plusieurs trajets peuvent parvenir à l'observateur et une source peut alors posséder de multiples images. Ces phénomènes constituent les manifestations les plus remarquables de l'effet de lentille gravitationnelle<sup>1</sup>. En 1979, Walsh et ses collaborateurs ont observé pour la première fois un système d'images multiples d'un quasar (Walsh et al. (1979)). Cette découverte a ouvert la voie au développement de la recherche des effets de lentille gravitationnelle.

Suivant la source et le déflecteur considéré, l'effet de lentille permet d'étudier une grande diversité de phénomènes astrophysiques. Des grands programmes de recherche de matière noire baryonique comme EROS (Afonso et al. (2003)) ou MACHO (Alcock et al. (1996)), ont basé leurs observations sur l'effet de micro-lentille produit par des corps de faible masse comme des naines brunes ou des planétoïdes. Ces expériences se basaient sur la variation d'intensité que peut produire par effet gravitationnel le transit d'un corps devant une étoile. Elles ont joué un rôle important sur la détermination de l'origine de la matière noire en fournissant une limite de la contribution des objets baryoniques de faible masse. Par ailleurs, l'effet de lentille gravitationnelle permet également d'étudier des galaxies situées à grand *redshift* grâce à l'amplification produite par des galaxies ou amas de galaxies d'avant-plans. On peut encore estimer la composante de matière noire des amas de galaxies en analysant la déflexion induite sur des galaxies d'arrière plan et en comparant la masse dé-

---

1. Dans ce cas on parle généralement d'effet de lentille gravitationnelle fort

duite à la masse baryonique visible. L'amas du Boulet par exemple fait l'objet de nombreuses études pour ce genre d'analyses (voir par exemple Clowe et al. (2004)). Il est également possible d'étudier la répartition de matière dans l'Univers à plus grande échelle en réalisant une mesure statistique de l'ellipticité d'un grand nombre de galaxies. En effet, en plus de l'ellipticité intrinsèque des galaxies qui s'annule si on moyenne sur un grand nombre d'objets, on peut observer dans certaines régions des déformations corrélées produites par effet de lentille gravitationnelle. La mesure de ces effets donne accès à la répartition de matière. Ce type d'analyse porte le nom de cisaillement cosmique (cosmic shear en anglais). On trouvera dans l'article de revue Bartelmann & Schneider (2001) un bon aperçu des applications de l'effet de lentille gravitationnelle en astrophysique. Finalement, l'effet de lentille gravitationnelle permet également d'étudier de façon statistique, la distribution de matière de l'Univers à partir du CMB (voir notamment l'article de revue Lewis & Challinor (2006)). C'est ce dernier point qui nous intéresse plus particulièrement et qui a motivé les développements de cette thèse.

Lorsqu'on s'intéresse à l'effet de lentille gravitationnelle, le CMB représente une source particulièrement intéressante. En effet, son émission s'étend sur tout le ciel et la surface de dernière diffusion est située à grand *redshift* ( $z=1100$ ) ce qui permet de sonder des zones à plus grand *redshift* qu'une étude basée sur les galaxies. Par ailleurs, la polarisation du CMB donne deux observables de plus. Enfin, les propriétés statistiques du CMB sont bien connues et sont décrites dans une bonne approximation par un champ gaussien.

L'étude de l'effet de lentille gravitationnelle du CMB est motivée par les conséquences directes sur les observables. En effet, la déflexion perturbe les propriétés statistiques du CMB. Elle crée notamment des non-gaussianités et modifie de plusieurs pourcents les spectres de puissance. Enfin, l'effet de lentille gravitationnelle produit des modes B de polarisation et constitue une des sources de confusion les plus importantes pour la mesure du signal provenant d'ondes gravitationnelles primordiales. Ces effets sont maintenant accessibles aux observations de PLANCK et doivent impérativement être pris en compte pour ne pas introduire de biais dans les processus d'estimation des paramètres cosmologiques à partir des données notamment.

## 3.2 Cadre général

### 3.3 Hypothèses

Dans le cas général, il est nécessaire d'intégrer l'ensemble des déflexions introduites par la matière rencontrée le long du trajet des photons. Le calcul doit être réalisé dans le cadre de la relativité générale afin d'évaluer les trajectoires perturbées par les modifications de la métrique en présence de matière. Dans cette étude, on se placera dans un modèle d'Univers de type  $\Lambda$ CDM.

Dans des conditions d'homogénéité parfaite, il n'est pas possible de générer des effets de lentille gravitationnelle. Dans la suite, on considère que les fluctuations de densité, à l'origine du phénomène de lentille gravitationnelle, sont des perturbations de la métrique et de faible amplitude donnant lieu à un effet de lentille faible.

Par ailleurs, on considère que les zones où se produisent les déflexions sont petites par rapport au trajet total. Par conséquent on peut diviser le parcours des photons en trois zones. Des trajets non perturbés (rectilignes pour un Univers euclidien) séparés par une zone d'interaction de taille négligeable devant les distances totales. Cette hypothèse correspond au régime des *lentilles minces* (voir figure 3.1). C'est une hypothèse largement vérifiée pour l'étude des effets de lentille gravitationnelle du CMB. Un amas de galaxie par exemple a une dimension typique de l'ordre du Mpc qui est négligeable devant les distances parcourues par les photons du CMB de plusieurs Gpc.

Finalement, on considère que les angles de déflexion sont faibles et que les photons ne s'écartent donc presque pas de leur trajectoire non perturbée. On peut alors évaluer le potentiel perturbateur sur le trajet que les photons auraient parcouru en l'absence de fluctuations de densité (c'est-à-dire une ligne droite pour un Univers homogène et plat). Cette hypothèse est connue sous le nom d'approximation de Born. Dans le cadre de cette approximation, on peut considérer que la déflexion ne se fait que dans le plan perpendiculaire à la ligne de visée. L'effet de lentille gravitationnelle « n'allonge » donc pas le trajet. Il existe en principe une composante longitudinale mais il a été montré dans Hirata & Seljak (2003b) qu'elle était négligeable étant donné la résolution et le niveau de bruit de PLANCK et des expériences à venir.

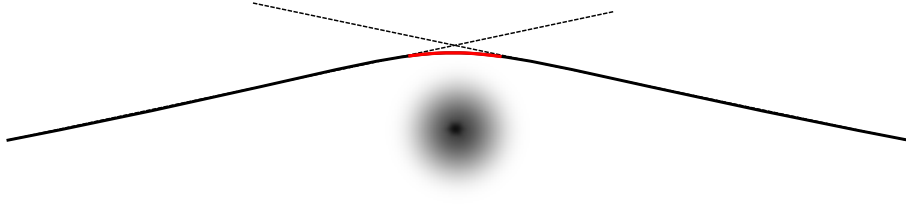


FIGURE 3.1 – **Approximation des lentilles minces.** Si la distance entre le déflecteur et la source et le déflecteur et l’observateur est suffisamment grande on peut se placer dans le régime des lentilles minces et considérer que la déflexion se produit de façon instantanée au niveau du déflecteur et que le reste du parcours se fait de façon non perturbée, c’est-à-dire en ligne droite pour un Univers euclidien. Une justification géométrique est donnée par le schéma ci-dessus dans lequel on voit que la zone (en rouge) sur laquelle le trajet réel (en trait plein) dévie fortement du trajet rectiligne tangent (en traits pointillés) est très faible comparé à l’ensemble du trajet.

### 3.3.1 Équation de lentille

Une lentille (L) dévie la trajectoire lumineuse des photons provenant d’une source (S) d’un angle  $\alpha$  (voir figure 3.2). La source apparaît donc à un observateur (O) comme l’image (I) située dans une direction différente. Il est possible d’exprimer cet angle en fonction des positions angulaires de la source et de son image et des distances angulaires entre l’observateur, la lentille et la source. Si on reprend les notations de la figure 3.2, dans l’approximation des petits angles, un calcul géométrique conduit alors à l’équation suivante :

$$\theta_I = \theta_S + \frac{D_{LS}}{D_{OS}} \alpha \quad (3.1)$$

qui correspond à l’équation *géométrique de lentille*. On pourra chercher à l’inverser afin de reconstruire le champ source, c’est-à-dire le champ observable en l’absence de déflexion par effet de lentille gravitationnelle. Le rapport  $\frac{D_{LS}}{D_{OS}}$  est maximal quand le déflecteur est placé à mi-chemin entre la source et l’observateur. Les déflecteurs situés dans cette région produisent donc les plus forts angles de déflexion.

Lorsqu’on s’intéresse au CMB il est plus intéressant de travailler à partir de la différence de position angulaire, introduite par l’effet de lentille gravitationnelle, au niveau de la surface de dernière diffusion. D’après la figure 3.2, on constate qu’une source repérée par la position angulaire  $\theta_S$  sera observée avec une position angulaire différente  $\theta_I$  à cause de la déflexion d’angle  $\alpha$  introduite par l’effet de lentille gravitationnelle. Cette différence de position angulaire s’écrit :

$$\delta\theta = \theta_I - \theta_S = \frac{D_{LS}}{D_{OS}} \alpha \quad (3.2)$$

En intégrant cette équation le long de la ligne de visée, on peut alors définir un champ de déflexion  $\mathbf{d}$  en tout point du ciel. Pour une source située à la distance comobile  $\chi_*$ , dans un Univers plat, on peut écrire (voir Lewis & Challinor (2006)) :

$$\mathbf{d}(\hat{\mathbf{n}}) = -2 \int_0^{\chi_*} d\chi \frac{\chi_* - \chi}{\chi_*} \nabla_{\perp} \Psi(\chi \mathbf{n}; \eta_0 - \chi) \quad (3.3)$$

où  $\hat{\mathbf{n}}$  est un vecteur unitaire sur la sphère définissant une position angulaire,  $\eta$  est le temps conforme et l’indice 0 désigne la valeur présente. Le champ  $\Psi$  correspond au champ de potentiel gravitationnel tri-dimensionnel. Pour le CMB on suppose que la source est située au niveau de la surface de dernière diffusion et que cette dernière possède une épaisseur négligeable. La distance  $\chi_*$  correspond donc à celle de la surface de dernière diffusion.

L’équation précédente correspond donc à un *réarrangement* de la position de chaque point source. Si on considère un champ à deux dimensions,  $P$ , subissant un effet de lentille gravitationnelle, on peut déterminer le champ image,  $\tilde{P}$ , en estimant la position image de chaque point source. Cette opération peut s’écrire sous la forme :

$$\tilde{P}(\hat{\mathbf{n}}) = P(g(\hat{\mathbf{n}})) \quad (3.4)$$

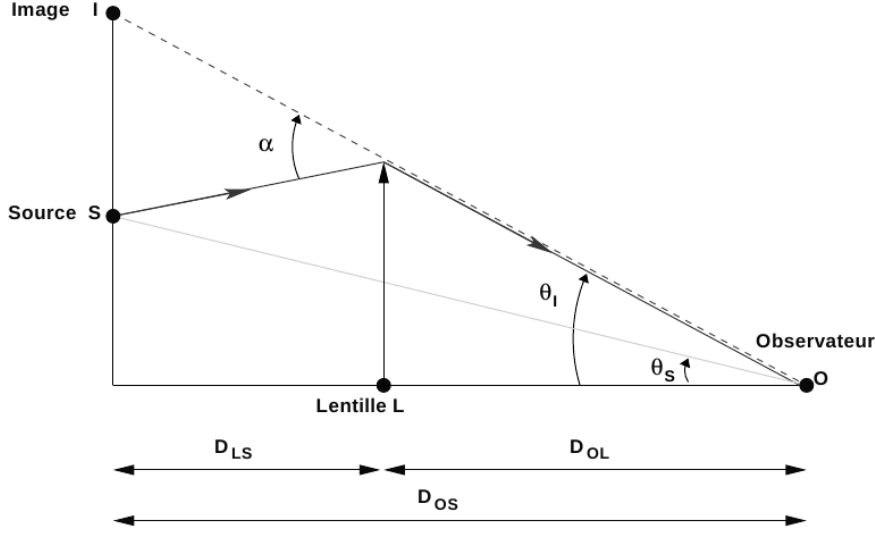


FIGURE 3.2 – **Notations pour l'équation de la déflexion** Une source (S) est repérée par un vecteur angulaire  $\theta_S$  par rapport à une direction arbitraire fixée pour l'observateur (O). Une lentille gravitationnelle (L) produit alors une déflexion d'angle  $\alpha$  et l'image de la source possède finalement une direction apparente  $\theta_I$ . On note  $D_{OL}$  la distance angulaire entre l'observateur et la lentille,  $D_{OS}$  la distance angulaire entre l'observateur et la source et  $D_{LS}$  la distance angulaire entre la lentille et la source.

où  $g$  correspond à la *fonction de réarrangement* qui est donnée par l'équation et s'exprime de la façon suivante :

$$g(\hat{\mathbf{n}}) = \hat{\mathbf{n}} + \mathbf{d}(\hat{\mathbf{n}}) \quad (3.5)$$

Cette équation décrit la transformation entre le champ source et le champ image. Il est courant de décrire cette transformation à l'aide de son Jacobien  $J$ , ou de la matrice d'amplification  $a = J^{-1}$  qu'on exprime généralement sous la forme (Bartelmann & Schneider (2001)) :

$$a^{-1} = J = \left( \frac{\partial g(\hat{\mathbf{n}})}{\partial \hat{\mathbf{n}}} \right) = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \quad (3.6)$$

Dans le cadre des effets de lentille faibles, le terme  $\kappa$  s'interprète comme une convergence, c'est-à-dire qu'il correspond à l'amplification isotrope d'un élément de surface. Il est relié au champ de déflexion par l'équation  $\kappa = \frac{1}{2} \nabla \cdot \mathbf{d}$ . Les termes  $\gamma_1$  et  $\gamma_2$  correspondent à un cisaillement (shear en anglais), c'est-à-dire un étirement dans deux directions propres perpendiculaires. Ils traduisent les déformations à surface constante.

Le cisaillement est l'observable la plus importante lorsqu'on étudie l'effet de lentille gravitationnelle des galaxies car il décrit la déformation des images. C'est la mesure statistique des corrélations des ellipticités des galaxies qui permet de détecter l'effet de lentille gravitationnelle. Cette mesure peut se faire sans connaître la distribution des galaxies sources. Pour le CMB, en revanche, on dispose de nombreuses contraintes sur la distribution du champ non lentillé. Ses propriétés sont bien connues. Il est ainsi intéressant de tenir compte en plus du cisaillement, des effets de convergences sur des échelles de la taille des perturbations. On s'intéressera alors généralement à la déflexion directement qui intègre l'ensemble des effets produits par le réarrangement des points par les lentilles gravitationnelles.

### 3.3.2 Potentiel de lentille

Dans le cadre de l'approximation des petits angles (effet de lentille gravitationnelle faible), l'intégrale de l'équation 3.3, est évaluée dans le cadre de l'approximation de Born, c'est-à-dire le long du chemin non perturbé. Dans ce cas, on peut relier le gradient transverse à un gradient angulaire sur la sphère à une distance  $\chi$  par la relation  $\nabla_{\perp} = \nabla \Psi / \chi$  (voir Lewis & Challinor (2006)). On peut

alors réécrire l'équation 3.3 sous la forme d'un gradient de potentiel :

$$\mathbf{d} = \nabla\phi \quad (3.7)$$

où le champ  $\phi$  désigne le *potentiel de lentille*. Il s'écrit sous la forme :

$$\phi(\mathbf{n}) = -2 \int_0^{\chi_*} d\chi \frac{\chi_* - \chi}{\chi_* \chi} \Psi(\chi \mathbf{n}; \eta_0 - \chi) \quad (3.8)$$

et représente l'intégrale le long de la ligne de visée du champ gravitationnel à trois dimensions.

Les propriétés statistiques du potentiel de lentille dépendent de la géométrie de l'Univers et du potentiel gravitationnel sur la ligne de visée. Si on considère que le champ gravitationnel suit une évolution linéaire à partir de perturbations initiales gaussiennes, on peut considérer dans une bonne approximation que le champ  $\phi$  est un champ scalaire gaussien. Dans la pratique, l'évolution récente des structures à petite échelle n'est pas linéaire mais le grand nombre de structures à petite échelle rencontrées sur la ligne de visée rend la distribution gaussienne (voir Hanson et al. (2010a)). Les non-linéarités modifient tout de même le spectre de puissance du potentiel de lentille principalement à haut  $\ell$ . Pour obtenir un calcul précis il est donc nécessaire de se baser sur des simulations des structures et de leur évolution et procéder par lancé de rayon. Il est tout de même possible de prendre en compte les corrections non linéaires au niveau du spectre de puissance à l'aide d'une approche paramétrique comme HALOFIT couplé au code CAMB.

Les propriétés statistiques du potentiel de lentille sont décrites par son spectre de puissance car c'est un champ gaussien. On peut alors écrire les relations suivantes :

$$\langle \phi \rangle = 0 \quad (3.9)$$

$$\langle \phi_{\ell m} \phi_{\ell' m'}^* \rangle = \delta_{\ell \ell'} \delta_{m m'} C_\ell^{\phi\phi} \quad (3.10)$$

Le spectre de puissance du champ  $\phi$  est relié à celui du champ de déflexion. D'après l'équation 3.7 en calculant le gradient dans l'espace des harmoniques sphériques, on obtient :

$$C_\ell^{dd} = \ell(\ell + 1) C_\ell^{\phi\phi} \quad (3.11)$$

On peut voir sur la figure 3.3 le spectre de puissance du potentiel de lentille, calculé avec CAMB à partir des paramètres cosmologiques estimés sur les données correspondant à sept années d'observation de WMAP.

### 3.3.3 Ordres de grandeur et validité des approximations

On trouvera une discussion intéressante sur la validité des différentes approximations sur lesquelles reposent les résultats précédant dans Lewis & Challinor (2006) et Hanson et al. (2010a). Nous reprenons ici un certain nombre d'ordres de grandeur qui confortent les hypothèses que nous avons utilisées. L'élément le plus important est la validité de l'approximation des petits angles. C'est la condition nécessaire pour utiliser l'approximation de Born qui permet d'évaluer le potentiel sur la trajectoire non perturbée des photons. Si cette approximation n'était pas vérifiée, il faudrait résoudre le problème beaucoup plus complexe de lancé de rayons sur la trajectoire, en tenant compte de l'ensemble des déflexions rencontrées. Ce qui nécessite de réaliser des simulations de la densité de matière à trois dimensions, en prenant en compte son évolution au cours du temps.

Afin d'estimer un ordre de grandeur de la déflexion des photons, on peut considérer que l'effet qu'ils subissent au cours de leur trajet est une succession d'interactions avec les grandes structures. La taille typique des puits de potentiel est de l'ordre de 300Mpc. Elle est donnée par le maximum du spectre de puissance de la matière. La déflexion moyenne introduite par ces structures est de l'ordre de  $10^{-4}$  radians (voir Hanson et al. (2010a)). La distance de la surface de dernière diffusion, à l'époque actuelle, étant de l'ordre de 14000Mpc, les photons du CMB qu'on observe ont subi une cinquantaine d'interactions en moyenne. L'écart type de l'amplitude des déflexions est donc de l'ordre de  $50^{1/2} \times 10^{-4} \sim 7.10^{-4}$  radians soit environ deux minutes d'arc. Ce résultat est du même ordre de grandeur que celui de la section 3.4.2 déterminé à partir du spectre de déflexion attendu avec la mesure des paramètres cosmologiques effectuée par WAMP, dans un modèle  $\Lambda$ CDM. L'approximation des petits angles devrait donc être bien vérifiée, ce qui renforce ainsi la validité de l'approximation de Born.



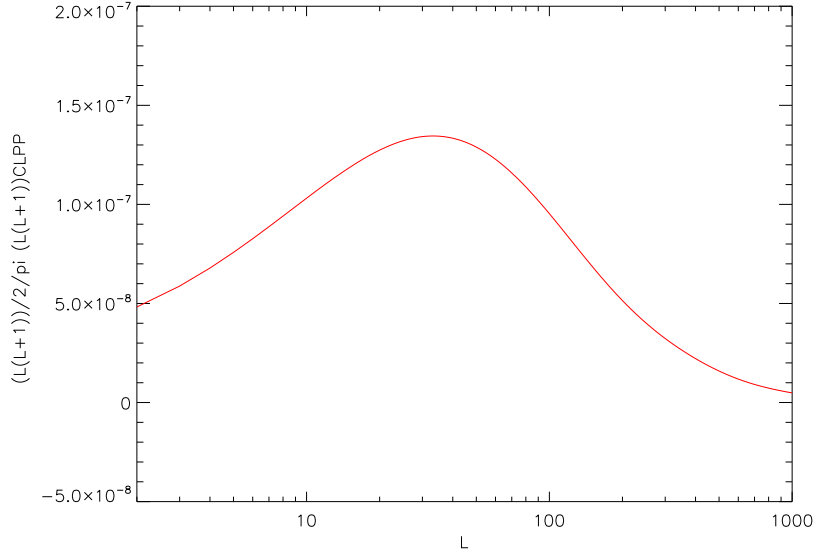


FIGURE 3.3 – **Spectres de puissance de la déflexion.** Le calcul est effectué avec CAMB dans un modèle  $\Lambda$ CDM à partir des paramètres cosmologiques déterminés par l'équipe de WMAP intégrant sept années d'observations.

On s'attend par ailleurs, à ce que la déflexion soit corrélée sur de grandes échelles. Comme on a vu dans la section 3.3.1, la déflexion possède une amplitude maximale quand le déflecteur est placé à mi-chemin entre la source et l'observateur. Les objets situés dans ces régions contribuent donc plus fortement à la déflexion. L'ordre de grandeur de l'échelle des corrélations peut donc être estimé en évaluant la taille angulaire d'une structure de 300Mpc, placée à mi-distance entre l'observateur et la surface de dernière diffusion. Ce qui donne une taille angulaire d'environ  $300/7000 \approx 0.043$  radians, soit  $\sim 2.5^\circ$  (voir Hanson et al. (2010a)). Cette valeur est en accord avec les résultats attendus à partir du spectre de la déflexion calculé avec les paramètres cosmologiques issus des mesures de WMAP. En effet sur la figure 3.3 on voit que le spectre de puissance angulaire de la déflexion est maximal vers  $\ell \sim 50$ , ce qui correspond à une échelle angulaire de quelques degrés.

Enfin, à très petite échelle angulaire (bien en dessous de la minute d'arc), on peut trouver des régions isolées du ciel où l'angle de déflexion est grand, notamment quand la ligne de visée est quasiment alignée avec un objet très massif. Cependant, seul un très petit nombre de rayons sont affectés. Dans tous les cas, cet effet est totalement négligeable dans le cas de PLANCK, étant donnée la taille du lobe instrumental. Pour les expériences futures, il est possible de suivre l'approche de Hirata & Seljak (2003b) en considérant ces régions comme des avant-plans à soustraire quand on s'intéresse à l'effet de lentille gravitationnelle sur le CMB.

## 3.4 Effet sur les observables du CMB

### 3.4.1 Effet sur les cartes du CMB

Comme on l'a vu précédemment, l'effet de lentille gravitationnelle défléchit faiblement la trajectoire des photons du CMB. La distribution d'intensité mesurée par les détecteurs ne correspond donc pas directement à celle de la surface de dernière diffusion. La température mesurée  $\tilde{T}$  dans une direction  $\hat{\mathbf{n}}$ , correspond à la température de la surface de dernière diffusion,  $T$ , provenant d'une direction légèrement différente  $\hat{\mathbf{n}}'$  donnée par l'équation suivante :

$$\tilde{T}(\hat{\mathbf{n}}) = T(\hat{\mathbf{n}}') = T(\hat{\mathbf{n}} + \mathbf{d}(\hat{\mathbf{n}})) \quad (3.12)$$

Cet effet conduit à un réarrangement des points de la source dont l'effet est illustré sur la figure 3.4. On peut y voir une carte carrée de 10 degrés de côté, d'une réalisation du CMB, ainsi que son image tenant compte des effets de lentille gravitationnelle produit par une réalisation du potentiel

de lentille. On trouve également la différence entre les cartes source et image, comparée à la carte du potentiel de lentille.

L'effet sur les cartes de polarisation est similaire. À partir des paramètres de Stokes  $Q$  et  $U$ , on peut définir les observables :

$$\pm X(\hat{\mathbf{n}}) = Q(\hat{\mathbf{n}}) \pm iU(\hat{\mathbf{n}}) \quad (3.13)$$

qu'on appelle paramètres de Stokes complexes. Si on désigne par  $\pm X$  ces champs au niveau de la surface de dernière diffusion alors le champ observé,  $\pm \tilde{X}$ , après avoir subi l'effet de lentille gravitationnelle décrit par le potentiel de lentille  $\phi$ , est donné par (Hu (2000)) :

$$\pm \tilde{X}(\hat{\mathbf{n}}) = \pm X(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})) \quad (3.14)$$

Les paramètres de Stokes complexes sont donc transformés de la même manière que la température.

### 3.4.2 Effet sur le spectre de puissance des anisotropies de température

La déflexion des photons du CMB par effet de lentille gravitationnelle modifie les propriétés statistiques des anisotropies de température mesurées. Le champ lentillé n'est plus totalement gaussien, en particulier son trispectre n'est plus nul (voir Hanson et al. (2011), Hu (2000) et Kesden et al. (2003)) et son spectre de puissance est modifié.

Il est possible d'estimer le spectre de puissance du CMB observé (c'est-à-dire tenant compte des effets de lentille gravitationnelle) en fonction du spectre primordial au niveau de la surface de dernière diffusion. Nous suivons la démarche de Hu (2000), qui permet dans une première approximation, de se faire une idée de l'impact de la déflexion sur le spectre de puissance des anisotropies de température tel qu'on pourrait l'observer en l'absence d'effet de lentille gravitationnelle. On se place pour cela dans l'approximation plane puis on commence par réaliser un développement limité de l'équation 3.12 au second ordre en  $\phi$ . On obtient alors l'équation :

$$\begin{aligned} \tilde{T}(\hat{\mathbf{n}}) &= T(\hat{\mathbf{n}} + \mathbf{d}(\hat{\mathbf{n}})) \\ &= T(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})) \\ &\approx T(\hat{\mathbf{n}}) + \nabla\phi(\hat{\mathbf{n}}) \cdot \nabla T(\hat{\mathbf{n}}) + \frac{1}{2} \nabla\phi(\hat{\mathbf{n}}) \cdot \nabla\phi(\hat{\mathbf{n}}) (\nabla \cdot \nabla T(\hat{\mathbf{n}})) + 0([\nabla\phi(\hat{\mathbf{n}})]^3) \end{aligned} \quad (3.15)$$

De façon générale, on peut exprimer le spectre de puissance du champ lentillé  $\tilde{T}$  sous la forme :

$$\langle \tilde{T}^*(\mathbf{l}) \tilde{T}(\mathbf{l}') \rangle = (2\pi)^2 \delta(\mathbf{l} - \mathbf{l}') C_l^{\tilde{T}\tilde{T}} \quad (3.16)$$

On peut alors réécrire l'équation 3.15 dans l'espace de Fourier de la façon suivante :

$$\tilde{T}(\mathbf{l}) = T(\mathbf{l}) + \int \frac{d^2\mathbf{l}_1}{(2\pi)^2} T(\mathbf{l}_1) L(\mathbf{l}, \mathbf{l}_1) \quad (3.17)$$

avec,

$$L(\mathbf{l}, \mathbf{l}_1) = \phi(\mathbf{l} - \mathbf{l}_1) (\mathbf{l} - \mathbf{l}_1) \cdot \mathbf{l}_1 + \frac{1}{2} \int \frac{d^2\mathbf{l}_2}{(2\pi)^2} \phi(\mathbf{l}_2) \phi^*(\mathbf{l}_2 + \mathbf{l}_1 - \mathbf{l}) (\mathbf{l}_2 \cdot \mathbf{l}_1) (\mathbf{l}_2 + \mathbf{l}_1 - \mathbf{l}) \cdot \mathbf{l}_1 \quad (3.18)$$

Finalement en injectant l'équation 3.17 dans 3.16 et en utilisant la relation 3.10, on obtient :

$$C_l^{\tilde{T}\tilde{T}} = (1 - l^2 R) C_l^{TT} + \int \frac{d^2\mathbf{l}_1}{(2\pi)^2} C_{|\mathbf{l}-\mathbf{l}_1|}^{TT} C_{\mathbf{l}_1}^{\phi\phi} [(\mathbf{l} - \mathbf{l}_1) \cdot \mathbf{l}_1]^2 \quad (3.19)$$

où  $R$  est donné par :

$$R = \frac{1}{4\pi} \int \frac{dl}{l} l^4 C_l^{\phi\phi} \quad (3.20)$$

Cette quantité est proportionnelle au carré de la valeur quadratique moyenne de la déflexion ( $R = 1/2 \langle |\nabla\phi|^2 \rangle$ ) et vaut  $3 \times 10^{-7}$  pour le spectre de puissance calculé à partir des paramètres cosmologiques estimés par WMAP. Ainsi la valeur quadratique moyenne de l'amplitude de la déflexion attendue, est de l'ordre de  $\sqrt{2} \times 3 \times 10^{-7} \times 180/\pi \times 60 \approx 2.7$  minutes d'arc.

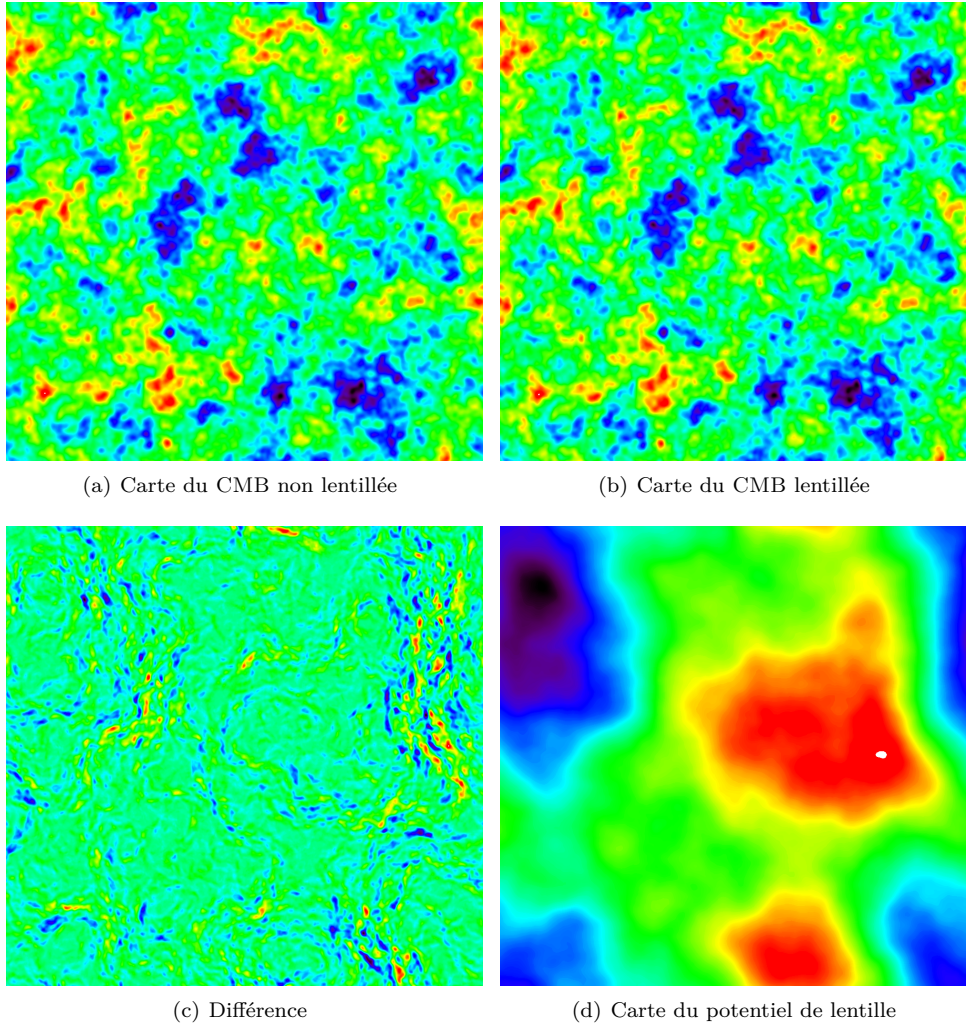


FIGURE 3.4 – **Effet de la déflexion sur le CMB.** Les pixels du champ lentillé (figure (b)) sont obtenus à partir d'un réarrangement des pixels de la carte non lentillée (figure (a)) par le champ de déflexion qui est le gradient du potentiel de lentille qu'on peut voir sur la figure (d). À partir de l'équation 3.12, au premier ordre en  $\phi$ , on peut voir que la différence entre les cartes lentillées et non lentillées est proportionnelle au gradient de  $\phi$  ( $T(\hat{\mathbf{n}}) - \tilde{T}(\hat{\mathbf{n}}) \approx \nabla\phi(\hat{\mathbf{n}}) \cdot \nabla T(\hat{\mathbf{n}})$ ). Cet effet est bien visible sur la figure (c) qui représente la différence entre les cartes du CMB lentillées et non lentillées. On peut remarquer que l'intensité est la plus intense dans les régions où le gradient de la carte du champ de potentiel de lentille est le plus marqué. Chaque carte correspond à une zone du ciel de 10 degrés de côté.

Le second terme de l'équation 3.19 s'interprète comme une convolution du spectre de puissance de température par le spectre du potentiel de lentille. La largeur caractéristique du noyau de convolution est donnée par la largeur à mi-hauteur du spectre de la déflexion. Cette largeur est de l'ordre de  $50\ell$ , ce qui va lisser les structures inférieures à cette échelle. Cet effet est bien visible sur la figure 3.6. Qualitativement, l'effet de lentille gravitationnelle a deux effets principaux sur le spectre de puissance de température : les plus grandes échelles produisent un lissage des pics et des bosses tandis que les déplacements à petite échelle introduisent de la puissance à haut  $\ell$ . On pourra trouver une discussion plus détaillée, ainsi qu'une méthode plus précise pour le calcul du spectre lentillé dans Lewis & Challinor (2006).

Enfin, il est important de noter qu'il existe une corrélation entre la température du CMB et le potentiel de lentille. Elle est produite par l'effet Sachs-Wolfe intégré (Hu (2000)). La corrélation se fait uniquement au très bas  $\ell$ , comme on peut le voir sur la figure 3.5.

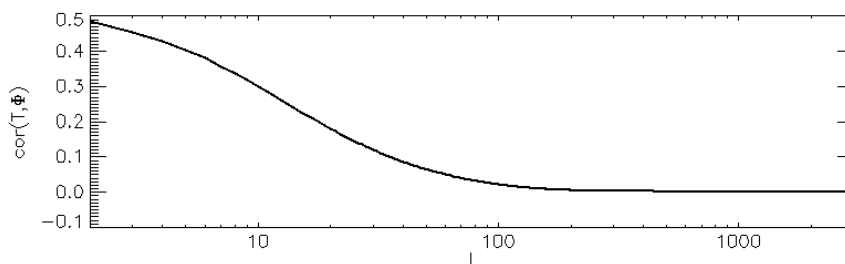


FIGURE 3.5 – **Spectre de corrélation entre la température du CMB et le potentiel de lentille.** Le calcul est effectué avec CAMB pour un modèle d'univers de type *LambdaCDM* à partir des paramètres cosmologiques estimés avec sept ans d'observation par WMAP.

### 3.4.3 Effet sur les spectres de polarisation du CMB

L'impact global sur les spectres de E et B est comparable à celui sur le spectre de température c'est-à-dire que la déflexion produit un lissage général des structures et un déplacement de la puissance à haut  $\ell$ . Par ailleurs, l'effet de lentille gravitationnelle de la polarisation du CMB produit une fuite de puissance de E dans B et constitue une source importante de contamination pour la mesure des modes B primordiaux. On trouvera le détail des équations dans Lewis & Challinor (2006).

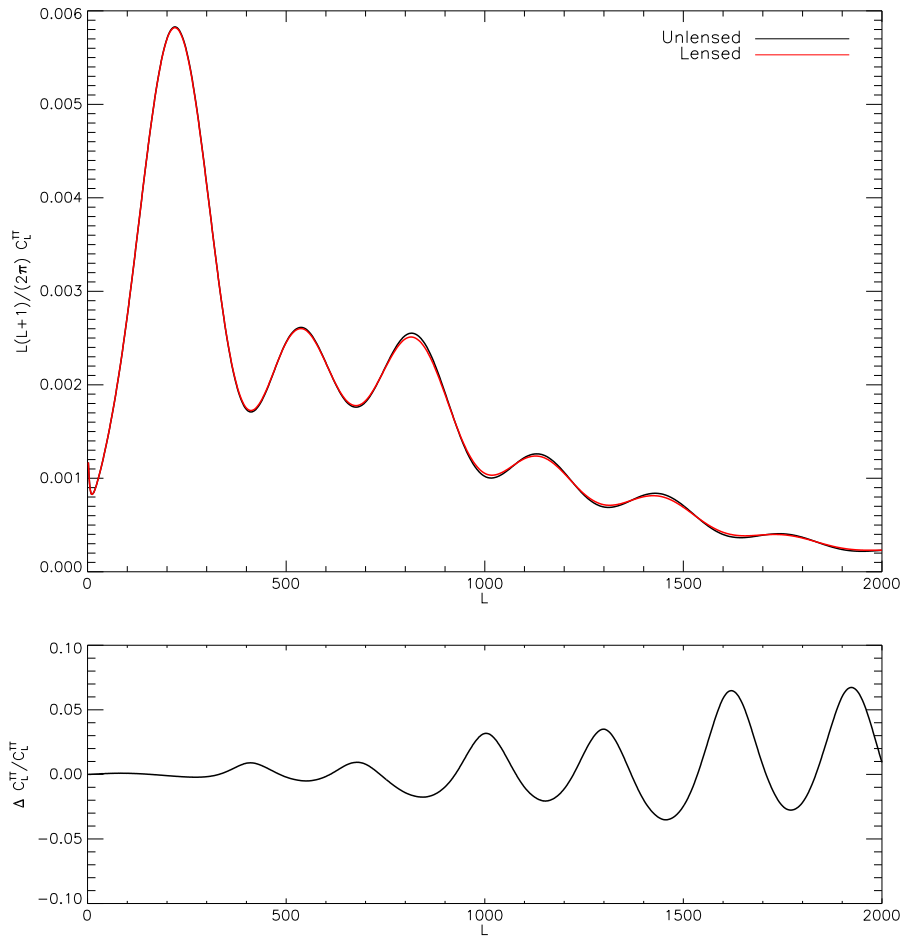


FIGURE 3.6 – **Impact de l’effet de lentille gravitationnelle sur le spectre de puissance de température.** Le calcul est effectué avec CAMB à partir des paramètres cosmologiques de WMAP. Les pics et les creux du spectre de puissance lentillé (en rouge) sont lissés. La différence relative entre les spectres lentillé et non lentillé, est tracée sur la courbe du bas. Elle dépasse les 5% à haut  $\ell$ .

## Deuxième partie

# Compression des données de l'instrument HFI de Planck



## Chapitre 4

# L'expérience Planck

Le satellite PLANCK, lancé le 14 mai 2009, est une mission de l'Agence Spatiale Européenne (ESA) visant à observer le rayonnement du fond diffus cosmologique (CMB), sur le ciel complet avec une sensibilité, une résolution et une couverture en fréquence inégalées. Il est prévu pour balayer quatre couvertures complètes du ciel afin de mesurer le spectre de puissance des anisotropies de température et de polarisation du CMB. Le but final est d'essayer de répondre aux questions fondamentales de la cosmologie moderne et de contraindre les modèles de formation et d'évolution de l'Univers et des grandes structures qui le compose.

De nombreuses composantes galactiques émettant aux longueurs d'ondes d'observation de PLANCK, des résultats particulièrement intéressants sont attendus également sur la physique de notre Galaxie, notamment sur la poussière interstellaire, le rayonnement synchrotron galactique, le rayonnement de freinage... Dans ce chapitre nous présenterons rapidement l'historique et les motivations scientifiques à l'origine du projet PLANCK. Puis, nous détaillerons les éléments qui composent le satellite en nous attardant plus particulièrement sur les deux instruments de mesure, basse fréquence et haute fréquence. En plus des références spécifiques données dans les sections suivantes, une description du satellite PLANCK et de ses objectifs scientifiques peut être trouvée dans le *Planck Blue Book* édité par l'ESA (Planck Science Team (2005)).

### 4.1 Historique et présentation

Le rayonnement du CMB, prédit par Gamow dès 1948 (Gamow (1948b)), a été détecté par hasard, pour la première fois par Penzias et Wilson en 1965 Penzias & Wilson (1965). Par la suite, de nombreuses expériences ont été menées pour caractériser son spectre mais aussi et surtout pour rechercher la présence d'éventuelles anisotropies. Le maximum de l'émission du CMB se fait dans le domaine millimétrique ou infrarouge lointain (voir figure 2.1). Il est relativement difficile de faire des observations à ces longueurs d'onde depuis la Terre, à cause de l'émission thermique des corps d'une part et à cause de l'absorption par l'atmosphère d'une grande partie du rayonnement d'autre part.

Pour s'affranchir du bruit causé par l'émission thermique des instruments il est nécessaire de les refroidir à des températures suffisamment basses en fonction du niveau de bruit requis. Cela nécessite le plus souvent l'emploi de technologies cryogéniques avancées. Il reste toutefois toujours un bruit de fond créé par l'émission de l'ensemble de l'environnement (bâtiments, sol, atmosphère...).

Comme on peut le voir sur la figure 4.1, l'absorption par l'atmosphère dans l'infrarouge est très importante, à cause de la présence de molécules très efficaces pour absorber le rayonnement à ces longueurs d'onde, notamment la vapeur d'eau et l'ozone. On peut réduire ces effets en plaçant les instruments en altitude ou aux pôles, dans des régions où l'atmosphère est particulièrement stable et possède une faible teneur en vapeur d'eau. Plusieurs expériences ont ainsi été menées depuis le sol dont certaines sont toujours en cours de prise de données. On peut citer par exemple ACBAR, CBI, DASI, SPT ou encore ACT en ne retenant que quelques unes des plus connues.

Pour limiter encore les effets de l'atmosphère, il est possible d'embarquer les détecteurs sur des ballons stratosphériques. Des expériences de ce type ont obtenu des résultats particulièrement intéressants, notamment BOOMERanG, MAXIMA ou encore Archeops qui a servi de prototype pour



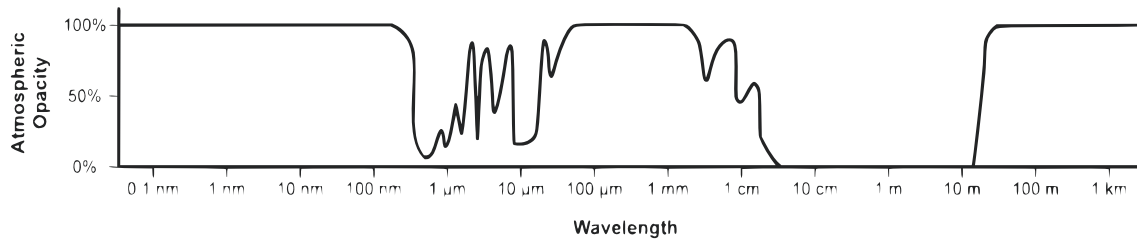


FIGURE 4.1 – **Transmission de l'atmosphère en fonction de la longueur d'onde.** On constate que l'atmosphère est presque totalement opaque à l'ensemble du spectre électromagnétique. On trouve des fenêtres importantes dans le visible et dans le centimétrique qui ont permis de réaliser un grand nombre d'observations depuis le sol à ces longueurs d'onde. Aux longueurs d'onde de PLANCK, dans le domaine millimétrique, on remarque qu'une grande partie du rayonnement est absorbée.

les détecteurs bolométriques et le réfrigérateur à dilution d'hélium de PLANCK. Cependant les vols en ballons n'excèdent pas quelques dizaines d'heures et la couverture du ciel est donc relativement restreinte. La seule façon de s'affranchir totalement des perturbations atmosphériques et d'obtenir une couverture complète du ciel, est de faire les observations depuis l'espace.

Lancé en novembre 1989, le satellite COBE (COsmic Background Experiment) est le premier satellite dédié à l'observation du CMB. Il a fait la mesure la plus précise du spectre du CMB et a permis la première mise en évidence d'anisotropies dans le rayonnement fossile. En 2001, la NASA a lancé le satellite WMAP (Wilkinson Microwave Anisotropy Probe) qui a largement amélioré la mesure des anisotropies de température et a permis de déterminer son spectre de puissance de façon très précise jusqu'à  $\ell$  égal à 1000. Le satellite PLANCK, lancé en mai 2009 est le satellite de troisième génération dédié à l'étude du CMB. En plus d'apporter une meilleure sensibilité et résolution angulaire il permet également de mesurer pour la première fois de façon significative le spectre de puissance de la polarisation du CMB.

La publication en avril 1992 par l'équipe de COBE de la détection de fluctuations dans le rayonnement du CMB à un niveau de l'ordre de  $10^{-5}$  est un événement majeur de la cosmologie moderne qui a suscité une grande effervescence à l'origine de l'intérêt croissant porté à l'étude du CMB. Ainsi, cette même année, deux missions spatiales ont été proposées à l'ESA dans le but d'étudier les anisotropies du CMB : une mission italienne COBRAS (COsmic BackgRound Anisotropies Satellite) et une mission française SAMBA (Satellite for the Measurement of Background Anisotropies). COBRAS devait faire des observations dans les longueurs d'onde de 30GHz à 100GHz avec des radiomètres refroidis par cryogénie. SAMBA devait, de son côté, faire des observations dans les longueurs d'onde entre 100GHz et 1000GHz avec des bolomètres cryogéniques placés au foyer d'un télescope de basse émissivité d'environ 80cm.

En 1994, l'ESA a décidé d'associer les deux projets dans un même satellite, jugeant qu'ils étaient relativement similaires et complémentaires en terme de longueur d'onde et a alors lancé l'étude de phase A pour un satellite COBRA/SAMBA. En 1996, dans le cadre du programme scientifique Horizon 2000, l'ESA a définitivement sélectionné la mission COBRAS/SAMBA comme mission dédiée à l'étude du CMB, en tant que troisième mission de classe moyenne. Un appel d'offre pour deux instruments a été lancé par l'ESA, auquel ont naturellement répondu un consortium international mené par les italiens du CNR de Milan pour l'instrument basse fréquence et un consortium international conduit par l'IAS pour l'instrument haute fréquence.

La mission a été rebaptisée PLANCK, en 1997, en l'honneur du scientifique allemand Max Planck (1858–1947) à qui on doit le modèle d'émission du corps noir. Pour des raisons de budget, l'ESA a décidé de regrouper le satellite d'observation HERSCHEL avec PLANCK, dans la même fusée. C'est en 1999 que la version définitive du satellite et de ses instruments a été acceptée.

Finalement, PLANCK a été lancé avec le satellite HERSCHEL le 14 Mai 2009 à 14h12m02s BST par une fusée Ariane 5 depuis la base de lancement de Kourou. Après un trajet d'un cinquantaine de jours, il a été placé sur son orbite finale autour du second point de Lagrange Soleil-Terre (L2) à une distance d'environ 1.5 million de kilomètres de la Terre. Après une phase de refroidissement, d'étalonnage et de vérification des performances en vol, les observations ont commencé le 13 août 2009 et devraient continuer jusqu'en janvier 2012.

## 4.2 Objectifs scientifiques

L'objectif scientifique de PLANCK est large. Le satellite a été conçu afin d'être capable de faire des mesures limitées par le bruit de photon, aux longueurs d'onde dominées par le CMB. Sa résolution variant de 5 minutes d'arc à 30 minutes d'arc est jusqu'à trois fois supérieure à celle de WMAP et sa sensibilité est jusqu'à dix fois meilleure suivant les voies. Par ailleurs sept des neuf bandes de fréquence de PLANCK sont sensibles à la polarisation linéaire du rayonnement. Sa large couverture spectrale a été optimisée afin de pouvoir séparer au mieux les différentes composantes astrophysiques et l'émission du CMB et de pouvoir détecter de façon optimale les amas de galaxies grâce à l'effet Sunyaev-Zel'dovich.

La mission première de PLANCK est dictée par l'ESA. Le consortium PLANCK s'est engagé à fournir à la communauté les données ordonnées en temps calibrées, de l'ensemble des voies de mesure, les cartes sur tout le ciel combinant tous les détecteurs à chaque fréquence, les cartes des différentes composantes émettant aux longueurs d'onde d'observation, séparées, à partir de l'ensemble des mesures (émission du CMB, de la poussière, du rayonnement synchrotron et de freinage, des raies du CO...) et enfin un catalogue de sources compactes (sources galactiques, galaxies lointaines, amas SZ...)

À terme, PLANCK devrait mesurer le spectre des anisotropies de température du CMB jusqu'à  $\ell$  égal à 3500, en étant limité uniquement par la variance cosmique, jusqu'à  $\ell$  égal à 2000. Il devrait également être capable de mesurer avec une grande précision le spectre du mode E de polarisation jusqu'à  $\ell$  égal à 1500. Finalement, si l'amplitude du mode tenseur n'est pas trop faible par rapport au mode scalaire, PLANCK pourrait mettre en évidence et mesurer pour la première fois le spectre du mode B de polarisation.

La précision sur la mesure du spectre de puissance des anisotropies de température du CMB devrait permettre d'estimer les paramètres cosmologiques avec une précision de quelques pourcents. C'est le résultat le plus important attendu avec PLANCK par les cosmologistes. La mesure du spectre de puissance des modes E de polarisation du CMB permettra par ailleurs de lever des dégénérescences entre certains paramètres. Suivant les modèles les plus favorables, PLANCK pourrait aussi être capable de mesurer les modes B de polarisation. Cela permettrait de mettre des contraintes sur l'inflation et serait une preuve de la présence d'ondes gravitationnelles primordiales. Enfin, grâce à la précision de ses mesures, PLANCK devrait pouvoir mettre de fortes contraintes sur le niveau de non-gaussianité du CMB. La plupart des modèles ne prédisant pas de non-gaussianité, une mise en évidence éventuelle serait un résultat particulièrement intéressant.

En plus des anisotropies primaires, c'est-à-dire des anisotropies présentes lors de la phase de recombinaison, le CMB contient également un certain nombre d'anisotropies secondaires produites durant le parcours des photons depuis la surface de dernière diffusion jusqu'aux détecteurs. Grâce à sa large couverture spectrale, PLANCK va ainsi être capable de mesurer l'effet Sunyaev-Zel'dovich thermique et cinétique par l'empreinte caractéristique qu'il laisse dans le spectre. Il va être sensible à l'effet Sachs-Wolfe intégré. Enfin, il va être capable de mesurer pour la première fois le spectre de déflexion de la matière grâce à l'effet de lentille gravitationnelle du CMB (voir chapitre 8). La détection de cet effet permet en outre de poser des contraintes sur la masse des neutrinos.

Finalement, PLANCK va participer grandement à l'amélioration de la connaissance de la physique galactique. Sa couverture spectrale est sensible à l'émission de la poussière de la Galaxie, du rayonnement de freinage, de l'émission synchrotron. Il va réaliser les premières mesures du ciel complet des émissions galactiques aux plus hautes fréquences d'observation. PLANCK va également permettre de réaliser un catalogue multi-fréquence de sources compactes (galaxies lointaines, amas de galaxies, nuages galactiques froids...)

## 4.3 Description du satellite

Le satellite PLANCK est placé sur une orbite de Lissajous au second point de Lagrange Terre-Soleil (L2), situé à environ 1,5 millions de kilomètres de la Terre. C'est un point gravitationnellement instable donc de légères corrections d'orbite doivent être effectuées régulièrement à l'aide de petits propulseurs à hydrazine. Ce sont également ces propulseurs qui ont été utilisés pour diriger le satellite vers sa position finale en L2.

Sur la figure 4.2 on peut voir les principaux éléments constituant le satellite : le panneau solaire,

le télescope et son baffle, les instruments de mesure, les boucliers thermiques et le module de service qui abrite entre autre, une partie de l'électronique des instruments de mesure, l'ordinateur de bord, le transpondeur et son antenne, le senseur stellaire, le détecteur de rayons cosmiques, les différents tanks contenant les réserves d' $^3\text{He}$ , d' $^4\text{He}$ , d' $\text{H}_2$  pour la cryogénie et d'hydrazine pour les propulseurs...

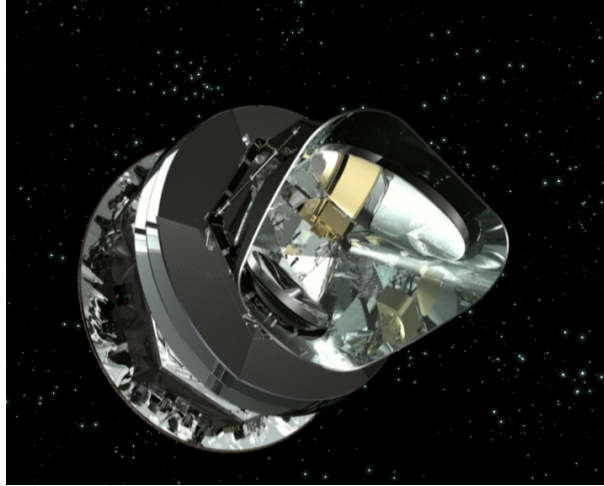


FIGURE 4.2 – **Vue d'artiste du satellite Planck.** On distingue les principales parties du satellite : les miroirs primaire et secondaire du télescope, le plan focal avec les détecteurs, le baffle du télescope, les trois boucliers thermiques, le module de service de forme hexagonale et le dos du panneau solaire.

### 4.3.1 Gestion thermique

Comme on peut le voir sur la figure 4.2, le module de service de PLANCK a une forme hexagonale. Les réserves de carburant et de liquides pour la cryogénie sont placés au centre de ce module dans un large cylindre. Les divers composants électriques et électroniques sont situés sur les faces extérieures pour faciliter la dissipation de chaleur.

Le panneau solaire est associé à des batteries et permet de fournir l'énergie à l'ensemble du satellite. C'est la seule source d'énergie électrique disponible. Il joue également un rôle non négligeable dans la gestion thermique du satellite.

Le satellite tourne sur lui-même autour d'un axe perpendiculaire au panneau solaire. La direction de l'axe de rotation est régulièrement modifiée (voir section 4.3.7). En moyenne elle tourne d'un degré par jour, ce qui compense le changement d'orientation relatif au soleil, induit par la révolution du satellite autour du soleil, de sorte que l'axe de rotation est toujours aligné avec le soleil. Ainsi, le satellite gravite autour du soleil en pointant toujours la même face au soleil (voir figure 4.3). L'ensemble du satellite se trouve donc en permanence à l'ombre du panneau solaire. Cela permet d'une part de protéger les détecteurs de la lumière directe provenant du soleil, de la Terre et de la Lune, mais d'autre part, cela joue également un rôle majeur dans la stratégie globale de contrôle de la température dans les différents étages du satellite.

Le module de service est la partie chaude du satellite. Elle est découplée thermiquement de la partie supérieure constituée du télescope et de son baffle et des modules cryogéniques qui sont refroidis radiativement et des détecteurs qui sont eux refroidis par des méthodes cryogéniques actives. Le module de service a une température d'équilibre élevée d'environ 300K. Une série de trois boucliers permettent de découpler thermiquement cet étage de la partie supérieure du satellite qui fonctionne à très basse température. Les trois boucliers sont formés par des structures réfléchissantes en forme de cône, ayant chacune un angle différent de  $5^\circ$ .

Le large baffle qui entoure le télescope assure un refroidissement radiatif. Il est formé par des structures très aérées d'aluminium, peintes en noir. Cela en fait un radiateur très efficace et entraîne un fort couplage thermique avec le vide. Le baffle participe donc au dispositif global de maîtrise de la température en permettant d'assurer un refroidissement radiatif. L'ensemble du dispositif permet ainsi d'obtenir un couplage radiatif fort avec l'espace grâce aux baffles et une très bonne isolation

entre les deux parties du satellite grâce aux boucliers. Le refroidissement passif radiatif du télescope et du baffle leur permettent d'atteindre une température d'équilibre avec l'espace environnant d'environ 50K. Pour atteindre la précision requise pour les mesures de PLANCK, ses détecteurs nécessitent d'être refroidis à des températures encore plus basses. Ainsi, en plus du refroidissement passif, le satellite PLANCK embarque un système cryogénique actif complexe à plusieurs étages.

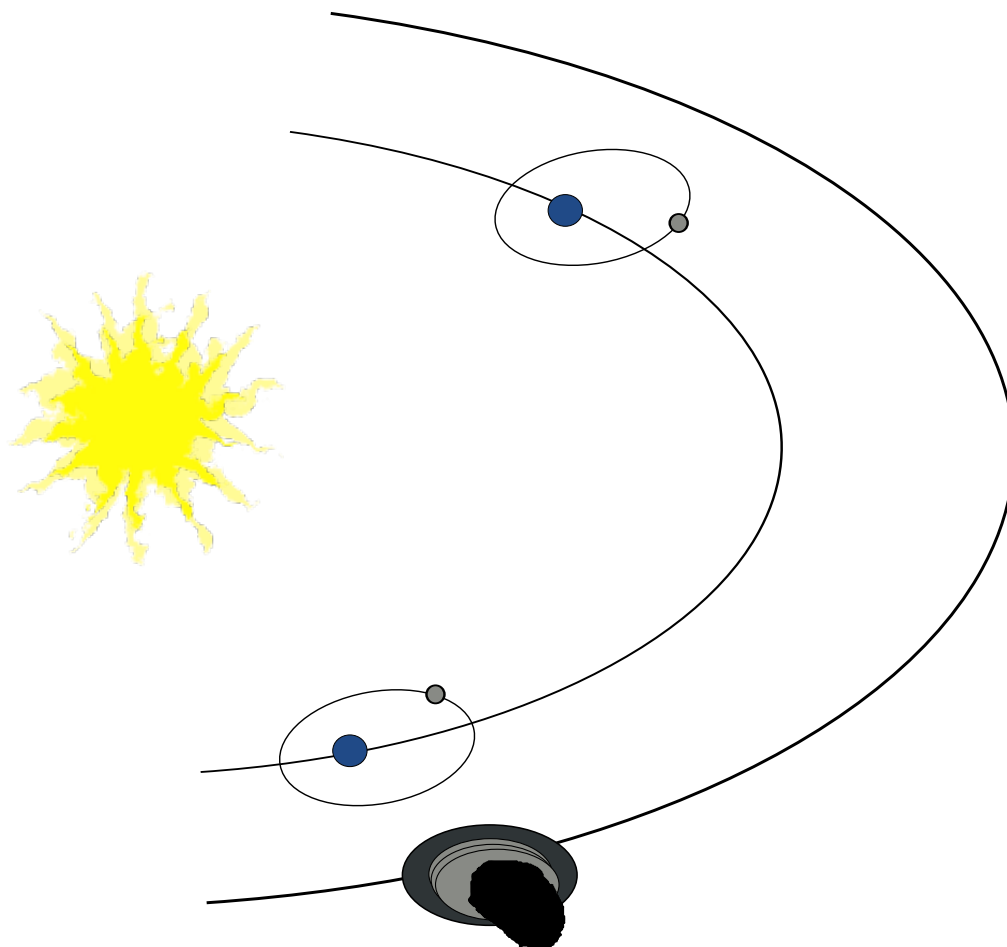


FIGURE 4.3 – **Orbite du satellite au point de Lagrange L2.** PLANCK orbite au second point de Lagrange Soleil-Terre à 1.5 million de kilomètres de la Terre. Il tourne sur lui même dans une direction perpendiculaire à son panneau solaire. Le satellite réalise des observations pendant environ une heure puis son axe de rotation est décalé de quelques minutes d'arc. On obtient un décalage d'environ un degré par jour par étapes successives. Cela correspond donc à un tour complet de l'axe de rotation en un an. Ainsi le satellite pointe toujours la même face au soleil et sa face arrière est protégée en permanence du rayonnement du Soleil, de la Terre et de la Lune. Sur le schéma ci-dessus les échelles ne sont pas respectées.

### 4.3.2 Le système cryogénique

Dans un premier temps l'instrument basse fréquence est refroidi à 20K grâce à un système de réfrigération basé sur une détente de Joule-Thomson d'un gaz d' $H_2$  en cycle fermé. Ce système utilise une série de six compresseurs à adsorption identiques, fonctionnant en parallèle chacun dans une phase particulière bien déterminée. À tout moment un compresseur est dans sa phase de désorption chaude, un autre dans sa phase de réchauffement, un autre encore dans sa phase de refroidissement pendant que les trois derniers sont froids et adsorbent le gaz. Cette configuration particulière, permet de s'assurer qu'aucune vibration n'est transmise aux détecteurs. Ce dispositif est également utilisé pour refroidir à 18K le premier étage de l'instrument haute fréquence.

On trouve ensuite un second étage refroidi à 4K par un réfrigérateur Joule-Thomson également

mais utilisant un compresseur mécanique. Le compresseur a été conçu pour limiter les vibrations mais des compensations électroniques sont ajoutées pour contrôler encore plus finement les vibrations résiduelles. Finalement les détecteurs sont portés à 0.1K grâce à un nouveau type de réfrigérateur utilisant une dilution d' $^3\text{He}^4\text{He}$ . Ce mélange subit d'abord une détente de Joule-Thomson qui permet d'obtenir un étage à 1.6K. C'est la dilution des deux isotopes qui est une réaction endothermique qui est utilisée pour arriver à la température de 0.1K. Elle est effectuée à l'intérieur de petits capillaires ce qui lui permet de fonctionner dans l'espace en apesanteur. Ce système de dilution innovant a été testé avec succès et validé dans le ballon Archéops.

L'ensemble de la chaîne cryogénique est résumé dans la figure 4.4. Après dilution et détente le mélange d'hélium est relâché dans le vide. Ce cycle ouvert est donc limité par les réserves initiales d'hélium et constitue le facteur déterminant la durée de vie du satellite. La précision requise pour les mesures de PLANCK nécessite une très grande stabilité de la température du plan focal. Il a donc été fabriqué dans un alliage spécial d'Iridium-Itrimum qui, en plus de satisfaire les contraintes mécaniques, possède une très grande capacité thermique. Cette propriété a été recherchée afin de fortement lisser les variations de température. De plus, plusieurs régulateurs électroniques (PID) sont utilisés dans les différents étages pour compenser en temps réel les variations afin d'obtenir une température encore plus stable.

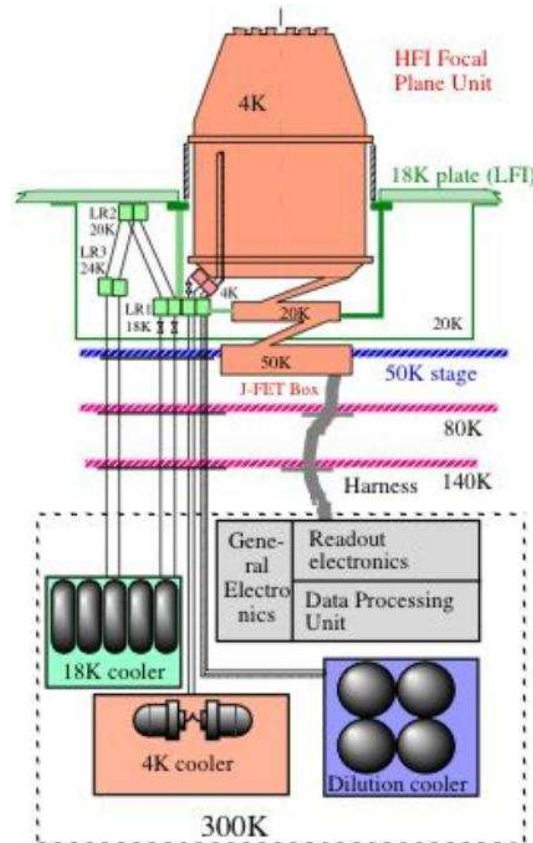


FIGURE 4.4 – **Schéma de la chaîne cryogénique de Planck.** La chaîne cryogénique de PLANCK est constituée d'un premier étage à 20K utilisant un réfrigérateur utilisant une détente de Joule-Thomson avec un compresseur à adsorption, également utilisé pour pré-refroidir à 18K l'étage suivant. Un étage à 4K refroidi par un réfrigérateur utilisant également une détente de Joule-Thomson avec un compresseur mécanique. Puis un étage à 1.6K réalisé par une détente de Joule Thomson d'hélium. Finalement le dernier étage est refroidi à 0.1k en utilisant la propriété endothermique de la dilution d' $^3\text{He}$  et d' $^4\text{He}$

### 4.3.3 Le système optique

Le télescope de PLANCK est un télescope hors-axe, dans une configuration aplanétique. D'autres configurations, qui permettent par exemple de mieux limiter l'astigmatisme sur l'axe optique, existent. Cependant, c'est une configuration aplanétique qui a été retenue car elle donne les meilleures performances compte tenu de la grande surface du plan focal de PLANCK.

Afin de ne pas être limité par la diffraction, étant données la résolution requise de quelques minutes d'arc et les longueurs d'onde d'observation de quelques millimètres, un calcul simple permet de déterminer que la taille minimale du miroir primaire doit être de l'ordre d'un mètre. La figure 4.5 donne une représentation schématique du télescope. Son diamètre projeté moyen est de  $1.5m$ . Il est constitué de deux miroirs elliptiques. Un miroir primaire de  $1.7m$  de diamètre et un miroir secondaire de  $0.85m$  de diamètre. Ils sont fabriqués dans un plastique renforcé par des fibres de carbone, ce qui permet de limiter le poids et leur confère les bonnes propriétés optiques à leur température d'équilibre de  $50K$ . La surface réfléchissante est couverte d'une couche d'aluminium pour assurer une réflectivité supérieure à  $0.995$  aux fréquences d'intérêt. Les défauts présents à la surface des miroirs sont inférieurs au  $\mu m$  et l'erreur sur la forme des miroirs varie entre  $7.5\mu m$  au centre et  $50\mu m$  sur les bords.

Un dispositif composé de cornets, d'un jeu de filtres et d'une lentille, permettant de focaliser le rayonnement, de sélectionner une bande de fréquence et de gaussianiser le lobe, est placé devant chaque détecteur (voir figure 4.6). La largeur à mi-hauteur (FWHM) du lobe principal de l'optique complète (y compris les cornets) pour chaque fréquence d'observation est de quelques minutes d'arc. Les détails par fréquences sont récapitulés dans le tableau 8.4. Les miroirs et le baffle qui les entoure ainsi que les boucliers sont fixés au module de service par une structure en fibre de carbone. En plus d'entrer en jeu dans le refroidissement passif comme on l'a vu précédemment, le baffle qui entoure le télescope sert également à limiter au maximum l'exposition des miroirs à toute lumière parasite, en particulier le rayonnement qui peut être capté par les lobes lointains de l'optique. Finalement, les deux instruments de PLANCK destinés à mesurer le rayonnement du CMB sont placés dans le plan focal du télescope.

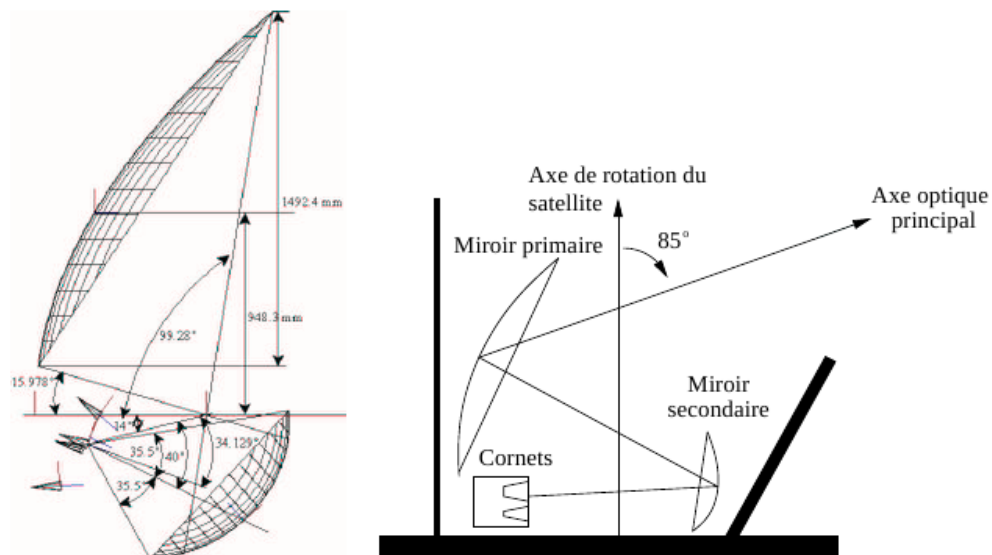


FIGURE 4.5 – **Schéma du télescope de Planck** C'est un télescope dans une configuration aplanétique. La direction vers laquelle il pointe fait un angle de  $85^\circ$  avec l'axe de rotation du satellite. Le miroir primaire a un diamètre moyen de  $1.7$  mètre et le miroir secondaire un diamètre moyen de  $0.85$  mètre.

### 4.3.4 Le senseur stellaire

Le senseur stellaire est un petit télescope à grand champ situé dans le module de service qui pointe dans la même direction que le télescope principal d'observation. Il a deux rôles fondamentaux.

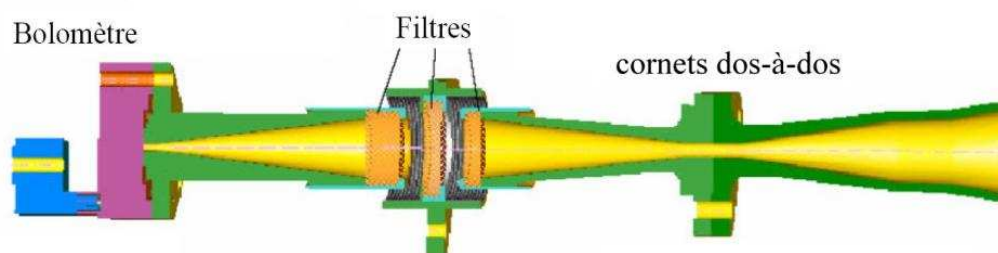


FIGURE 4.6 – **Schéma du dispositif placé avant les bolomètres.** Devant chaque bolomètre un dispositif constitué de cornets placés dos à dos (Back to back horns), suivis par des filtres et lentilles puis d'un dernier cornet, permet de focaliser le rayonnement vers le bolomètre, de sélectionner la bande de fréquence et de gaussianiser le lobe de l'optique.

Le premier est de déterminer en temps réel les mouvements du satellite (notamment la vitesse de rotation sur lui même, les variations de la direction de l'axe de rotation) afin de les asservir. Le second est de permettre de mesurer la direction absolue d'observation des détecteurs.

Pour effectuer ces deux tâches, le senseur stellaire réalise des images à grand champ du ciel. Le déplacement d'un certain nombre d'objets ponctuels présents dans le champ, est alors utilisé pour contraindre le mouvement de rotation. Pour déterminer la position vers laquelle pointe chaque détecteur, deux étapes complémentaires sont nécessaires. La première est de connaître la direction exacte vers laquelle pointe le télescope. On utilise alors les images du senseur stellaire que l'on corrèle avec un catalogue d'objets dont on connaît la position de façon précise (principalement des étoiles de notre Galaxie). Cela permet de déduire la direction d'observation du télescope de façon absolue avec une précision de 0.5 minute d'arc. La seconde étape est de caractériser le plus finement possible la position des détecteurs dans le plan focal du télescope. On se base pour cela sur des sources ponctuelles, assez brillantes dans toute la gamme de longueurs d'onde d'observation, qui après avoir été vues par l'ensemble des détecteurs servent de référence pour déterminer la position relative des différentes voies. Cette étape est appelée la reconstruction du plan focal. La combinaison de ces deux opérations permet de mesurer avec une grande précision la position pointée sur le ciel.

Le senseur stellaire joue donc un rôle essentiel pour la qualité des données, en assurant le contrôle précis du mouvement du satellite et en permettant de connaître la zone du ciel observée par les détecteurs. La qualité du repérage et la stabilité du mouvement sont des éléments fondamentaux notamment pour la production des cartes afin de pouvoir reprojeter les données de façon précise en particulier lorsqu'on combine les observations provenant de plusieurs voies différentes, observant à la même longueur d'onde. Enfin PLANCK a la responsabilité de générer un catalogue d'objets compacts qui nécessite également une grande précision sur la position. Par mesure de sécurité, le satellite embarque donc un second senseur stellaire redondant qui peut prendre le relais de façon automatique en cas de panne.

### 4.3.5 Le SREM

Les rayons cosmiques sont des particules de haute énergie que l'on observe dans tout notre environnement interplanétaire. Ils peuvent atteindre des énergies supérieures à  $10^{21} eV$ . À basse et moyenne énergie, leur flux est très important, puis il diminue avec l'énergie. Lors de leur interaction avec la matière (ou des organismes vivants), ils peuvent déposer localement une grande quantité d'énergie. Ils représentent donc une source importante de problèmes (vieillesse de l'électronique, altération des matériaux ou tissus vivants) pour toutes les applications spatiales. Au niveau du sol, sur Terre, le flux est beaucoup moins élevé que dans l'espace, en partie grâce au champ magnétique terrestre qui joue un rôle de bouclier en déviant les particules mais également car la plupart des rayons cosmiques interagissent avec les molécules ou atomes des hautes couches de l'atmosphère avant d'atteindre le sol.

Beaucoup d'hypothèses restent encore mal contraintes, notamment quant à leur origine, leur distribution, leur processus d'accélération, leur interaction avec le champ magnétique solaire et terrestre, leur composition précise. Dans le but de répondre à ces questions, dans le cadre d'un

programme large d'étude des rayons cosmiques, l'ESA a développé un détecteur de particules pouvant être embarqué sur des expériences spatiales : le SREM pour Standard Radiation Environment Monitor (Buehler et al. (1996)). Il est capable de détecter des protons d'énergie supérieure à 30 MeV et des électrons d'énergie supérieure à 1 MeV et de calculer le spectre en énergie sur 16 intervalles. Par ailleurs, il possède un cône d'ouverture de  $\pm 45^\circ$  afin d'obtenir une information sur la direction des particules.

Le SREM a été conçu dans le but d'être produit en série afin d'être embarqué tel quel sur un grand nombre de satellites et fournir ainsi des données depuis différentes positions dans le système solaire. Cela a justifié les fortes contraintes qui ont été imposées : son faible encombrement, son poids de 2.5 kg, sa consommation électrique de moins de 2W... En 2010, on retrouve un SREM embarqué avec succès sur les missions, Proba-1, Integral, Rosetta, GIOVE-B, Herschel et Planck. Le SREM de PLANCK se trouve dans le module de service.

#### 4.3.6 Communication avec la Terre

Le satellite communique avec le sol par l'intermédiaire d'un transpondeur émettant dans la bande X. L'antenne est alignée le long de l'axe de rotation du satellite. Le signal est capté sur terre par différentes antennes radio, puis est transmis au centre de contrôle de l'ESA. Durant la phase d'observation, la réception des données se fait tous les jours, pendant un créneau de trois heures au cours de la période où le satellite est visible par l'antenne. Les données reçues sont ensuite transférées au centre de traitement principal de l'ESA, le MOC (pour Main Operation Center), situé à Darmstadt en Allemagne. Pendant cette période, des télécommandes peuvent être envoyées au satellite pour changer ses paramètres et interagir avec les différents éléments embarqués. Les données sont stockées dans le satellite après compression. La mémoire interne permet de stocker jusqu'à 48h de données ce qui assure une marge relative en cas d'éventuels problèmes lors de la transmission.

#### 4.3.7 Stratégie de pointage

La façon dont un détecteur observe le ciel est primordiale pour une expérience visant à étudier le CMB. Plus la surface de ciel couverte est grande plus la variance sur l'estimation du spectre de puissance de température et polarisation est faible. Ainsi, il faut essayer d'avoir la couverture angulaire la plus large possible. Pour cela, des observations depuis un satellite sont idéales.

Par ailleurs, même si on suppose le bruit constant durant toute la durée de la mission, si la couverture du ciel n'est pas homogène, les endroits où le ciel a été observé plus de fois, auront un bruit plus faible car d'autant plus moyenné. Dans la stratégie de pointage, il faut donc prendre en compte l'homogénéité de la couverture. Finalement, pour mesurer la polarisation du rayonnement, on fait des mesures de polarisation linéaire dans différentes directions. Plus la répartition des angles de mesure est élevée et plus la précision sur la détermination de la polarisation est importante. Ainsi, une stratégie de pointage doit également optimiser la répartition de la direction de mesure des angles de polarisation. C'est en prenant en compte tous ces paramètres que la stratégie de pointage de PLANCK a été choisie.

Le satellite observe donc le ciel en tournant sur lui même dans une direction perpendiculaire aux panneaux solaires. Le télescope pointe à  $85^\circ$  de l'axe de rotation (voir figure 4.5). La vitesse de rotation est d'environ un tour par minute. Les observations effectuées au cours de chaque tour sont appelées des *cercles*. Régulièrement, après une période d'un peu moins d'une heure, la direction de l'axe de rotation est décalée d'environ 2,5 minutes d'arc. Cette manoeuvre est appelée dépointage. L'ensemble des cercles, entre deux périodes de dépointage, observe donc la même zone du ciel. On se réfère généralement à ces groupes de cercles sous le nom d'anneau (ou *ring* en anglais).

La direction de l'axe de rotation du satellite est décalée d'à peu près un degré par jour au total et fait donc un tour en un an. Le fait de repasser un grand nombre de fois sur la même région du ciel, dans chaque *ring*, avant de changer la direction de l'axe de rotation permet de diminuer le bruit de la mesure et permet aussi de mettre en évidence des effets systématiques qui varient dans le temps sur des échelles plus petites que l'heure. C'est donc un élément essentiel de la stratégie de pointage.



## 4.4 L'instrument basse fréquence (LFI)

L'instrument basse fréquence (LFI pour Low Frequency Instrument) de PLANCK utilise une série de radiomètres micro-ondes pour mesurer l'intensité et la polarisation linéaire du rayonnement du CMB. C'est cette technologie qui a également été utilisée pour l'instrument DMR du satellite Cobe puis pour le satellite WMAP. Les détecteurs de LFI représentent donc la troisième génération de radiomètres embarqués sur un satellite afin d'étudier le CMB. Par rapport aux précédents détecteurs, de bien meilleures performances ont été atteintes aussi bien en terme de sensibilité, de stabilité, que de gestion du bruit et des effets systématiques.

Dans le satellite PLANCK, les radiomètres de LFI sont situés au foyer du télescope et opèrent à une température de  $20K$ . Des cornets fixés au plan focal du télescope servent à collecter les photons vers les détecteurs et à sélectionner le rayonnement dans une bande de fréquence déterminée. Au foyer de chaque cornet, on trouve deux radiomètres, sensibles uniquement au rayonnement polarisé linéairement dans deux directions respectives perpendiculaires. Les différents détecteurs de LFI sont répartis sur le pourtour du plan focal et entourent ceux de l'instrument HFI qui sont placés au centre (voir figure 4.7).

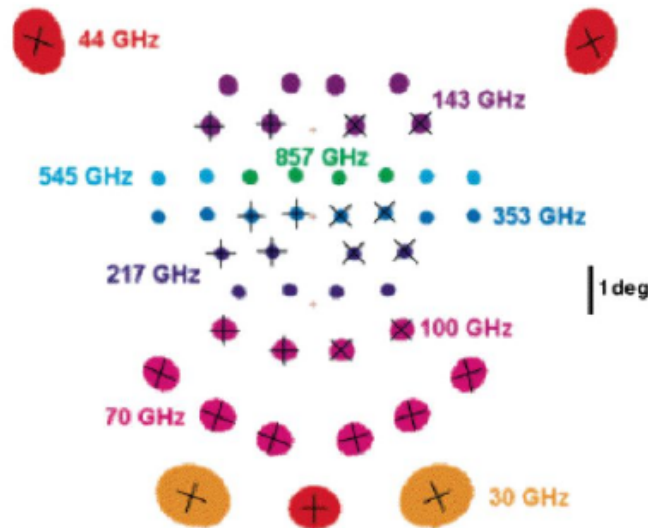


FIGURE 4.7 – **Schéma du plan focal de Planck.** Les bolomètres de l'instrument HFI se trouvent au centre. Les radiomètres de l'instrument LFI sont répartis sur le bord du plan focal. La direction de polarisation des bolomètres PSB et des radiomètres est représentée par des traits noirs.

L'instrument LFI fait des observations dans des bandes de fréquence centrées en 30 GHz, 44 GHz et 70 GHz avec des bandes passantes respectives de 6 GHz, 8.8 GHz et 14 GHz. Les radiomètres utilisés sont sensibles à des variations relatives de quelques  $10^{-6}$  aussi bien en terme d'intensité totale que de polarisation et leur sensibilité est inférieure au  $mKHZ^{-1/2}$  pour l'ensemble des voies. Finalement, la résolution angulaire respective des différents canaux est de 33, 24 et 14 minutes d'arc.

Ces performances remarquables sont assurées par un système complexe de récepteurs différentiels associés à des amplificateurs à très bas bruit fonctionnant à des températures cryogéniques. Le système différentiel combine le signal du ciel avec le signal d'un corps noir de référence. Cette mesure permet de s'affranchir du bruit en  $1/f$  induit par les variations de gain et les fluctuations de température dans les amplificateurs. On peut montrer que les performances d'un tel système sont d'autant meilleures que la température de la référence est proche de celle du signal. LFI utilise donc une référence placée sur l'étage à  $4K$  de l'instrument HFI. Par ailleurs, pour limiter la dissipation de chaleur dans le plan focal refroidi à  $20K$ , les radiomètres sont divisés en deux groupes indépendants dont le signal amplifié est transmis, grâce à des guides d'ondes, à la partie électronique secondaire déportée dans un étage supérieur du satellite qui fonctionne à une température de  $300K$  (voir figure 4.8). LFI possède au total vingt-deux voies distinctes dont quatre sont dédiées aux observations à

30GHz, six à 44GHz et douze à 70GHz.

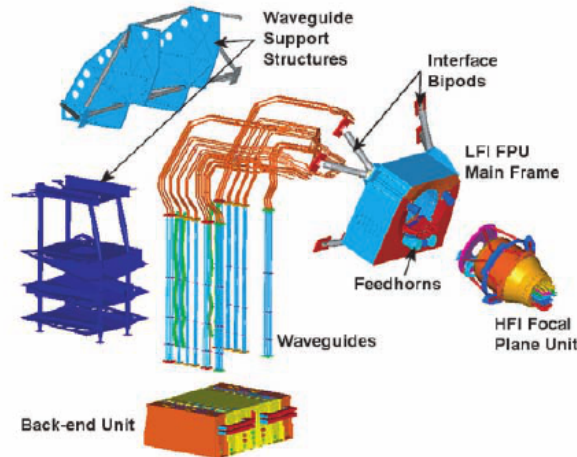


FIGURE 4.8 – Schéma des détecteurs de l'instrument LFI. Les détecteurs de LFI sont situés sur le bord du plan focal du télescope. Les détecteurs de HFI sont au centre. On peut voir les guides d'onde qui conduisent le signal des radiomètres de la partie cryogénique à 20K, vers la partie électronique déportée dans le module de service qui fonctionne à 300K.

Parmi toutes les bandes de fréquence de PLANCK, on s'attend à ce que celle à 70GHz soit la moins contaminée par les émissions des avant-plans polarisés (principalement l'émission de la poussière galactique). Elle est donc d'une importance particulière pour la mesure de la polarisation du CMB. C'est pour cela que ces voies sont les plus nombreuses sur l'instrument LFI. Par ailleurs, il était également prévu au départ que l'instrument LFI possède des voies à 100GHz qui auraient été redondantes avec un des canaux de l'instrument HFI. Cela aurait permis de faire des comparaisons directes aux mêmes fréquences avec des détecteurs utilisant des technologies différentes et donc de mettre en évidence ou de caractériser des effets systématiques éventuels, présents sur un seul des deux types de détecteurs. Cependant ces voies ont été supprimées en 2003 pour des raisons de budget.

## 4.5 L'instrument haute fréquence (HFI)

L'instrument haute fréquence (HFI pour High Frequency Instrument) de PLANCK est composé d'une série de détecteurs utilisant des bolomètres pour mesurer l'émission du CMB, deux instruments cryogéniques qui permettent de refroidir les bolomètres jusqu'à une température de 0.1K et un ordinateur, appelé DPU pour le terme anglais *Data Processing Unit* (Unité de traitement des données), qui gère la chaîne électronique de mesure et le traitement après numérisation (en particulier la compression), qui commande les instruments cryogéniques et qui communique avec l'ordinateur central du satellite.

Les détecteurs de l'instrument HFI se situent dans la partie centrale du plan focal du télescope de PLANCK et sont entourés par les radiomètres de LFI (voir figure 4.7). Comme pour les détecteurs de LFI, des cornets, des filtres et des lentilles guident les photons et sélectionnent leur longueur d'onde avant leur arrivée sur les bolomètres (voir figure 4.6). HFI fait ses observations dans six bandes de fréquence, centrées en 100GHz, 143GHz, 217GHz, 353GHz, 545GHz et 857GHz.

Les bolomètres permettent de réaliser des mesures avec une très grande sensibilité, comme démontré sur des expériences telles que BOOMERanG ou ARCHEOPS. À très basse température, la sensibilité des bolomètres utilisés dans PLANCK augmente quand la température diminue. Afin d'obtenir le niveau de précision nécessaire pour atteindre les objectifs de la mission, les bolomètres de PLANCK fonctionnent donc à une température de 0.1K. Cela permet d'atteindre une sensibilité de quelques  $\mu K$ . Cette faible température entraîne également une forte diminution du bruit thermique ce qui améliore la précision des mesures. Le refroidissement est assuré par un ensemble de plusieurs

systèmes cryogéniques décrits dans la partie 4.3.2.

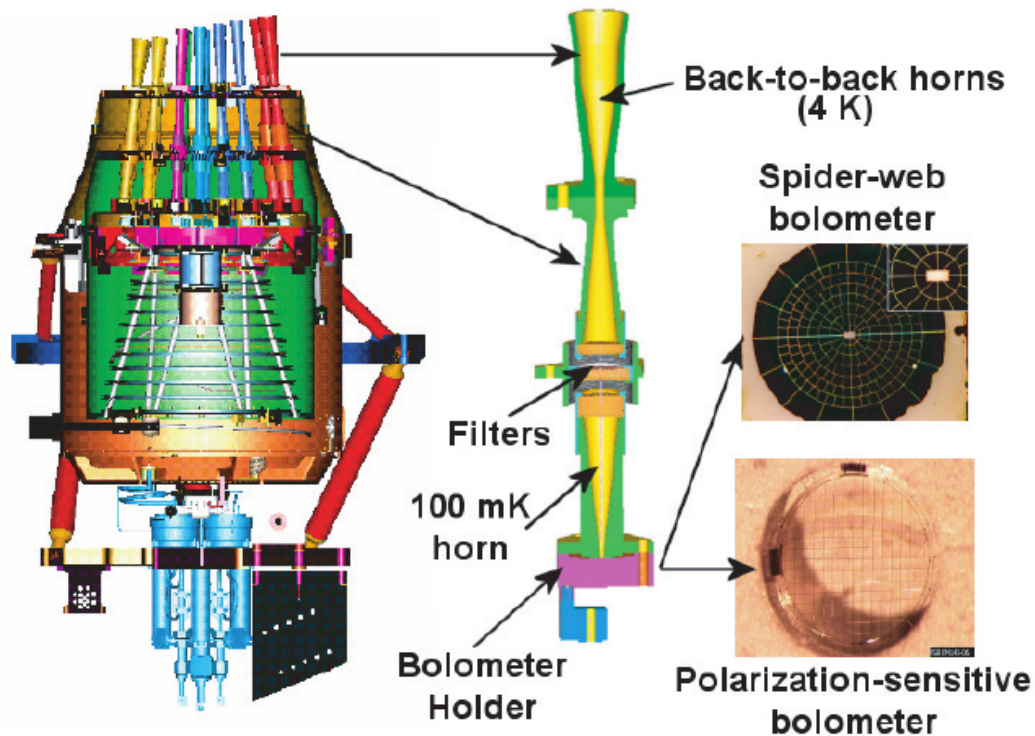


FIGURE 4.9 – Schéma des détecteurs de l'instrument HFI. Sur la gauche on peut voir la partie cryogénique de HFI. On distingue, dans la partie supérieure, l'ensemble des cornets fixés sur l'étage à 4K, le système de dilution en spirale vers le plan focal. Le schéma du centre représente l'ensemble des filtres et cornets permettant de focaliser le rayonnement et de sélectionner la bande de fréquence. Sur la droite on trouve des photos des deux types de bolomètres. En haut, les bolomètres de type spiderweb et en bas, ceux de type PSB.

Les bolomètres sont placés à la base des cornets. Ils sont constitués d'un absorbeur, qui est chauffé par le rayonnement et d'une thermistance qui permet de mesurer la température de l'absorbeur et d'en déduire la puissance déposée. L'absorbeur est en contact thermique permanent avec un réservoir à température constante ce qui lui permet d'évacuer l'énergie au fur et à mesure et de ne pas chauffer indéfiniment. Comme on peut le voir sur la droite de la figure 4.9, l'absorbeur est constitué d'une série de fils. Il existe deux topologies différentes qui sont à la base de deux catégories de bolomètres, les bolomètres de type *spiderweb* et PSB (Polarisation Sensitive Bolometer). Les fils des absorbeurs des *spiderwebs* sont disposés en forme de toile d'araignée, d'où leur nom. Ils absorbent le rayonnement quelque soit sa polarisation. Les bolomètres de type PSB ont des fils uniquement dans deux directions perpendiculaires. Seule la composante du rayonnement dont la direction de polarisation est parallèle au fil va être absorbée. Cela permet de mesurer deux directions de polarisation perpendiculaires. En repassant sur la même position du ciel avec des orientations différentes, on peut alors reconstruire totalement la polarisation. Seules les bandes de fréquences jusqu'à 353GHz possèdent des bolomètres polarisés. Avec ses deux instruments, PLANCK est donc capable d'observer la polarisation linéaire du rayonnement dans les bandes de 30GHz à 353GHz. Aucune voie n'est sensible à la polarisation circulaire mais les modèles ne prédisent pas de mécanismes pouvant produire ce type de polarisation.

L'ensemble absorbeur et thermistance mesure seulement quelques millimètres, ce qui est relativement compact. HFI embarque au total 20 bolomètres de type *spiderweb* et 16 de type PSB (4 PSB à 100GHz, quatre PSB et quatre spiderweb à 143GHz, 217GHz et 353GHz et quatre *spiderweb* à 545GHz et 857GHz). Leur répartition dans le plan focal est détaillée dans la figure 4.7. Les PSB nécessitent une voie de lecture pour chaque composante de polarisation. Cela donne donc 52 voies de lecture pour les détecteurs, réparties sur les différentes fréquences d'observation. À cela il faut ajouter un certain nombre de voies dédiées au suivi du bon fonctionnement de l'instrument

(suivi de la température des différents étages, suivi du vieillissement de l'électronique, suivi des effets systématiques...). Pour cela deux bolomètres supplémentaires, de type PSB, sont présents mais sont protégés de tout rayonnement. Cela permet de mesurer en temps réel une partie des variations de réponse des bolomètres, indépendantes du signal. On trouve également un ensemble de 8 thermomètres répartis sur les étages à 0.1K, 1.6K et 4K ainsi que 8 thermomètres permettant de suivre l'évolution des PID et finalement une mesure d'une résistance et d'une capacité de test est effectuée pour suivre l'évolution de l'électronique. Cela donne un total de 72 voies. L'électronique de lecture des différentes voies est séparée en douze ceintures indépendantes de six voies chacune. Les voies de chaque ceinture sont distribuées entre les voies à différentes fréquences et les voies de maintenance. Cela permet en cas de dysfonctionnement d'une ceinture de ne pas perdre l'ensemble des observations dans une bande de fréquence donnée.

Le détail de l'électronique de mesure des bolomètres est donné dans la figure 4.10. On peut distinguer trois parties principales : le boîtier JFET, l'unité de pré-amplification et l'électronique de lecture. Le boîtier JFET est un élément essentiel de l'électronique bas bruit de PLANCK. Il permet de réaliser une adaptation d'impédance afin de passer dans la suite de la chaîne électronique dans une régime à faible impédance qui rend plus facile la gestion du bruit. Le boîtier JFET fonctionne à la température de 100K qui est celle qui minimise le bruit des composants JFET. Ce boîtier contient également une partie de l'électronique de polarisation des bolomètres. L'étage suivant, l'unité de pré-amplification (PAU pour Pre-Amplification Unit), a pour rôle principal d'amplifier le signal par un facteur 1000. Finalement le dernier élément de la chaîne de lecture le REU (Readout Electronic Unit) s'occupe de l'échantillonnage et de la conversion analogique numérique du signal. Il génère aussi les tensions triangles qui sont utilisée pour générer le courant de polarisation des bolomètres et s'occupe de la synchronisation des différents sous-systèmes. Il embarque enfin un FPGA qui réalise le traitement numérique du signal (détection synchrone, filtrage...).

L'électronique de polarisation et de lecture des bolomètres a été optimisée afin d'être capable de détecter les faibles variations du CMB de l'ordre de  $10^{-4}$  à  $10^{-6}$  par rapport au niveau moyen (voir Piat (2000)). En particulier, les bolomètres sont polarisés par un courant carré qui permet d'obtenir un signal modulé autour d'une valeur moyenne. L'électronique de lecture des bolomètres échantillonne le signal 90 fois par période du courant de polarisation des bolomètres. La quantification se fait sur 16bits. La stratégie retenue pour obtenir les mesures brutes de chaque voie, consiste à moyennner ces échantillons sur chaque demie-période du courant de polarisation. On obtient alors 2 échantillons bruts, modulés, par période. Le signal mesuré est donc modulé autour d'une valeur de référence. On appelle  $DSN^+$  et  $DSN^-$  respectivement les valeurs moyennées dans la demi-période supérieure et inférieure à la ligne de base. La fréquence de modulation est proche de 90Hz, on obtient donc une fréquence d'échantillonnage pour les DSN deux fois plus élevée, ce qui donne pour PLANCK une fréquence d'échantillonnage exacte de 180.38Hz. Afin de travailler sur des nombres entiers, les DSN sont calculés en pratique, uniquement comme la somme des échantillons et ne sont pas divisés par le nombre d'échantillons qui les compose. Ils ne représentent donc pas exactement une moyenne. Le nombre d'échantillon de la demi-période qui sont effectivement utilisés pour le calcul des DSN est un paramètre qui a été optimisé (voir Montier (2005)) mais qui vaut au maximum 45 (car il y a 90 échantillons par période). On voit que la somme de 45 éléments codés sur 16bits nécessite donc d'être codée sur un minimum de 22bits. Afin de minimiser le débit de données, on utilise le moins de bits possible. Finalement, chaque échantillon de PLANCK, à la sortie de l'électronique de lecture (REU), est donc un DSN codé sur 22bits.

L'ensemble des éléments de l'instrument HFI est coordonné par le DPU (Data Processing Unit). C'est l'élément central qui permet de contrôler et de faire fonctionner ensemble les différents sous-système de HFI. Il est chargé de la communication avec l'ordinateur central, du traitement des télécommandes qui lui sont transmises par l'ordinateur central, de la gestion des sous systèmes, en particulier l'électronique de lecture des bolomètres, l'électronique du réfrigérateur à 0.1K, l'électronique du compresseur, du réfrigérateur à 4K... Finalement, le DPU est chargé de traiter les DSN à la sortie de l'électronique de lecture. Il doit notamment gérer la compression et la mise en forme des données sous forme de paquets standardisés qui vont être envoyés à l'ordinateur central pour être stockés et envoyés à la Terre pendant les périodes de communication (voir section 5). Le protocole impose des règles de traitement strictes. Après la réception de chaque télécommande, le DPU doit signaler la bonne réception. Si la télécommande concerne un sous-système, comme le REU ou le compresseur à 4K, le DPU doit transmettre la télécommande, indiquer si elle a été reçue ou non,

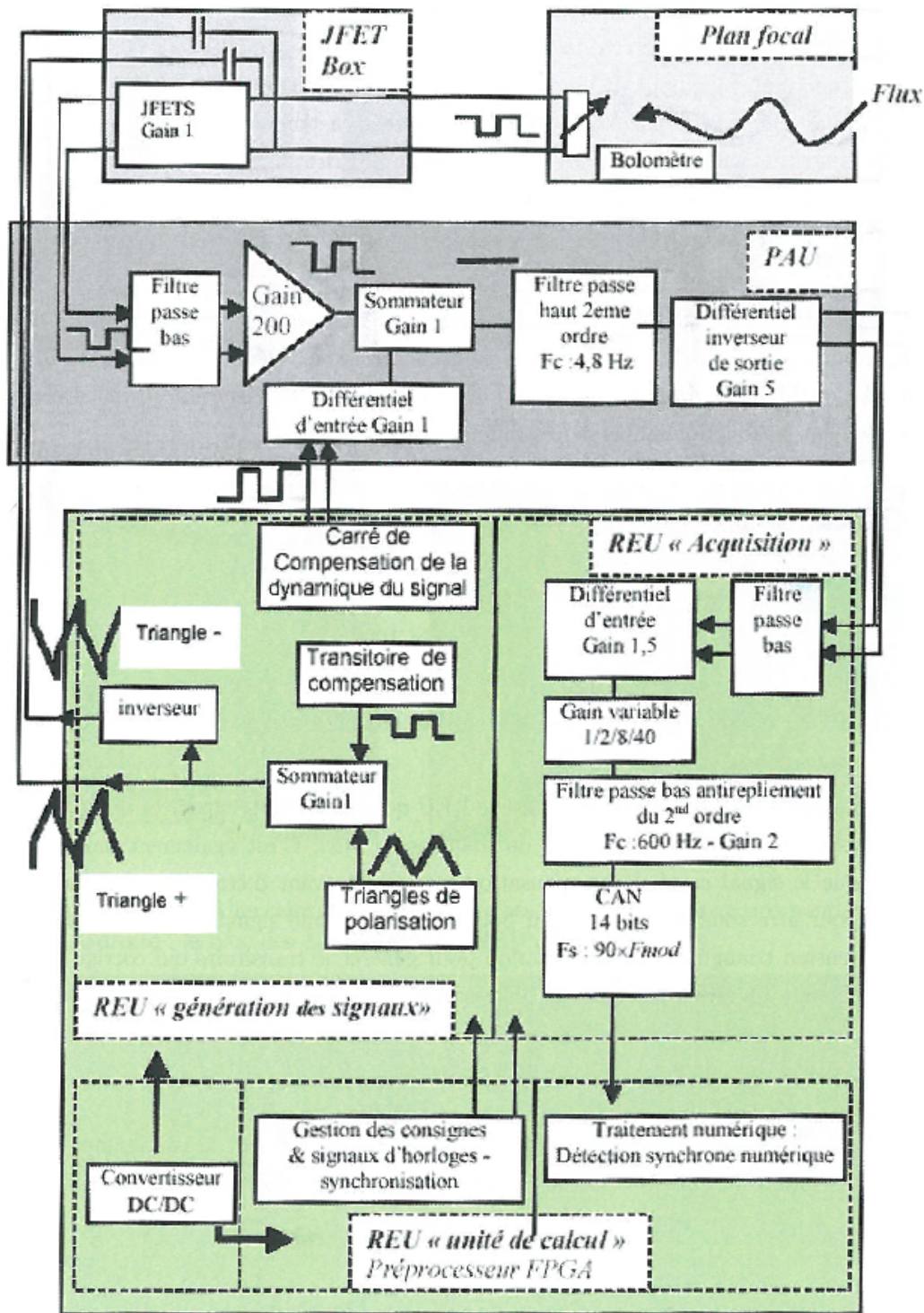


FIGURE 4.10 – Schéma de l'électronique de lecture et de polarisation des bolomètres. Elle est composée de trois parties principales : le boîtier JFET, l'étage de préamplification (PAU) et l'unité d'acquisition (REU) qui est également chargée de générer l'intensité de polarisation des bolomètres.

signaler lorsque la télécommande a été exécutée et donner l'état à la fin de l'exécution de la télécommande. Si la télécommande doit être exécutée par le DPU, il doit l'exécuter, signaler la fin de

l'exécution et envoyer l'état de l'exécution.

Le lancement du satellite a été effectué le 14 mai 2009, par une fusée Ariane 5, depuis le centre spatial de l'ESA à Kourou. Le rôle du centre spatial s'arrête après la mise sur sa trajectoire du satellite. La gestion et le contrôle général du satellite sont ensuite effectués par le MOC (Main Operation Center) de l'ESA, situé à Darmstadt en Allemagne. C'est le MOC qui est chargé de la gestion de la trajectoire, la gestion du pointage, la rotation du satellite sur lui-même, des communications. Les instruments sont eux sous la responsabilité des laboratoires qui les ont conçus (l'IAS pour HFI et le CNR de Milan pour LFI). Ce sont eux qui sont chargés de s'assurer de leur bon fonctionnement. La commande à distance du satellite et de ses composants se fait par l'intermédiaire de télécommandes. Elles sont envoyées pendant la période quotidienne de communication avec le satellite. L'envoi de télécommandes suit un protocole standardisé, strict. Seuls les opérateurs du MOC sont autorisés à envoyer des télécommandes. Lorsqu'il est nécessaire d'interagir avec un des instruments pour changer des paramètres par exemple, le laboratoire responsable doit en faire la demande au MOC qui se chargera de générer les télécommandes. Les télécommandes sont reçues par l'ordinateur central qui les exécute ou les fait suivre aux instruments si elles leur sont dirigées.

Sur Terre, le traitement des données est organisé en quatre niveaux :

**niveau L1** : il est chargé de traiter les paquets télémétriques bruts reçus directement du satellite pour en extraire les données scientifiques et les données de maintenance. Les données scientifiques doivent être décompressées. Le niveau L1 formate finalement les données et les injecte dans la base de données dans un format standardisé accessible aux niveaux supérieurs.

**niveau L2** : il effectue le traitement des données brutes ordonnées en temps (notamment le traitement du bruit, la démodulation...) et la génération des cartes par fréquence (calibration, projection...)

**niveau L3** : son rôle est de fournir les cartes des différentes composantes astrophysique émettant aux fréquences de PLANCK, de fournir des catalogues d'objets compacts et de déterminer les mesures cosmologiques.

**niveau LS** : il ne réalise pas directement le traitement des données réelles. C'est un niveau de simulation des instruments et du ciel nécessaire aux études des effets systématiques. Il participe donc à la compréhension et au traitement des effets instrumentaux.



## Chapitre 5

# Compression des données de l'instrument HFI à bord du satellite

### 5.1 Position du problème

#### 5.1.1 Contraintes liées aux expériences spatiales

La conception et la réalisation d'une expérience spatiale nécessitent la prise en compte de problèmes particuliers, souvent inexistantes pour des observations depuis la Terre. En effet, l'ensemble des composants électriques, électroniques et mécaniques d'un satellite, subit des contraintes très fortes, notamment à cause des vibrations et de l'accélération subies lors du décollage, mais aussi de par le bombardement permanent et intense de rayons cosmiques qui dépose des grandes quantités d'énergie et qui dégrade les matériaux. Des composants spécifiques ayant obtenus une qualification spatiale doivent donc être utilisés. Ils nécessitent parfois l'utilisation de matériaux ou de techniques de production optimisés. Dans tous les cas, ils doivent obéir à des standards et subir une longue série de tests, de compatibilité, de fiabilité ou de résistance aux contraintes, afin d'obtenir leur qualification, avant de pouvoir être utilisés dans une expérience spatiale. C'est un processus long et coûteux qui explique la grande différence de fonctionnalité et de puissance des éléments embarqués dans un satellite comparés à ceux employés sur Terre.

L'ensemble de ces problèmes liés au lancement et à l'environnement interplanétaire, ajoutés à l'impossibilité de pouvoir intervenir physiquement pour résoudre une défaillance matérielle entraîne l'ajout systématique de sécurités supplémentaires lors de la conception. On trouve par exemple un grand nombre de composants possédant un exemplaire redondant, pouvant prendre le relais en cas de panne. Les programmes informatiques obéissent également à des règles spéciales. Ils prennent notamment en compte, par exemple, la possibilité de corruption de la mémoire. D'une manière générale, l'ensemble des opérations est conçu en prenant en compte l'éventualité d'un problème temporaire.

#### 5.1.2 Chaîne de transmission des données

C'est à la suite d'un compromis entre le coût de réalisation, la faisabilité et la prise en compte des sécurités supplémentaires liées aux expériences spatiales que l'ensemble de la chaîne de transmission et de stockage a été conçu. On peut voir les principaux éléments de cette chaîne sur la figure 5.1.

Le transpondeur se charge des communications avec la Terre. Le signal est capté au sol, par de larges antennes radio puis transmis au MOC (Main Operation Center de l'ESA). La réception par une antenne donnée, n'est possible que lorsque le satellite entre dans son champ de vision. Ainsi, la localisation de l'antenne utilisée pour chaque communication change en fonction des disponibilités, de la période de l'année et de l'heure. Dans la phase normale d'observation, la communication avec le satellite a lieu une fois par jour et permet de recevoir l'ensemble des données collectées depuis le dernier échange. Une courte période donne par ailleurs accès aux mesures en temps réel. Le débit



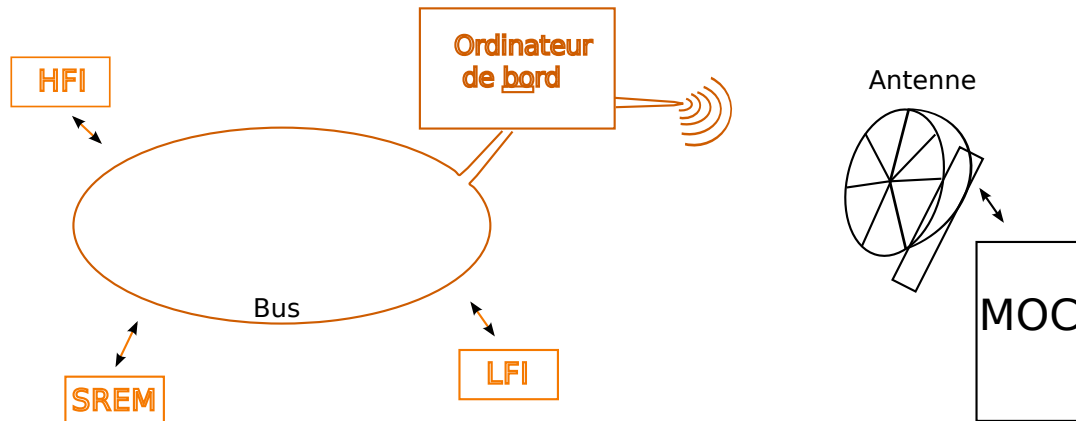


FIGURE 5.1 – **Schéma de la chaîne de transmission des données.** Les données provenant des instruments sont empaquetées puis transmises à l'ordinateur de bord par le bus. C'est le DPU qui est chargé de générer les paquets pour l'instrument HFI. Les données sont ensuite transmises par le transpondeur, vers une antenne située sur Terre, puis vers le MOC (Main Operation Center de l'ESA).

de la connexion est prévu pour pouvoir transmettre les informations en trois heures environ. Pour s'assurer de ne perdre aucune information en cas de problème de communication, il a été imposé de pouvoir stocker, à bord du satellite, l'ensemble des informations à transmettre, pendant une période de 48 heures. Cela assure une certaine marge pour résoudre un éventuel problème avant la période suivante de transmission ou pour libérer une antenne pour communiquer en urgence avec le satellite, sans compromettre la prise de données. En dehors des périodes de communication il est impossible de connaître l'état du satellite. Ainsi, une large série de tests est également effectuée de façon systématique, au cours de chaque échange, afin de s'assurer du bon fonctionnement global du satellite.

Entre les périodes de communication, les données sont stockées dans l'ordinateur de bord. Son espace de stockage est partagé entre les différents sous systèmes (HFI, LFI, SREM...). L'ensemble des instruments communique avec l'ordinateur de bord par l'intermédiaire d'un bus industriel (Bus 1553). Les données doivent être formatées sous forme de paquets standardisés, appelés paquets PUS (Paquet Utilisation Standard), avant de pouvoir transiter sur le bus.

### 5.1.3 Transport des paquets

Chaque paquet contient un en-tête spécifique d'une taille fixe de 8 octets, qui permet de vérifier l'intégrité du contenu, qui identifie le type de données transportée ainsi que la date et l'heure de création et la taille totale du paquet. Les données utiles suivent l'en-tête et sont organisées de façon normalisée et spécifique au type de paquet. La taille totale, en-tête compris, est limitée à 1024 kilo octets par paquet. Le bus peut transmettre au maximum 80 paquets par seconde. Sa bande passante limitée est distribuée de façon stricte entre les différents éléments. Chaque sous-système se voit allouer un nombre de paquets par unité de temps limité.

Pendant la prise de données scientifiques, le DPU est placé dans le mode observation. Dans ce mode, le DPU est chargé d'empaqueter les données issues des détecteurs de HFI ainsi que des données de maintenance, puis d'envoyer les paquets à l'ordinateur central par le bus. L'instrument HFI possède au total 52 voies issues des bolomètres d'observation, ainsi que 2 voies provenant de bolomètres de tests placés à l'abri de toute lumière. On trouve en plus 18 voies supplémentaires associées à une résistance et une capacité de test et à des thermomètres qui contrôlent la température des différents étages cryogéniques (voir section 4.5). L'ensemble des données issues de ces 72 voies est empaqueté en continu et envoyé vers l'ordinateur de bord, dans un type de paquets particuliers appelés « paquets science ». Les paquets de maintenance sont appelés « paquets HSK » (pour le terme anglais housekeeping). Ils sont générés à intervalle régulier et contiennent des données permettant de suivre l'évolution de l'état de l'instrument, notamment le statut des sous-système (l'étage cryogénique à 4K, le REU...), de leur programme, de leur mémoire, la valeur de leurs paramètres, leur

température, des mesures de chaleur dissipée, des mesures de puissance des principaux composants électriques... Au total ce sont plus de 1300 paramètres qui sont regroupés dans neuf paquets HSK1 à HSK9. La fréquence d'émission est variable en fonction du paquet HSK considéré. Les paquets de type HSK1 sont envoyés toutes les 2.4 secondes, les HSK3 toutes les 14 secondes, les autres HSK sont envoyés toutes les minutes. Ces paquets de maintenance sont prioritaires. En cas de problème, ce sont eux qui sont transmis en priorité au bus car leur rôle est de permettre de suivre de façon très précise l'évolution de l'ensemble des composants et sont donc nécessaires à la compréhension des problèmes.

#### 5.1.4 Gestion du débit de données

Pour être certain qu'aucun des sous-systèmes ne dépasse l'espace disponible qui lui est alloué dans l'ordinateur central, leur flux de données est limité. Le débit total du DPU est ainsi fixé à 80koctet/s en moyenne. Par ailleurs, le DPU, est autorisé à envoyer un maximum de 15 paquets par seconde sur le bus, ce qui permet de transmettre 120kbits/s au maximum, en-tête compris, si tous les paquets sont totalement remplis et qui est suffisant pour le débit qui lui est alloué.

Pour ne pas dépasser la limite de 80kbits/s en moyenne, une contrainte logicielle a été ajoutée dans le DPU. Son rôle est d'arrêter l'envoi des « paquets science » si le débit total de tous les types de paquets intégré depuis un temps donné est trop important. Le débit est intégré périodiquement sur des cycles d'environ 45 minutes correspondant approximativement à la durée moyenne d'un *ring*. La durée des cycles est un paramètre et peut être modifié. Si le débit intégré, mesuré depuis le début du cycle d'intégration, dépasse le débit total attendu pour le cycle complet, alors, aucun paquet scientifique n'est envoyé jusqu'à la fin du cycle. Cela autorise à avoir temporairement une augmentation du débit tout en empêchant de saturer la mémoire de l'ordinateur central. Toutefois, même si une augmentation du débit est envisageable sur une courte période par rapport au débit moyen alloué au DPU, le nombre maximal de paquets ne peut jamais dépasser la limite fixée à 15. Dans le cas extrême où les données ne peuvent pas être transmises sur ces 15 paquets, le DPU possède un tampon en mémoire qui lui permet de stocker jusqu'à 32 paquets. Au-delà de cette limite, les paquets contenus dans le tampon depuis le plus de temps seront perdus.

Le mécanisme précédent est relativement souple. Il permet de limiter les données perdues en cas d'augmentation du débit, il permet une augmentation temporaire du débit et surtout il empêche de perdre les données sur une zone particulière du ciel. En l'absence de cette contrainte, si le débit était trop élevé pendant un temps trop long, après avoir rempli le quota de mémoire dans l'ordinateur de bord, on perdrait toutes les informations sur la fin de l'observation avant le prochain transfert vers le sol. Grâce à ce système, dans le cas d'une augmentation trop importante du débit sur une longue période de temps, les pertes sont réparties sur chaque cycle couvrant l'ensemble des observations entre deux communications avec la Terre. Par ailleurs, grâce à la redondance dans la stratégie de pointage (voir section 4.3.7), les conséquences de ces pertes, réparties dans le temps, sont d'autant moins importantes.

#### 5.1.5 Compression des données

Le débit total de 80kbits/s alloué au DPU est partagé entre les « paquets science » et les paquets de maintenance. Ces derniers, nécessitent un débit moyen de 4kbits/seconde. Il reste donc 76kbits/s pour transmettre les « paquets science ». Cependant, comme on l'a vu dans la section 4.5, à la sortie du REU, chaque échantillon du signal est un nombre codé sur 22bits (voir également section 5.3.1). Étant donné le taux d'échantillonnage de 180.38Hz, un débit d'environ 280kbits/s est nécessaire pour transmettre les données issues des 72 voies de l'instrument HFI. Il est donc indispensable de compresser ces données afin de ne pas saturer la bande passante.

Dans la partie utile des paquets PUS, le signal est stocké par tranche de 254 échantillons, une voie après l'autre. Lorsque la taille maximale d'un paquet est atteinte, un nouveau paquet est généré et ainsi de suite jusqu'à la dernière voie. Les données de chaque voie sont précédées par un en-tête spécifique qui identifie la voie, le nombre d'échantillons, qui contient les paramètres de compression... Pour l'ensemble des voies, l'en-tête représente un débit moyen, proche de 11kbits/s. En moyenne, les paquets de maintenance plus l'ensemble des en-têtes nécessitent donc un débit proche de 15kbits/s. Il reste alors, au maximum, environ 65kbits/s pour transmettre les données issues des 72 voies de

l'instrument HFI. L'algorithme de compression doit donc être capable d'atteindre un facteur de compression supérieur à 4.

Le signal mesuré par l'instrument HFI est composé du signal d'origine astrophysique auquel s'ajoute le bruit produit par la chaîne de mesure complète. Le signal astrophysique est la somme de l'émission des différentes composantes : le CMB, l'émission de la poussière, l'émission synchrotron, le rayonnement de freinage, l'émission des sources compactes... (voir section 2.5). Plusieurs composantes à l'origine du bruit de lecture des bolomètres ont été identifiées : le bruit de photon, le bruit de phonon, le bruit de Johnson, du bruit en  $1/f$  (voir par exemple Mather (1982)). Il faut également mentionner le bruit poissonien produit par l'impact des rayons cosmiques. Le développement de l'électronique de lecture de PLANCK a mis l'accent sur la réduction du bruit afin d'obtenir une mesure avec le meilleur rapport signal sur bruit pour parvenir aux objectifs ambitieux de la mission. Durant la phase d'étalonnage de PLANCK il a été montré que le bruit des bolomètre était relativement blanc<sup>1</sup>. On observe cependant une composante en  $1/f$  à des fréquences inférieures au Hertz, qui est responsable de dérives lentes.

À partir de ces résultats, un algorithme de compression a été développé au LAL par Stéphane Plaszczyński et François Couchot et implémenté dans le DPU. Il se base sur l'hypothèse que le signal peut être approximé, sur une courte période, par un bruit blanc gaussien. Le bruit en  $1/f$  est responsable des principales déviations par rapport à un bruit blanc, mais il faut également tenir compte des variations du signal astrophysique notamment l'émission de la poussière galactique et la composante dipolaire du CMB. Cependant si l'échelle de temps considérée est suffisamment courtes, les dérives lentes du bruit en  $1/f$  ainsi que les variations fortes du signal sont suffisamment faibles pour que l'hypothèse précédente représente une bonne approximation. Les données sont donc traitées par tranches de compression dont la durée a été optimisée en tenant compte de ce critère. Par ailleurs l'impact des rayons cosmiques et les sources ponctuelles élargissent la distribution qui n'est plus tout à fait gaussienne. Les données sont compressées avec cet algorithme par le DPU, ce qui permet d'atteindre un débit suffisamment faible pour ne pas saturer la bande passante.

## 5.2 Algorithme de compression

### 5.2.1 Choix de l'algorithme

L'algorithme de compression des données à bord du satellite a été choisi afin de minimiser l'impact de la compression sur le signal d'intérêt scientifique et ses propriétés statistiques, tout en permettant d'obtenir un débit inférieur aux limites imposées par la bande passante et la capacité de stockage de l'ordinateur de bord. Son implémentation tient compte des fortes contraintes matérielles, notamment celles liées à des expériences spatiales : faibles ressources en mémoire vive, faible puissance de calcul, faible bande passante du bus permettant de communiquer avec l'ordinateur de bord, faible capacité de stockage dans l'ordinateur de bord, bande passante et temps de communication avec le sol limités, traitement en temps réel... L'algorithme de compression se base sur la quantification en amplitude d'un signal supposé gaussien et de moyenne nulle. Les données sont traitées par tranches de compression de 254 échantillons. Elles sont d'abord démodulées puis leur offset est estimé et soustrait sur chaque tranche de compression. L'amplitude des échantillons démodulés est ensuite quantifiée avec un pas fixe. Chaque échantillon est finalement codé sur un nombre variable de bits en fonction de l'intervalle dans lequel tombe son amplitude après quantification. Les intervalles les plus probables utilisent le nombre de bits le plus faible, puis le nombre de bits utilisé augmente lorsque la probabilité d'obtenir un échantillon dans l'intervalle diminue.

Cet algorithme ne permet pas une compression sans perte. C'est-à-dire que le signal reconstruit n'est pas identique bit à bit à l'original. Cependant un ensemble d'études théoriques Widrow & Kollár (2008) et de simulations (Présentations au PLANCK Core Team de Stéphane Plaszczyński) montrent que l'impact sur les données est connu et contrôlable. Les conséquences de la compression sur les données de l'instrument HFI de PLANCK se résument à celles introduites par la quantification du signal, qui sont des erreurs d'arrondis. La quantification est une opération non linéaire dont les effets ne sont pas toujours intuitifs. Il faut noter toutefois qu'un signal physique doit subir une conversion numérique avant de pouvoir être traité de façon numérique. Ainsi, même en l'absence de compression, les données ont nécessairement subi une première quantification.

1. Note interne de Désert et Macias-Perez : *HFI noise fit*

Les travaux pionniers sur les effets de la quantification d'un signal ont été menés par Widrow (Voir notamment Widrow & Kollár (2008)). Il est à l'origine de plusieurs théorèmes qui permettent de prédire l'influence de la quantification sur les propriétés statistiques d'un signal. Nous allons détailler dans la suite les effets introduits par la quantification d'un signal à valeurs réelles et montrer que l'algorithme de compression n'introduit pas de non gaussianités mais rajoute simplement un bruit de quantification qui a un spectre blanc.

## 5.2.2 Opérateur de quantification

L'opérateur de quantification  $\mathcal{Q}$  agit sur un signal à valeurs réelles  $s$ . L'unité de base de quantification est  $q$ . L'opérateur de quantification produit alors en sortie un signal  $s'$  dont les valeurs varient par pas de  $q$ . L'expression du signal après quantification en fonction du signal d'entrée est donnée par la formule :

$$s' = q \times \left\lfloor \frac{s + q/2}{q} \right\rfloor \quad (5.1)$$

où l'opérateur  $\lfloor \cdot \rfloor$ , représente l'arrondi à l'entier inférieur le plus proche (floor). C'est une opération non linéaire. Après quantification le signal d'origine ne peut plus être déduit de façon exacte du signal de sortie de l'opérateur  $\mathcal{Q}$ . Si on note  $\nu$  l'erreur commise lors de la quantification alors pour tout signal  $s$  on a la relation :

$$\nu = s - s' \quad (5.2)$$

avec  $\nu$  vérifiant  $|\nu| \leq q/2$  pour toute valeur du signal d'entrée. Ainsi l'erreur maximale est bornée par  $q/2$ . Lorsque le pas de quantification devient arbitrairement petit, l'erreur tend donc vers zéro. Sur la figure 5.2, on peut voir la sortie de l'opérateur de quantification en fonction du signal d'entrée. L'erreur associée est tracée sur la figure 5.3.

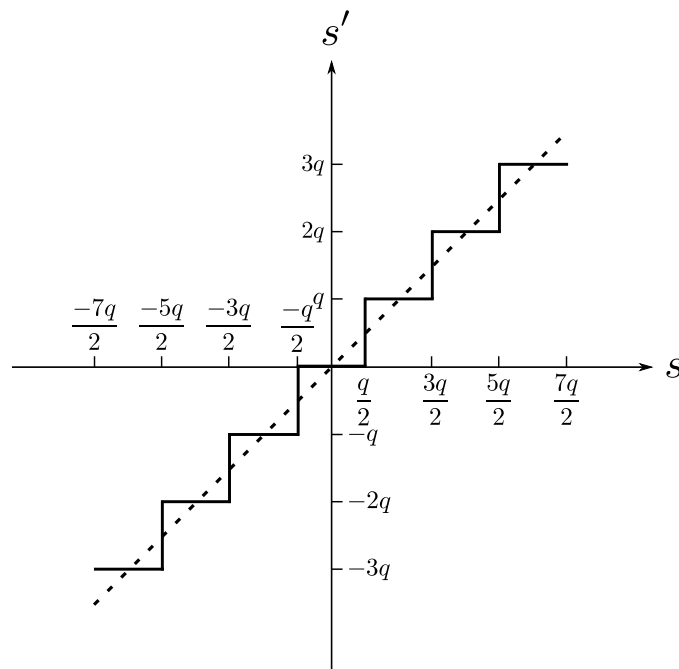


FIGURE 5.2 – **Effet de l'opération de quantification sur un signal à valeurs réelles.** Le signal quantifié varie par pas de  $q$ . C'est une opération irréversible. Il n'est pas possible de restaurer le signal d'entrée à partir du signal de sortie. Par rapport au signal en entrée, l'erreur maximale sur le signal quantifié est  $q/2$ . La distance entre la droite en pointillés et la courbe en marches d'escalier (en traits pleins) est proportionnelle à l'erreur de quantification.

## 5.2.3 Théorèmes de quantification

Il existe une forte analogie entre l'échantillonnage d'un signal et la quantification. En effet, si on considère un signal aléatoire  $s$  et le signal quantifié correspondant  $s'$ , alors, si la densité de

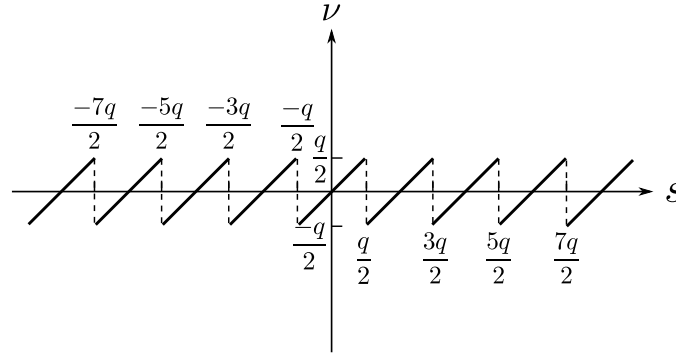


FIGURE 5.3 – **Erreur de quantification sur un signal à valeurs réelles.** Elle varie de façon discontinue et dépend de la valeur du signal non quantifié. Elle est bornée par  $\pm q/2$ .

probabilité de  $s$  est donnée par la fonction  $f_s$ , la probabilité que le signal se trouve dans un des intervalles de quantification  $i$ , est :

$$p_i = \int_{i \times q - q/2}^{i \times q + q/2} f_s(x) dx \quad (5.3)$$

Par définition, après quantification, tous les éléments tombant dans cet intervalle prennent la même valeur  $q_i = i \times q$ . La probabilité  $p_i$  correspond donc également à la probabilité que le signal quantifié prenne la valeur discrète  $q_i$ . La densité de probabilité du signal quantifié prend donc des valeurs discrètes qui valent l'intégrale de la densité de probabilité du signal non quantifié sur l'intervalle de quantification (voir figure 5.4). L'effet de la quantification est alors de réaliser un échantillonnage de la densité de probabilité. On parle généralement dans ce cas d'*échantillonnage continu*, c'est-à-dire que la valeur n'est pas simplement évaluée de façon discrète mais la fonction est intégrée sur toute la largeur de l'échantillon. En gardant cela à l'esprit on peut alors appliquer les lois de l'échantillonnage à la densité de probabilité d'un signal quantifié afin d'estimer l'influence de la quantification sur les propriétés statistiques d'un signal.

Cette analogie a été développée dans Widrow & Kollár (2008) et donne lieu à plusieurs théorèmes de quantification. Si la transformée de Fourier de la densité de probabilité du signal, ou fonction caractéristique,  $\Phi_s(u)$ , est limitée en fréquence, avec :

$$\Phi_s(u) = 0, |u| \geq \frac{\pi}{q} \quad (5.4)$$

alors il est possible de reconstruire la densité de probabilité d'un signal à partir des valeurs quantifiées de ce signal. Cela constitue le premier théorème de quantification. Dans le cas moins restrictif où :

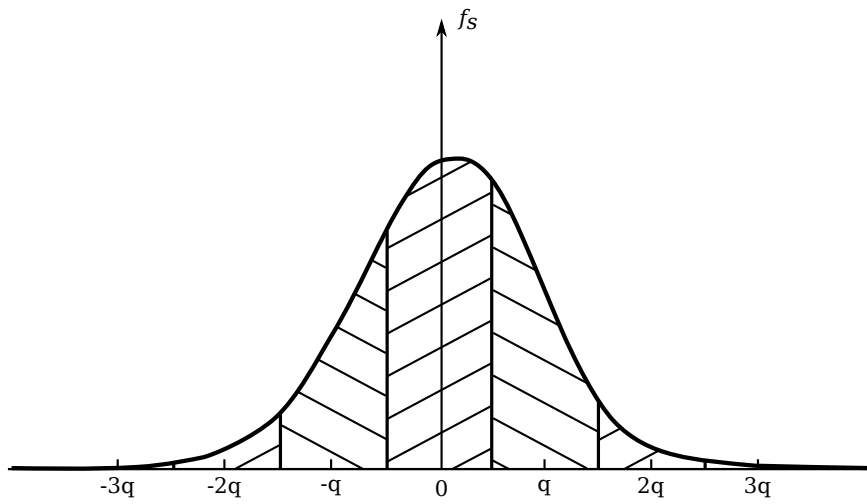
$$\Phi_s(u) = 0, |u| \geq \frac{2\pi}{q} \quad (5.5)$$

on n'est plus capable de reconstruire la densité de probabilité de façon exacte. Cependant, on peut montrer qu'il est toujours possible d'estimer tous les moments de la distribution ce qui donne lieu au deuxième théorème de quantification. Il a été montré que dans ces conditions, l'effet de la quantification pouvait être modélisé par l'ajout d'un bruit uniforme sur l'intervalle  $[-q/2, q/2]$ , indépendant du signal, appelé « pseudo bruit de quantification » (ou PQN pour le terme anglais Pseudo Quantification Noise). Si on note  $n$  ce bruit, alors le signal quantifié,  $s'$  aura les mêmes propriétés statistiques que :

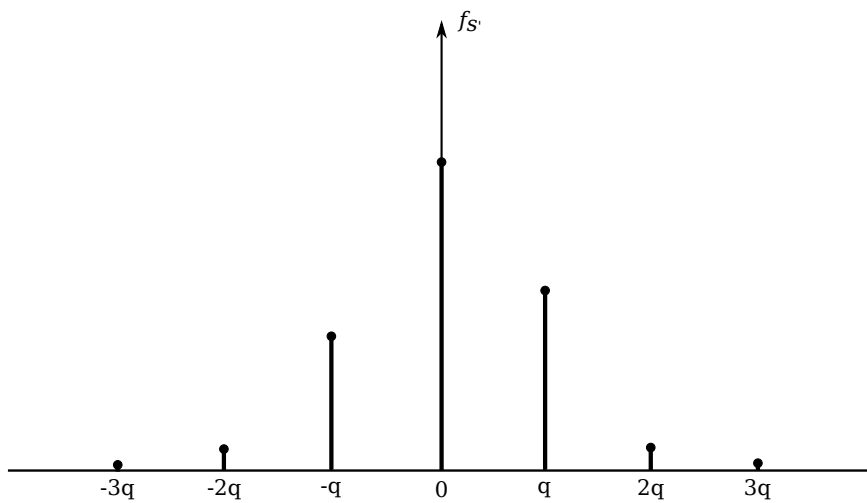
$$s'' = s + n \quad (5.6)$$

La densité de probabilité du pseudo bruit de quantification est représentée sur la figure 5.5. On peut en déduire que sa valeur moyenne est nulle :

$$\begin{aligned} \langle \nu \rangle &= \int_{-\infty}^{\infty} f_\nu(s) s ds \\ &= \int_{-q/2}^{q/2} \frac{1}{q} s ds \\ &= 0 \end{aligned} \quad (5.7)$$



(a) Densité de probabilité du signal



(b) Densité de probabilité du signal quantifié

FIGURE 5.4 – **Effet de la quantification sur la densité de probabilité du signal.** La densité de probabilité du signal quantifié prend des valeurs discrètes, d'amplitude égale à l'intégrale de la densité de probabilité du signal initial, sur les intervalles de quantification. La densité de probabilité du signal quantifiée peut donc être considérée comme une version échantillonnée de la densité de probabilité initiale du signal avant quantification. L'échantillonnage n'est cependant pas réalisé sur des valeurs discrètes mais correspond à l'intégrale sur un intervalle. On parle alors d'*échantillonnage continu*.

Sa variance est donnée par :

$$\begin{aligned}
 \langle \nu^2 \rangle &= \int_{-\infty}^{\infty} f_{\nu}(s) s^2 ds \\
 &= \int_{-q/2}^{q/2} \frac{1}{q} s^2 ds \\
 &= \frac{q^2}{12}
 \end{aligned} \tag{5.8}$$

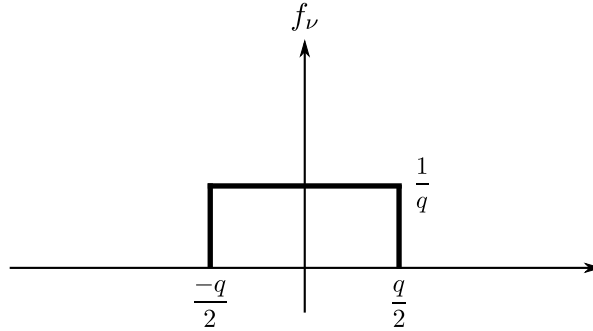


FIGURE 5.5 – **Densité de probabilité du bruit de quantification.** Elle est répartie uniformément sur l'intervalle  $[-q/2, q/2]$ .

On peut donc utiliser les résultats du théorème de la quantification, pour calculer les moments du signal quantifié en modélisant l'erreur introduite par la quantification par le bruit  $n$ . Les premiers moments du signal quantifié  $s'$  exprimés en fonction de ceux du signal non quantifié  $s$ , sont alors donnés par Widrow & Kollár (2008) :

$$\begin{aligned}
 E[s'] &= E[s] \\
 E[(s')^2] &= E[s^2] + \frac{q^2}{12} \\
 E[(s')^3] &= E[s^3] + \frac{q^2}{4} E[s] \\
 E[(s')^4] &= E[s^4] + \frac{q^2}{2} E[s^2] + \frac{q^4}{80}
 \end{aligned} \tag{5.9}$$

Les corrections par rapport aux moments non quantifiés sont connues sous le nom de corrections de Shepard. Ces résultats donnent les déviations sur les moments de la distribution, introduites par la compression.

## 5.2.4 Application à un bruit blanc gaussien

On considère un signal de densité de probabilité gaussienne de moyenne nulle et d'écart type  $\sigma$ . Sa fonction caractéristique  $\Phi_s(u) \propto \exp[-u^2\sigma^2/2]$  possède en principe des fréquences infinies. Les théorèmes de quantification ne s'appliquent donc pas. Cependant, si on choisit le pas de quantification judicieusement, on peut être arbitrairement proche des conditions du second théorème car la fonction caractéristique chute très rapidement. On peut définir le pas de quantification en unité de  $\sigma$  :

$$q = Q\sigma \tag{5.10}$$

Il est alors possible de montrer Widrow & Kollár (2008) que si le pas de quantification  $q$  est choisi de façon à avoir  $q \lesssim \sigma$  alors, l'erreur sur l'estimation des moments est négligeable.

Finalement, il a aussi été montré dans Widrow & Kollár (2008) que l'effet de la quantification sur le spectre de puissance d'un signal n'est pas nul. Pour un signal gaussien, la modélisation des effets de la quantification par un bruit de type PQN reste valide dans certaines conditions, suivant la valeur de  $Q$ . Les calculs indiquent que pour  $q \lesssim \sigma$ , les corrélations introduites sont, dans une

bonne approximation, équivalentes à un bruit blanc additionnel de variance  $q^2/12$ , non corrélé au signal Banta (1965). Ainsi, l'impact de la compression sur les propriétés statistiques d'un signal est entièrement défini par le rapport  $q/\sigma$ . Des simulations ont été réalisées par Stéphane Plaszczyński (présentations aux Core Team) et confirment ces résultats : les non-gaussianités introduites par la compression sont négligeables et le bruit de quantification est blanc.

### 5.2.5 Application aux données de l'instrument HFI

Sur de courtes périodes de temps, le bruit des bolomètres est relativement bien approximé par un bruit blanc gaussien car les dérives lentes du bruit en  $1/f$  sont négligeables. Par ailleurs, si on choisit un temps suffisamment court pour les tranches de compression, le signal varie peu et en soustrayant un offset, la distribution effective obtenue est proche d'une distribution gaussienne de moyenne nulle, élargie par l'impact des rayons cosmiques et l'émission des sources compactes. Cependant si on choisit le rapport  $q/\sigma$  en soustrayant les queues de la distribution, de façon à être dominé par l'écart type du bruit instrumental uniquement, on reste dans les conditions du théorème de quantification. En effet, l'élargissement de la fonction de distribution par les rayons cosmiques ou les sources ponctuelles, donne une fonction plus étroite en Fourier.

Les paramètres de compression, pour l'ensemble des voies de l'instrument HFI de PLANCK, ont alors été choisis pour que l'effet de la compression soit inférieur à celui des autres effets instrumentaux. Pour les thermomètres, le paramètre  $Q$  a été fixé à 0.5 tandis que pour les bolomètres il a été fixé à 0.4. L'estimation de  $\sigma$  est celle du bruit blanc en soustrayant la contribution des rayons cosmiques. Ainsi, dans ces conditions, la compression introduit donc un bruit blanc additionnel d'écart type  $\sqrt{(0.5\sigma)^2/12} = 0.144\sigma$  et  $\sqrt{(0.4\sigma)^2/12} = 0.145\sigma$  pour les thermomètres et les bolomètres respectivement. Ce qui correspond à un bruit total d'écart type  $\sqrt{\sigma^2 + (0.144\sigma)^2} = 1.01\sigma$ , soit un accroissement du bruit total de seulement 1% dans les deux cas.

## 5.3 Compression des données de l'instrument HFI

Dans cette section nous nous concentrons sur les détails de l'implémentation de la compression à bord du satellite. Dans un premier temps, nous décrivons la nature du signal brut à compresser. Puis nous donnons les détails de l'algorithme utilisé. Nous précisons ensuite les différentes étapes réalisées afin d'obtenir un signal convenable pour appliquer l'algorithme retenu. Enfin nous discutons des différents types d'erreur auxquelles peut mener la compression.

### 5.3.1 Signal brut de l'instrument HFI

L'instrument HFI possède un total de 72 voies : 54 bolomètres, 16 thermomètres ainsi qu'une résistance et une capacité de test pour suivre le vieillissement de l'électronique. Les bolomètres sont les seuls détecteurs capables d'atteindre la sensibilité nécessaire aux mesures de PLANCK dans la gamme de fréquence de l'instrument HFI (100GHz, 143GHz, 217GHz, 353GHz, 545GHz et 857GHz). Comme expliqué dans la section 4.5, le principe de fonctionnement d'un bolomètre repose sur la mesure de la variation de puissance déposée par les photons grâce à la variation de température qu'elle induit dans une résistance, qui conduit à une variation d'impédance qui se traduit finalement par une variation de tension. La mesure du signal se fait donc par la mesure de la variation de tension aux bornes des bolomètres. Afin de diminuer au maximum le bruit électronique et de pouvoir mesurer des variations de température de quelques  $\mu K$ , les bolomètres sont refroidis à 100mK. Les thermomètres de HFI ont été développés spécialement pour suivre les variations de températures des différents étages cryogéniques afin de contrôler leur régulation pour permettre d'atteindre la stabilité nécessaire aux mesures. Ils fonctionnent sur le même principe que les bolomètres. Ils utilisent la variation de la résistance provoquée par la variation de température de l'élément avec lequel ils sont en contact thermique.

Afin d'obtenir les meilleures performances, l'électronique de lecture a nécessité un long travail d'optimisation, en particulier, le montage électronique de polarisation des bolomètres qui joue un rôle

2. Historiquement, dans l'expérience PLANCK, on parle souvent du rapport  $\sigma/q$ , qui est un des paramètres de compression utilisé par le satellite. Dans la suite nous utiliserons uniquement le rapport  $q/\sigma$  qui a plus de sens, dans le cadre de l'application du second théorème de quantification à un bruit blanc gaussien.



essentiel pour la stabilité des mesures. La configuration retenue est une polarisation différentielle des bolomètres qui possède de nombreuses qualités Gaertner et al. (1997). Les bolomètres sont polarisés par un courant carré couplé à une tension carrée de compensation, en opposition de phase, obtenue par dérivation d'une tension triangle par une capacité. Afin d'obtenir un signal parfaitement stable, le montage comporte également une polarisation transitoire en courant qui compense la réponse des bolomètres aux deux autres composantes. L'allure de la contribution de ces trois composantes est représentée sur la figure 5.6 pour une période de modulation. Leur somme produit un signal relativement stable qu'on appelle la *ligne de base*. Les variations de température donnent un signal modulé autour de cette valeur. La fréquence de modulation est donnée par la fréquence du signal de polarisation  $F_{mod}$ .

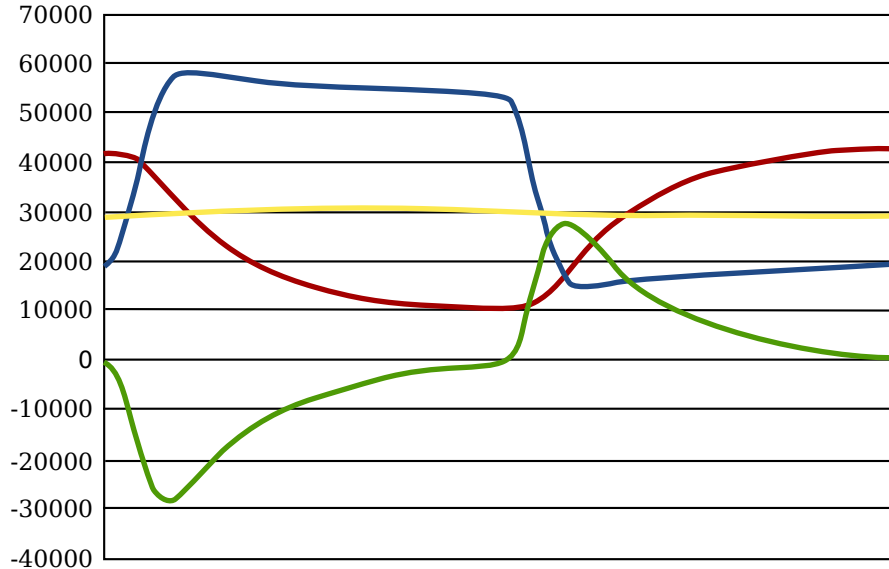


FIGURE 5.6 – **Allure des composantes du signal brut d'un bolomètre.** La réponse au courant de polarisation est visible en rouge, celle de la tension de compensation en tracée en bleu tandis que la réponse au contribution du courant transitoire est visible en vert. La réponse globale d'un bolomètre est la somme des trois et correspond à la courbe jaune. Elle est relativement stable. L'allure est tracée pour une période de modulation. (D'après Montier (2005))

La mesure du signal correspond à la mesure des variations de tension par rapport à la ligne de base. Le signal analogique est d'abord fortement amplifié, puis échantillonné à une fréquence de  $90F_{mod}$  et numérisé par un convertisseur analogique numérique qui utilise une quantification sur 16bits. On obtient alors un signal numérique, appelé signal brut sur-échantillonné, dont les valeurs varient entre 0 et  $2^{16} = 65536$ . L'unité de base de quantification est appelé ADU (pour le terme anglais de Analog to Digital Unit). Afin de répartir équitablement la gamme dynamique entre la partie modulée positivement et négativement, la conversion analogique numérique a été conçue pour que ligne de base soit proche de la valeur moyenne qui vaut 32768 ADU. Le signal brut utilisé dans l'analyse des données de l'instrument HFI correspond à la somme des échantillons du signal sur-échantillonné, sur chaque demi-période de modulation (voir figure 5.7). Cette opération est effectuée par le REU (Readout Electronic Unit voir section 4.5). Le signal brut correspond donc à un signal modulé de fréquence  $2F_{mod}$ . On définit alors la parité du signal : les échantillons bruts pairs correspondent à la somme sur les données sur-échantillonnées dans la phase de modulation supérieure à la ligne de base et les échantillons impairs, ceux de la phase de modulation inférieure.

Dans la pratique la réponse du système électronique à l'inversion de polarité n'est pas instantanée. Deux paramètres ont alors été introduits pour le calcul de la somme :

- $N_{blank}$  qui correspond au nombre de points au début de chaque demi-période qui ne sont pas pris en compte dans la somme. Il permet d'éviter la zone transitoire où le signal n'est pas nécessairement stabilisé suite au changement de polarisation.
- $S_{phase}$  qui introduit un déphasage du début de la somme par rapport au début de chaque demi-période.

Le couple  $(S_{phase}, N_{blank})$  a été optimisé afin de maximiser le rapport signal sur bruit des données

brutes.

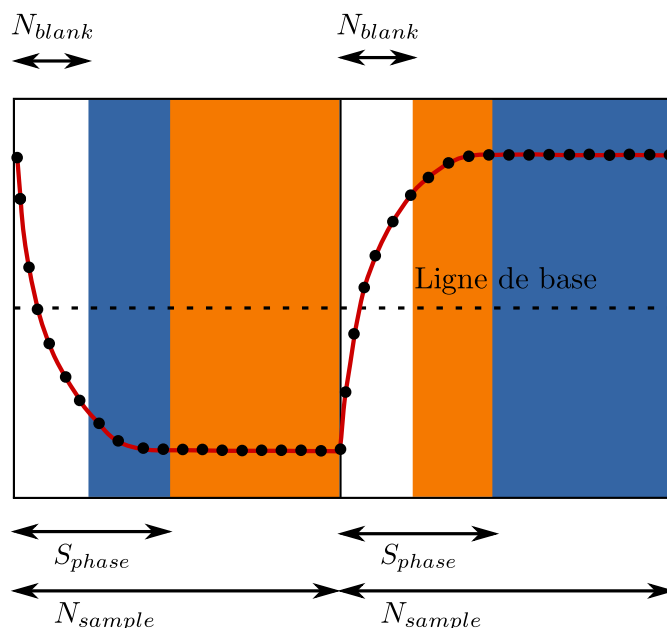


FIGURE 5.7 – **Période du signal brut sur-échantillonné.** Le signal mesuré correspond à la variation de tension par rapport à la ligne de base. Il est échantillonné avec une fréquence de  $90F_{mod}$ , où  $F_{mod}$  représente la fréquence de modulation. Les échantillons bruts du signal scientifique sont obtenus en sommant les données sur-échantillonnées sur une demi-période de modulation. Les  $N_{blank}$  premiers échantillons ne sont pas pris en compte. Par ailleurs un déphasage de  $S_{phase}$  est ajouté. Ces paramètres ont été optimisés afin de maximiser le rapport signal sur bruit. Chaque période de modulation donne finalement un échantillon scientifique brut pair (somme des échantillons dans la zone bleu) et un échantillon impair (somme des échantillons dans la zone orange).

On a  $N_{sample} = 45$  échantillons sur chaque demi-période de modulation, le signal brut correspond donc au maximum, dans le cas où  $N_{blank} = 0$ , à la somme de 45 échantillons codés sur 16bits, ce qui nécessite au minimum 22bits afin de pouvoir coder l'ensemble des valeurs possibles ( $2^{21} < 45 \times 2^{16} < 2^{22}$ ). Afin de minimiser la taille mémoire nécessaire, les échantillons bruts issus de la somme du signal sur-échantillonné, sont donc codés sur 22bits. Pour les observations, la valeur de  $N_{blank}$  retenue vaut 5. À la sortie du REU, les échantillons bruts de HFI sont donc des nombres codés sur 22bits, variant entre 0 ADU et  $40 \times 2^{16} = 2621440$  ADU. La ligne de base est centrée autour d'une valeur proche de  $(40 \times 2^{16})/2 = 1310720$  ADU qui dépend de la position exacte de la tension servant d'origine à la conversion analogique numérique comparée à la tension exacte de la ligne de base.

### 5.3.2 Application aux données de Planck

Dans cette section nous décrivons les détails de l'implémentation de l'algorithme de compression à bord du satellite. Nous donnons dans un premier temps les contraintes matérielles qui ont limité la complexité des calculs réalisables. Nous décrivons ensuite les principales étapes réalisées par le DPU afin de compresser les données : l'estimation de la ligne de base et la démodulation des données, l'estimation et la soustraction de l'offset puis le codage final des données sur le dictionnaire de compression et le regroupement des données dans les paquets téléométriques pour l'envoi sur Terre.

#### Contraintes matérielles

Les détails de l'implémentation de l'algorithme de compression doivent tenir compte des contraintes matérielles. Le DPU ne possède que de faibles ressources, notamment en mémoire et en puissance de calcul. Elles doivent être partagées entre les différentes tâches de façon équitable, proportionnelle à l'importance des calculs et dans la limite disponible. Le DPU ne dispose que de 128koctet de mémoire pour charger l'ensemble du système et les données en cours de traitement. Le

programme effectuant la compression des données doit ainsi avoir une empreinte mémoire la plus faible possible. Par ailleurs, l'opération de compression doit être réalisée en temps réel. C'est-à-dire qu'elle doit pouvoir s'exécuter dans un temps borné en tenant compte des autres opérations concurrentes. L'unité de calcul du DPU est cadencée à 20MHz ce qui limite le nombre d'opérations réalisables pour chaque tranche de compression.

### Principales étapes de la compression

À bord du satellite, le signal brut, issu des 72 voies de l'instrument HFI, est compressé puis empaqueté et envoyé à l'ordinateur de bord par le DPU. Ce dernier réalise plusieurs étapes importantes, successives, afin d'obtenir le signal compressé, avant de produire les paquets télémetriques qui seront envoyés sur Terre. Les données de chaque voie sont traitées par tranches de 254 échantillons, appelées tranches de compression. Pour chaque tranche de compression, la ligne de base de modulation de l'électronique est estimée, les données brutes sont démodulées puis leur offset est estimé et soustrait. L'amplitude du signal ainsi obtenu est alors quantifiée et chaque intervalle est codé sur un mot d'un dictionnaire non ambigu nécessitant au total moins de bits pour être codé. Finalement ce signal quantifié est envoyé à l'ordinateur de bord sous forme de paquets standardisés contenant toute l'information afin de pouvoir reconstruire le signal brut. On peut résumer ce traitement de la façon suivante :

- découpe du signal en tranche de compression
- estimation de la ligne de base et démodulation des données brutes
- estimation et soustraction de l'offset
- quantification et codage sur un dictionnaire
- regroupement des données dans des paquets
- envoi sur Terre
- décompression

Les principales étapes sont résumées sur la figure 5.8

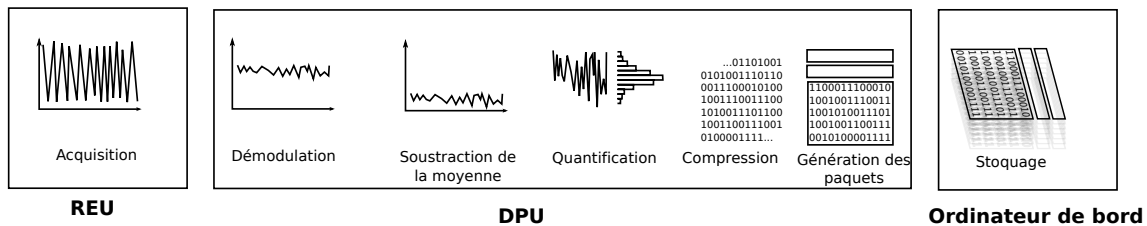


FIGURE 5.8 – **Schéma des étapes de la compression des données.** La compression des données est effectuée par le DPU. À la sortie du REU les données sont traitées par tranches de compression de 254 échantillons. Les données brutes sont démodulées, leur moyenne est soustraite puis l'algorithme de compression leur est appliquée. Pour cela une quantification en amplitude est effectuée dans un premier temps. À chaque intervalle de quantification est associé un mot sur un dictionnaire utilisant un nombre de bits variable, fonction de la probabilité que les échantillons tombent dans l'intervalle. Chaque échantillon quantifié est alors écrit avec ce dictionnaire. Les données ainsi compressées sont empaquetées puis transmises à l'ordinateur de bord où elles sont stockées avant d'être envoyées sur Terre.

### Estimation de la ligne de base de modulation

Comme on vient de le voir, la première étape réalisée par le DPU consiste à démoduler les données brutes de la tranche de compression considérée. Pour cela, il est nécessaire d'estimer la ligne de base de modulation de la chaîne électronique. De façon générale, on observe que cette ligne de base est relativement stable. Cependant, dans certaines conditions, des perturbations peuvent l'affecter temporairement. En particulier les impacts de rayons cosmiques les plus puissants sont capables de modifier la réponse de l'électronique, ce qui peut provoquer une variation de la valeur de la ligne de base de quelques ADU pendant un temps qui peut être supérieur à la durée d'une tranche de compression. Ce phénomène est bien compris et relativement bien reproduit par les modèles. Pour l'exploitation scientifique des données, une démodulation tenant compte de la réponse

de l'électronique au signal est utilisée. Elle est basée sur un filtre numérique fini glissant. Ce filtre a été obtenu grâce à une modélisation précise de la réponse de l'électronique et aux mesures des paramètres des bolomètres. C'est une opération trop complexe en terme de nombre de calcul, pour être effectuée par le DPU, étant données les faibles ressources disponibles dans le satellite et les contraintes de temps réel.

Une mauvaise démodulation ajoute au signal une oscillation dont l'amplitude est proportionnelle à l'erreur, ce qui modifie la répartition en amplitude des données de la tranche de compression et les échantillons tombent en général dans un intervalle différent. Ils sont ensuite codés sur un nombre de bits différent, ce qui dans la plupart des cas va augmenter le débit nécessaire. L'unique conséquence d'une mauvaise démodulation, est donc de diminuer l'efficacité de la compression en terme de débit mais elle n'affecte pas la qualité des données.

Ainsi, pour faciliter les calculs, à bord du satellite, pour démoduler les données avant la compression, la ligne de base de modulation de l'électronique est supposée constante sur chaque tranche de compression, même si cette hypothèse n'est pas toujours vérifiée. La ligne de base est alors estimée en calculant la moyenne des échantillons de la tranche de compression. Cette procédure conduit à une erreur sur l'estimation qui peut être approximée dans le cas d'un signal purement gaussien d'écart type  $\sigma$ , comme la variation de la moyenne estimée à partir de 254 échantillons. Elle vaut  $\sigma/\sqrt{254} \approx \sigma/16$ . Pour un rapport  $q/\sigma$  de 0.5, étant données les valeurs de  $q$  (voir tableau C.1), l'erreur sur la moyenne est d'une dizaine d'ADU dans le pire des cas. Dans la pratique, de fortes variations du signal produites par exemple par l'impact des rayons cosmiques, peuvent modifier sensiblement la distribution des points et conduire à une estimation encore plus biaisée de la ligne de base de l'électronique.

Lorsque l'énergie déposée par les rayons cosmiques n'est pas trop importante, ils ne modifient pas la ligne de base. Cependant, dans tout les cas on observe une brusque augmentation de la température suivie d'une décroissance exponentielle caractéristique du bolomètre qui correspond au refroidissement (voir par exemple la figure 5.16). Ce pic de température biaise l'estimation de la ligne de base à partir de la moyenne des échantillons. Pour limiter ce biais, les éléments de plus forte et de plus faible amplitude ne sont pas pris en compte dans le calcul de la moyenne. Lors du développement de l'algorithme de compression, des études ont montré que cela constituait un bon compromis entre la complexité des calculs et l'augmentation du débit induite par le biais résiduel.

Finalement à partir de la ligne de base estimée, les données sont démodulées. Les échantillons démodulés  $\hat{x}$ , sont obtenus à partir des échantillons modulés  $x$ , de la façon suivante :

$$\hat{x}_{2p} = 2b - x_{2p} \quad (5.11)$$

$$\hat{x}_{2p+1} = x_{2p+1} \quad (5.12)$$

où  $p$  est un entier positif et  $b$  est la valeur de la ligne de base estimée. Ainsi, les échantillons impairs (dans la partie positive de modulation) restent inchangés tandis que les échantillons pairs sont obtenus à partir de l'opposé de l'écart d'amplitude par rapport à la ligne de base des échantillons modulés.

### Estimation de l'offset

Le dictionnaire permettant de coder la valeur des différents intervalles de compression a été implémenté afin de donner la meilleure efficacité de compression, en terme de réduction du débit pour une signal proche d'un signal gaussien de moyenne nulle. Il est donc nécessaire de soustraire l'offset des données démodulées avant de compresser le signal afin d'obtenir une distribution centrée autour de 0. L'offset est obtenu à partir du signal démodulé en calculant la moyenne des échantillons. L'impact des rayons cosmiques peut biaiser cette estimation. Afin de limiter ce problème, le calcul de la moyenne pour l'estimation de l'offset, ne tient pas compte des échantillons de la tranche de compression dont l'amplitude est supérieure à  $10\sigma$ . Cette méthode permet de ne pas tenir compte des échantillons trop éloignés de la moyenne et augmente la précision sur l'estimation de l'offset.

### Codage des données pour la compression

L'algorithme de compression prend en entrée un signal qui est supposé être un bruit blanc gaussien. Cette hypothèse est bien vérifiée sur des temps courts. En effet, le signal est en réalité

Intervalle	Mot binaire associé	Bits	Probabilité (*)
-1	00	2bits	0.1465
0	01	2bits	0.1585
+1	10	2bits	0.1465
-2	111111	6bits	0.1156
+2	110111	6bits	0.1156
-3	111011	6bits	0.0779
+3	110011	6bits	0.0779
-4	111100	6bits	0.0448
+4	110100	6bits	0.0448
-5	111101	6bits	0.0220
+5	110101	6bits	0.0220
$[-6, \dots, -22]$	110000XXXX	10bits	$< 0.0092$
$[+6, \dots, +22]$	111000XXXX	10bits	$< 0.0092$
$[-23, \dots, -1047]$	110001XXXXXXXXXXXX	16bits	$< 10^{-17}$
$[+23, \dots, +1047]$	111001XXXXXXXXXXXX	16bits	$< 10^{-17}$
$[-1048, \dots, -66581]$	110010XXXXXXXXXXXXXXXXXXXX	22bits	indéfectable
$[+1048, \dots, +66581]$	111010XXXXXXXXXXXXXXXXXXXX	22bits	indéfectable
$[-66584, \dots]$ (Underflow)	111110	6bits	} Erreurs de compression
$[+66584, \dots]$ (Overflow)	110110	6bits	
Plus de 64 éch. sur 22bits	110110	6bits	

(\*) en supposant une distribution gaussienne et un rapport  $q/\sigma = 0.4$

TABLE 5.1 – **Tableau récapitulatif du dictionnaire pour la compression.** Afin de compresser le signal, l'amplitude de chaque échantillon est quantifiée. À chaque intervalle de quantification est associé un mot binaire. La taille du mot varie de 2 bits à 22bits. Avant la compression, chaque échantillon est codé sur 22bits. Le tableau ci-dessus donne la correspondance entre l'intervalle de quantification et le mot binaire. Il contient également la probabilité qu'un échantillon tombe dans l'intervalle considéré en supposant que le signal est gaussien et que le pas de quantification  $q$ , a été choisit de façon à avoir un rapport  $q/\sigma = 0.4$ . On voit que la plupart des échantillons sont codés sur des mots de moins de 10 bits. Cela permet alors d'utiliser moins d'espace et d'atteindre le taux de compression nécessaire.

composé de l'émission du CMB à laquelle s'ajoutent toutes les composantes qui émettent aux longueurs d'onde d'observation plus le bruit de la mesure. Si on considère un temps suffisamment court, les composantes autres que le CMB ne varient pas beaucoup et les dérives lentes du bruit sont également très faibles. On observe cependant une modification significative de la distribution des échantillons, produite par l'impact des rayons cosmiques et l'observation des sources ponctuelles.

Après cette étape de pré-traitement, décrite ci-dessus, l'algorithme de compression décrit dans la section 5.2 est utilisé pour compresser les valeurs. L'amplitude de chaque échantillon démodulé est quantifiée, avec un pas de quantification  $q$  fixe. Le pas est déterminé de façon indépendante sur chaque voie afin d'obtenir un rapport  $q/\sigma$  correct. À chaque intervalle de quantification est associé un mot binaire, appartenant à un dictionnaire pré-établi, détaillé dans le tableau 5.1. Le codage sur le dictionnaire utilise moins de bits que celui des échantillons bruts ce qui permet d'obtenir un signal compressé.

La probabilité qu'un échantillon tombe dans un intervalle donné dépend de la nature statistique du signal. Étant donné le taux d'échantillonnage de HFI de 180.38Hz, le nombre d'échantillons par tranche de compression et la vitesse de rotation du satellite sur lui-même d'un tour par minute, chaque tranche de compression dure 1.4 secondes et couvre une distance angulaire de  $8.4^\circ$  sur le ciel. À l'échelle d'une tranche de compression, le signal varie peu, notamment les dérives lentes du bruit en  $1/f$ , la composante dipolaire du CMB, ou même le signal galactique (excepté dans les régions du plan galactique pour ce dernier). Le signal est donc proche d'un bruit blanc gaussien. La distribution réelle est tout de même élargie par l'impact des rayons cosmiques et le signal des sources ponctuelles. La taille de chaque mot du dictionnaire a été optimisée en tenant compte de la répartition statistique d'un signal gaussien.

L'étape de pré-traitement permet d'obtenir un signal proche d'un bruit blanc gaussien de moyenne

nulle ce qui permet à l'algorithme de compression d'atteindre ses meilleures performances aussi bien en terme d'efficacité de compression que d'impact sur le signal. Les intervalles pour lesquels la probabilité est la plus élevée (les intervalles proches de zéro), sont codés sur des mots de 2, 6 puis 10 bits, tandis que ceux dont la probabilité est la plus faible, utilisent 16 bits puis 22 bits.

D'après le tableau 5.1, on voit que l'algorithme permet de coder  $2 \times (1 + 4 + 16 + 1024 + 65536) + 1 = 133163$  intervalles différents. L'amplitude maximale est donc  $\pm 65581.5q$

### Regroupement des données dans les paquets téléométriques

Finale­ment, pour chaque tranche de compression, les échantillons compressés de toutes les voies, sont regroupés dans les « paquets science » puis envoyés à l'ordinateur de bord qui va les stocker dans sa mémoire, avant de les transmettre à la Terre. Les informations nécessaires à la reconstruction du signal accompagnent les données compressées dans les paquets formés par le DPU afin de pouvoir reconstruire le signal brut sur Terre. Chaque paquet contient un en-tête principal qui l'identifie et qui permet de s'assurer entre autre de la validité de son contenu. Les données de chaque voie, pour une tranche de compression donnée, sont regroupées de façon séquentielle à l'intérieur des paquets. Elles possèdent un en-tête spécifique qui permet notamment d'identifier la voie de laquelle proviennent les mesures et qui contient également la valeur de l'offset et la valeur de la ligne de base, estimées sur la tranche de compression considérée. Cela permet par la suite, dans le centre de traitement, sur Terre, de reconstruire le signal compressé, en appliquant la procédure inverse de celle qu'on vient de décrire.

Les valeurs de la ligne de base et de l'offset sont codées sur un nombre à virgule flottante de 32 bits. Elles sont transmises dans l'en-tête associé à la tranche de compression de chaque voie dans les paquets téléométriques. Sur Terre cette valeur est nécessaire pour reconstruire les TOI du signal brut à partir des données compressées.

### 5.3.3 Erreurs de compression

Dans certaines conditions particulières, la compression peut conduire à des erreurs, entraînant une perte de données irrécupérable. Cependant, les conditions extrêmes dans lesquelles on peut rencontrer ce type d'erreurs, sont très peu probables. Par ailleurs, elles conduisent dans la plupart des cas à la perte de quelques échantillons par tranche de compression uniquement ce qui est totalement négligeable dans ce cas.

Les erreurs de compression attendues sont de trois types. En fonction de la valeur des paramètres de compression, la gamme dynamique des échantillons quantifiés peut être inférieure à la gamme dynamique du signal ce qui peut provoquer une saturation de la compression. Par ailleurs, deux contraintes ont été ajoutées pour répondre à des situations extrêmes produisant un débit trop important : une limite du nombre d'échantillons à fort poids et une limite de débit moyen intégré. Le principal avantage de ces mesures de protection est qu'elles permettent de contrôler en avance dans quelles conditions les pertes de données ont lieu afin d'en limiter les conséquences.

### Erreurs par dépassement de la limite en amplitude

Le signal observé par PLANCK varie fortement avec la fréquence et la direction d'observation. Aux plus basses fréquences d'observation de l'instrument HFI, l'émission est dominée par celle du CMB. Ses variations d'intensité sont faibles comparées aux autres sources. Elles sont dominées par les variations du dipôle. Lorsque la fréquence augmente, l'émission de la poussière va progressivement dominer. Cette émission varie fortement avec la direction d'observation. Son intensité est très élevée dans les régions proches du plan galactique et beaucoup plus faible aux hautes latitudes galactiques. Ces variations sont les plus marquées pour les observations des voies à 545GHz et 857GHz de HFI. L'amplitude du signal couvre donc une gamme dynamique différente suivant les fréquences d'observation. La gamme dynamique de l'électronique de lecture étant limitée par celle de la conversion analogique numérique et identique pour l'ensemble des voies (voir section 5.3.1), afin de ne pas obtenir de saturation, le gain appliqué aux voies à haute fréquence est moins important. Cela produit alors un niveau de bruit également plus faible. Par ailleurs, chaque bolomètre et son électronique de lecture, possède un niveau de bruit qui lui est propre même avec un gain identique. Le pas de quantification en unité ADU est donc différent d'une voie à l'autre et doit être déterminé

indépendamment sur chacune d'elle. On peut voir dans le tableau C.1 que la variation est faible pour un ensemble de détecteurs à la même fréquence. Les voies à 545GHz et 857GHz, qui ont le gain le plus faible, ont également un pas de quantification plus faible que les autres voies.

Le signal brut en sortie du REU est la somme de 40 échantillons codés sur 16bits (voir section 5.3.1). L'amplitude maximale d'un échantillon brut est donc  $m = 40 \times 2^{16} = 2621440$  ADU. Le signal brut possède ainsi des valeurs comprises dans l'intervalle  $[0, \dots, 2621440]$  ADU. L'algorithme de compression est capable de coder des valeurs entre  $\pm 66584.5q$  (voir section 5.3.2). Si l'amplitude d'un échantillon du signal obtenu après démodulation et soustraction de l'offset sur la tranche de compression considérée, dépasse la limite précédente, l'algorithme de compression générera une erreur de compression.

La limite en amplitude de la quantification dépend de la valeur du pas de quantification  $q$ . Les variations d'amplitude du signal à quantifier dépendent elles de la ligne de base utilisée pour démoduler les données. Si on note  $b$  la valeur de la ligne de base estimée pour la tranche de compression considérée, alors, d'après la formule 5.11 qui donne l'expression des échantillons modulés en fonction des échantillons démodulés, l'amplitude  $A_d$ , du signal démodulé variera donc :

$$\begin{aligned} 2b - m < A_p < m & \quad \text{si } b < m/2 \\ 0 < A_p < 2b & \quad \text{sinon} \end{aligned}$$

Dans la pratique la valeur de  $b$  est très proche de  $m/2$ . Cependant, l'estimation de cette valeur sur une tranche de compression peut parfois être biaisée en présence d'un signal dont l'amplitude varie fortement à l'échelle d'une tranche de compression. Grâce à la technique employée pour estimer cette valeur (voir section 5.3.2), le biais reste limité. Pour le calcul d'ordre de grandeur qui suit, on peut considéré que  $b$  vaut donc  $m/2$ .

La valeur de l'offset ne change pas la dynamique du signal. Cependant si sa valeur est fautive l'hypothèse d'un signal à moyenne nulle n'est plus vérifiée. On peut considéré dans un premier temps le pire des cas où l'ensemble des échantillons sont soit positifs, soit négatifs. Dans ces conditions, pour  $q > 40 \times 2^{16}/66584.5$  soit  $q > 40$ , l'ensemble des valeurs possibles peut être codé par l'algorithme de compression. La gamme dynamique de la compression est plus large que celle du signal. On ne peut donc pas avoir d'erreur de compression par dépassement de la limite en amplitude. Dans le tableau C.1 on compte 42 voies satisfaisant ce critère. Pour les autres, l'algorithme de compression ne sera pas capable de coder l'ensemble de la gamme dynamique disponible et on pourra rencontrer des erreurs de compression, si l'amplitude est trop importante (*overflow*) ou trop faible (*underflow*). Dans le cas où l'offset est bien estimé conduisant à un signal à compresser bien centré autour de 0, ce qui est généralement le cas dans une bonne approximation, on peut coder l'ensemble des variations d'amplitude tant que  $q < 20$ . Étant données les valeurs de  $q$  du tableau C.1 et la très faible probabilité d'obtenir une amplitude si importante, on ne s'attend pas à rencontrer un grand nombre de ces erreurs. Dans tous les cas, ce type d'erreur conduit uniquement à la perte des échantillons dont la valeur absolue de l'amplitude est trop élevée. Il est donc peu probable de perdre l'ensemble des échantillons d'une tranche de compression. On verra dans la section 5.6.3 que ce type d'erreur n'a été observé qu'une fois.

### Erreurs par dépassement de la limite des échantillons à fort poids

L'algorithme de compression est doté d'un dispositif permettant de limiter le nombre d'échantillons par tranche de compression utilisant un trop grand nombre de bits. Si pour une voie donnée, plus de 64 échantillons de la tranche de compression en cours de traitement, ont été codés sur 22bits, tous les échantillons suivants qui auraient dû être codés sur 22bits, généreront une erreur de compression (voir tableau 5.1). Cela assure que même dans le pire des cas, où tous les éléments devraient être codés sur 22bits, le débit total pour la tranche de compression soit limité à  $64 \times 22 + 190 \times 6 = 2548$  bits : soit une moyenne de 10bits par échantillon. Ce système permet de s'assurer de ne pas saturer la bande passante en cas de dysfonctionnement d'une voie dont le niveau de bruit augmenterait fortement ou en cas de problèmes causés par les rayons cosmiques par exemple. Il conduit au maximum à la perte de 190 échantillons par tranche de compression.

### Erreurs par dépassement du débit moyen (EOS)

Un dispositif permet d'arrêter totalement l'acquisition des données dans certaines conditions particulières. Il mesure la quantité de données au cours de cycles de temps donné. Si l'espace en mémoire nécessaire pour stocker l'ensemble des paquets produits depuis le début du cycle dépasse l'espace maximal autorisé pour un cycle, alors, l'envoi des « paquets science » à l'ordinateur central, est interrompu par le DPU jusqu'à la fin du cycle. Ce mécanisme est appelé le *End of Slew* (EOS). Dans le cas où le débit nécessaire est trop important ce mécanisme permet de répartir les pertes de données sur l'ensemble des *rings*. Grâce à la redondance des observations dans chaque *ring*, le EOS limite l'impact de la perte de données.

### Identification des erreurs dans la base de données

Sur Terre les données sont décompressées au niveau L1. Les échantillons issus d'une erreur de compression sont remplacés dans la base de données contenant le signal par des 0. Par ailleurs, ces échantillons sont également marqués comme non valides dans la base de données ( NOT WRITTEN). Il n'existe aucun moyen, à partir de la base de données, de savoir quel type d'erreur de compression est à l'origine de la perte d'un échantillon particulier. Les erreurs issues du déclenchement d'un EOS sont en principe plus facilement identifiables car ce mécanisme conduit à l'arrêt total de l'envoi des « paquets science ». Ainsi, on observe simultanément une perte de données sur toutes les voies. Le nombre d'échantillons perdus doit, par ailleurs, être multiple de la taille d'une tranche de compression. L'objet ONE\_OR\_MORE\_BOLO, de la base de données peut être utilisé pour mettre en évidence ce type d'erreur. Sa valeur vaut 1 lorsque les échantillons d'au moins une voie sont disponibles. Malheureusement, certains problèmes de transmission ou dans le traitement des données au niveau L1, peuvent également conduire à la perte de certains paquets, ce qui se traduit par un effet similaire au déclenchement d'un EOS. Il n'est pas possible de distinguer ces cas sans contacter directement les responsables de L1.

## 5.4 Réglages des paramètres de compression

Le satellite PLANCK a été lancé le 14 mai 2009 mais la période d'observation régulière n'a débuté que le 13 août 2009. Elle a été précédée par une longue phase de tests des instruments, de calibration en vol et de réglages des paramètres. Après le lancement, les instruments ont progressivement été refroidis au cours du parcours du satellite vers son orbite finale qui a duré près d'un mois. Cette phase qui précède le début des observations est indispensable afin d'obtenir la meilleure qualité des mesures et la plus longue durée de vie de la mission. La température nominale des différents étages a été atteinte le 3 juillet 2009. Dans un premier temps, il a fallu s'assurer du bon fonctionnement de l'ensemble des éléments constituant le satellite. Les tests de calibration ont par la suite permis de connaître la réponse des détecteurs dans leurs conditions d'utilisation finale, en vol. Finalement l'ensemble des réglages nécessaires, afin de placer le satellite et ses instruments dans un état de fonctionnement optimal ont été effectués, notamment les réglages relatifs aux différents étages cryogéniques (le débit de la dilution d'He par exemple), l'intensité de polarisation des bolomètres et également les réglages des paramètres de compression... Beaucoup de ces réglages influencent directement les niveaux de bruit des instruments. Les réglages de la compression qui dépendent fortement du niveau de bruit ont donc été effectués dans les derniers jours.

### 5.4.1 Procédure de réglage des paramètres

Les seuls paramètres dont dépend l'algorithme de compression sont, l'écart type du signal  $\sigma$  et le pas de quantification  $q$  (voir section 5.2). C'est le rapport des deux qui va fixer l'efficacité et l'impact sur le signal de la compression. Dans la section 5.2, on a montré que si le rapport  $Q = q/\sigma$  était inférieur ou égal à 1, alors, l'effet de la compression sur le spectre de puissance était équivalent à celui d'un bruit blanc et que la compression ne produisait pas de non gaussianité. Il a donc été décidé de fixer le rapport  $q/\sigma$  à 0.4 pour l'ensemble des bolomètres et à 0.5 pour les thermomètres. Des simulations ont montrées que cela constituait le meilleurs compromis, permettant d'obtenir un débit inférieur à la limite de 80kbits/s tout en réduisant au maximum le bruit de quantification.



Effectuer le réglage des paramètres de compression consiste à déterminer la valeur de deux des trois paramètres précédents, ou toutes combinaisons de ces paramètres. Un fois ces deux paramètres transmis au DPU, celui-ci en déduit la valeur de  $q$  (en unité ADU) et l'utilise pour quantifier le signal à compresser. Il a été choisi de transmettre les valeurs de  $\sigma$  et  $\sigma/q$  au satellite. Le changement de ces valeurs est effectué par l'envoi de télécommandes particulières au satellite. Une procédure normalisée a été établie par l'ESA pour l'envoi de l'ensemble des réglages des instruments. Une fois les paramètres déterminés, il est alors nécessaire d'effectuer une requête aux opérateurs du MOC qui sont les seuls autorisés à interagir avec le satellite. Ils génèrent finalement les télécommandes appropriées qui sont envoyées au cours de la communication quotidienne avec le satellite.

#### 5.4.2 Réglages pendant la phase pré-observation

Sur les figures 5.9(a) à 5.9(h) on peut observer de fortes variations du rapport  $q/\sigma$ , avant le début du moins d'août 2009. Dans cette phase de contrôle, de tests et de calibration des instruments, qui a précédé la phase d'observation régulière, les réglages de la compression ont été ajustés fréquemment. Au cours de cette période, de nombreux paramètres des instruments ont été modifiés, afin de placer le satellite et ses instruments dans un état de fonctionnement optimal, ce qui a conduit à des variations des niveaux de bruit des différentes voies et qui a nécessité d'adapter en permanence les paramètres de compression. Pour suivre l'évolution globale du signal les réglages de la compression n'ont pas besoin d'être optimaux. La création de non gaussianités par exemple, qui résulte d'un mauvais réglage de la compression ne pose pas de problèmes particuliers comme c'est le cas pour l'étude précise du CMB. Cependant, comme des tests et des réglages critiques ont été effectués, il est indispensable d'éviter toute perte de données. Au cours de la phase qui a précédé les observations, les paramètres de compression ont donc été ajustés principalement dans le but de ne pas produire d'erreurs de compression et de ne pas déclencher de EOS.

#### 5.4.3 Outils de surveillance quotidienne

Avant le lancement, un ensemble d'outils de contrôle a été développé, pour suivre le fonctionnement du satellite et de ses instruments. Ils permettent d'abord de s'assurer de la bonne réception des données mais également de vérifier le bon fonctionnement des différents éléments du satellite. Une surveillance systématique est effectuée chaque jour pendant la période de réception des données. Cela permet de détecter les éventuelles anomalies apparues depuis la dernière période de communication avec le satellite. Une courte période permet en plus de suivre l'évolution des mesures, en temps réel. En cas de problème, des télécommandes peuvent être envoyées au satellite afin de modifier son état et d'éviter les dysfonctionnements. Un des outils de surveillance, développé par Stéphane Plaszczyński avec l'aide des ingénieurs de l'IAS, est chargé de surveiller le fonctionnement de la compression des données. Pour cela, il fournit des statistiques sur les neuf intervalles de quantification centraux, puis sur le nombre d'échantillons codés sur 10bits, 16bits et 22bits. Cela permet de suivre l'évolution globale, de la répartition statistique de l'amplitude du signal au cours du temps. Connaissant le rapport  $q/\sigma$  on peut en déduire directement la probabilité qu'un échantillon tombe dans un intervalle de quantification. Cet outils permet donc de vérifier de façon simple l'état des réglages de la compression.

#### 5.4.4 Réglages initiaux

Pour effectuer les réglages initiaux, il est nécessaire dans un premier temps d'estimer  $\sigma$  en unité ADU puis de déterminer la valeur du pas de quantification (toujours en unité ADU) à partir du rapport précédemment fixé. Cette estimation doit se faire en vol lorsque l'ensemble des réglages qui peut influencer les niveaux de bruit a été effectué. En effet, de nombreux paramètres peuvent modifier l'écart type du signal notamment la température, le gain de l'électronique de lecture... Dans un premier temps les réglages sont effectués à partir de la valeur de l'écart type calculée par le DPU que l'on trouve dans les paquets de maintenance. Cette estimation est approximative car elle ne tient pas compte de la déviation introduite par les rayons cosmiques ou les sources ponctuelles par exemple. Dans un second temps on utilise l'outil de surveillance pour rajuster la valeur des paramètres. On utilise pour cela la valeur de l'intervalle le plus peuplé, qui est celui centré en zéro. Si on suppose que le signal a une densité de probabilité gaussienne d'écart type  $\sigma$ , connaissant le

rapport  $q/\sigma$ , on peut calculer la probabilité qu'un échantillon tombe dans un intervalle donné. Pour l'intervalle centré en zéro, la probabilité  $p_0$  est donnée par :

$$\begin{aligned}
p_0 &= \int_{-q/2}^{q/2} \frac{1}{\sigma\sqrt{2}} e^{-\frac{x^2}{2\sigma^2}} dx \\
&= \int_0^{q/2} \frac{1}{\sigma\sqrt{2}} e^{-\frac{x^2}{2\sigma^2}} dx + \int_{-q/2}^0 \frac{1}{\sigma\sqrt{2}} e^{-\frac{x^2}{2\sigma^2}} dx \\
&= \int_0^{q/2} \frac{1}{\sigma\sqrt{2}} e^{-\frac{x^2}{2\sigma^2}} dx - \int_0^{-q/2} \frac{1}{\sigma\sqrt{2}} e^{-\frac{x^2}{2\sigma^2}} dx \\
&= \frac{1}{2} \operatorname{erf}\left(\frac{q/2}{\sigma\sqrt{2}}\right) - \frac{1}{2} \operatorname{erf}\left(\frac{-q/2}{\sigma\sqrt{2}}\right) \\
&= \frac{1}{2} \operatorname{erf}\left(\frac{q}{\sigma 2\sqrt{2}}\right) + \frac{1}{2} \operatorname{erf}\left(\frac{q}{\sigma 2\sqrt{2}}\right) \\
&= \operatorname{erf}\left(\frac{q}{\sigma 2\sqrt{2}}\right)
\end{aligned} \tag{5.13}$$

où la fonction erf est la fonction d'erreur définie par :

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz \tag{5.14}$$

En utilisant alors le développement limité de l'inverse de la fonction erf en 0, donné par :

$$\operatorname{erf}^{-1}(x)|_0 \approx \frac{\sqrt{\pi}}{2} x + o(x^3) \tag{5.15}$$

on trouve :

$$\frac{q}{\sigma} \approx \sqrt{2\pi} p_0 \approx 2.5 p_0 \tag{5.16}$$

Ainsi connaissant la probabilité de l'intervalle central on peut en déduire très simplement le rapport  $q/\sigma$ .

Cependant, le signal réel ne possède pas une distribution purement gaussienne. De nombreux échantillons peuplent les queues de la distribution. Pour limiter leur influence, le rapport  $q/\sigma$  est estimé de la façon suivante :

$$\frac{q}{\sigma} = \frac{2.5 * p_0}{1 - p_{out}} \tag{5.17}$$

où  $p_{out}$  est la probabilité des éléments peuplant les intervalles de quantification au-delà de  $[-5q - q/2; 5q + q/2]$ .

Finalement, pour effectuer le réglage de la compression sans avoir à estimer  $\sigma$  il suffit alors de comparer la valeur du paramètre  $Q$ , estimée par l'outil de surveillance et d'en déduire le facteur multiplicatif à apporter afin d'obtenir le bon rapport.

### 5.4.5 Réglages fins

Les outils de surveillance permettent de suivre l'évolution de la compression, cependant la méthode d'estimation du rapport  $q/\sigma$  est basée sur le peuplement de l'intervalle centré en zéro uniquement. Cependant, le signal de l'instrument HFI de PLANCK est affecté par les rayons cosmiques, l'émission des sources ponctuelles ou l'émission de la poussière dans le plan de la galaxie, ce qui peut biaiser l'estimation de l'offset comme on l'a vu dans la section 5.3.2. L'intervalle central de quantification ne coïncide donc pas toujours avec le maximum de la distribution. J'ai donc développé un outils permettant d'estimer les paramètres de compression de façon plus précise. Le principe repose sur les mêmes hypothèses que la compression. On suppose que le signal est gaussien mais que l'estimation de l'offset peut être biaisée. L'ensemble des échantillons peut alors être codé sur un intervalle différent de celui dans lequel il devrait être. On se base alors sur l'histogramme des différents intervalles pour estimer de façon précise les paramètres de compression.

On commence par déterminer la valeur la plus probable dans l'histogramme afin d'obtenir une première approximation du centre de la distribution. On compare ensuite l'amplitude relative des

intervalles avoisinant l'intervalle le plus peuplé à l'histogramme attendu pour une distribution gaussienne et on estime finalement l'écart type  $\sigma$  et le rapport  $q/\sigma$  à partir de la méthode des moindres carrés. On répète finalement cette opération sur un grand nombre de tranches de compression afin de moyenner les éventuelles erreurs. Cette méthode est plus robuste que la précédente, qui est basée uniquement sur le peuplement de l'intervalle central. Cette méthode basée sur l'histogramme m'a permis de vérifier la bonne valeur des paramètres de compression.

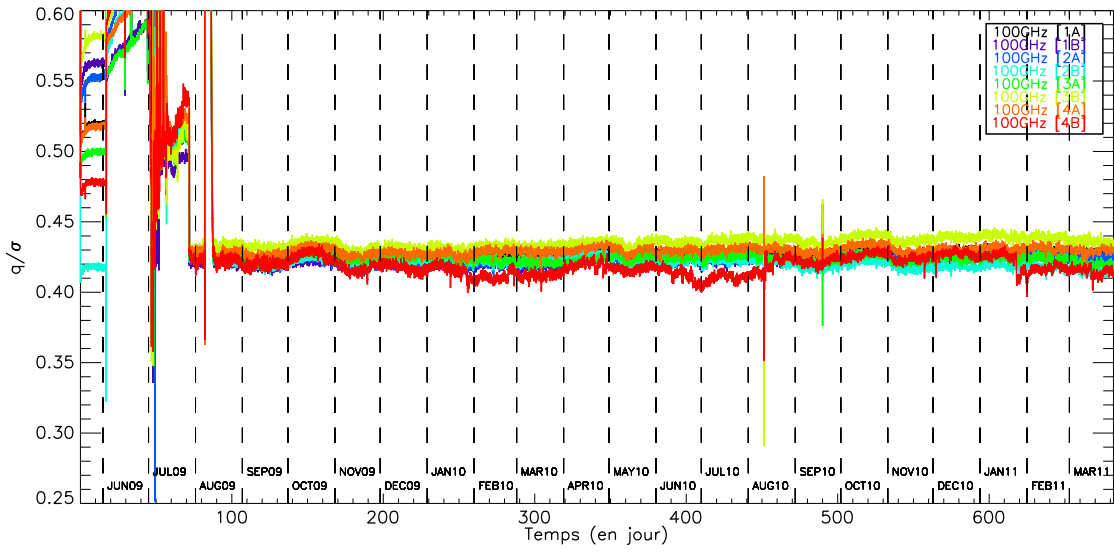
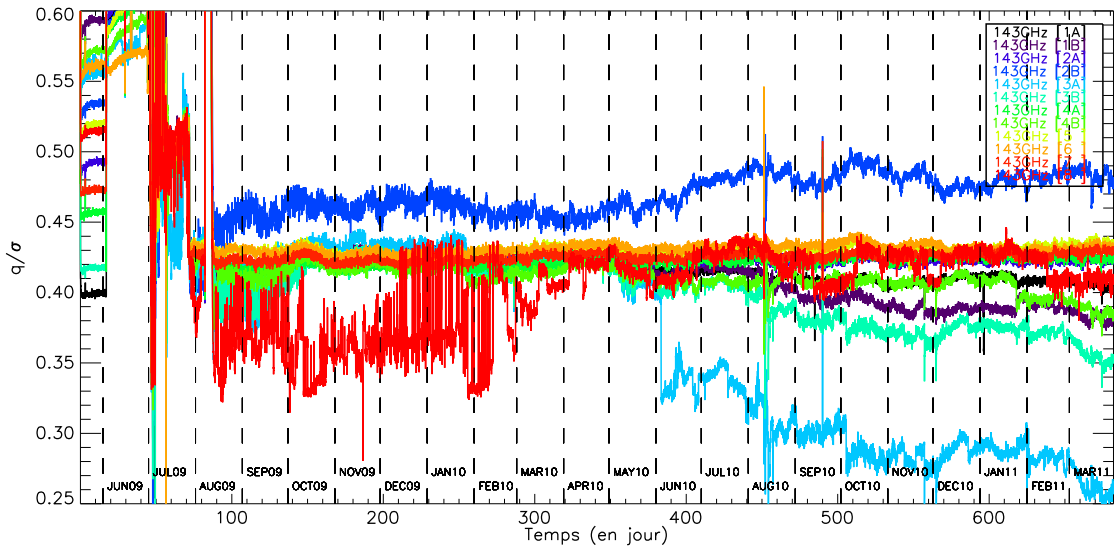
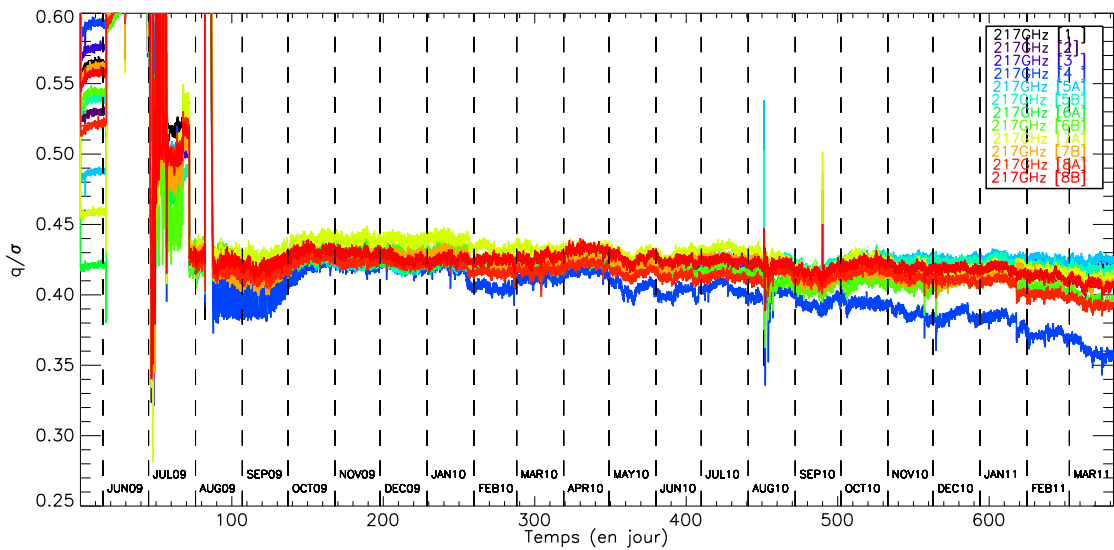
#### 5.4.6 Suivi de l'évolution des paramètres de compression

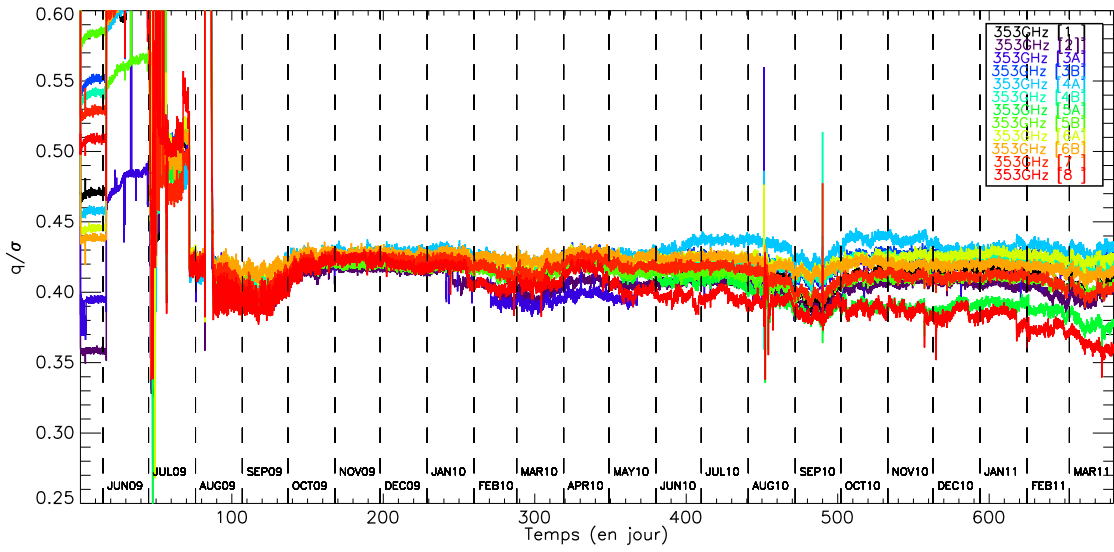
En plus de la surveillance quotidienne effectuée pendant la phase de communication avec le satellite, il est nécessaire de suivre l'évolution de la compression sur de longues périodes de temps afin de mettre en évidence des comportements non désirés. J'ai développé pour cela des programmes permettant de suivre l'évolution des paramètres au cours du temps. La méthode se base pour cela sur des objets de la base de données, contenant les statistiques de la fréquence des intervalles de quantification. Ils sont calculés au niveau de L1 et sont stockés dans un groupe nommé COMP\_STAT. Seule la fréquence des intervalles entre  $-5q$  et  $5q$  est donnée, puis la fréquence intégrée de l'ensemble des échantillons codés sur 10bits, 16bits et 22bits. Enfin, ce groupe contient également la fréquence des erreurs de compression.

Ces fréquences sont calculées pour chaque tranche de compression. Cependant, il n'est pas possible d'utiliser directement ces valeurs à cause de la manière dont sont traités les paquets au niveau L1. En effet, les données de chaque tranche de compression sont stockées de façon séquentielle dans les paquets PUS. Cependant, lorsque la taille totale des données et l'en-tête d'un paquet atteint la taille maximale d'un paquet, qui vaut 1024 octets, la suite des données est transmise dans un nouveau paquet. La suite des données se trouve donc dans le paquet suivant. Au niveau L1, les fréquences des intervalles des tranches de compression transmises sur plusieurs paquets sont calculées sur chaque morceau de la tranche séparément, comme si tous les échantillons étaient présents, ce qui biaise les calculs. En moyenne les données nécessitent 9 paquets. On trouve alors approximativement 8 voies sur 72, soit plus de 10%, qui possèdent des fréquences mal calculées.

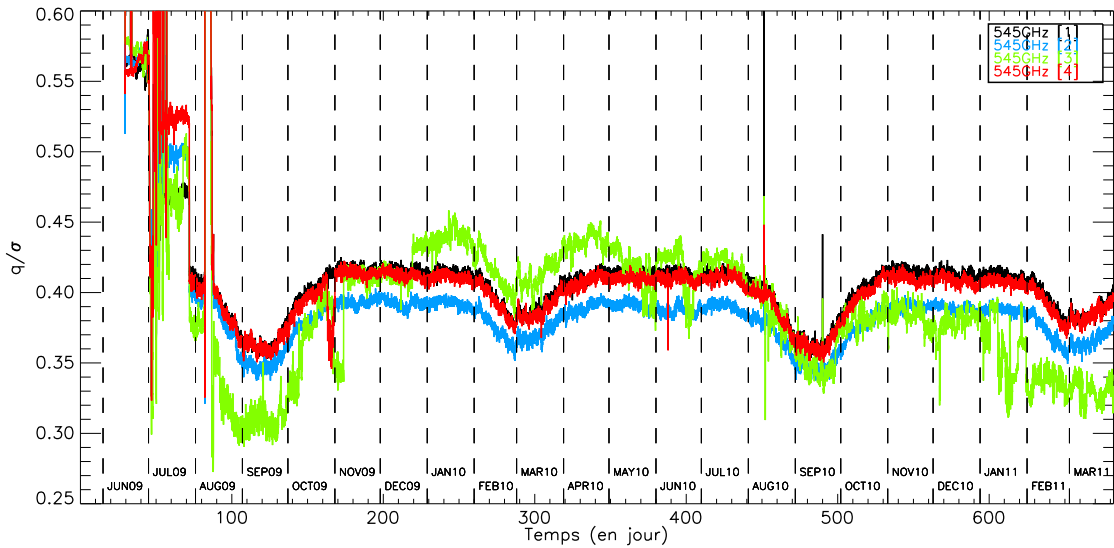
Afin de répondre à ce problème, nous commençons par recréer un groupe contenant les fréquences corrigées, possédant exactement une valeur par tranche de compression, même sur celles qui sont séparées sur plusieurs paquets. Pour cela, il faut commencer par identifier les échantillons des objets du groupe COMP\_STAT qui possèdent des valeurs incorrectes. Nous nous basons sur la différence de temps entre deux échantillons successifs. Tant que cette valeur n'est pas un multiple de la durée d'une tranche de compression, nous corrigeons les valeurs. La fréquence pour la tranche de compression complète se calcule facilement à partir des fréquences pour chaque morceau. En effet, connaissant la durée de chaque morceau, on en déduit la proportion de la tranche complète que cela représente.

Enfin, en utilisant ce groupe et la méthode d'estimation du rapport  $q/\sigma$  détaillée dans la section 5.4.4, nous avons pu tracer l'évolution de ce rapport depuis le lancement du satellite, pour l'ensemble des voies de l'instrument HFI. Elle est résumée dans les figures 5.9(a) à 5.9(h). On peut voir qu'à partir du 29 juillet 2009, date à laquelle les paramètres de compression ont été fixés à leur valeur nominale, l'évolution du rapport  $q/\sigma$  est relativement stable sur l'ensemble des voies. Pour les bolomètres, on observe une évolution globale similaire pour la plupart des voies à une fréquence donnée. On remarque cependant certaines voies possédant des variations significatives par rapport au comportement global des autres voies à la même fréquence. Pour certaines, ces variations sont temporaires, pour d'autres on peut les mettre en évidence depuis le début des observations. Ces voies présentent systématiquement un bruit additionnel, appelé bruit télégraphique qui se traduit par une variation brutale du niveau de bruit entre au moins deux états différents. Suivant les périodes considérées, le niveau de bruit n'est donc pas le même, ce qui conduit à une variation du rapport  $q/\sigma$ . On peut observer que dans la plupart des cas, pour ces voies, le réglage des paramètres de compression initiaux tient compte du niveau de bruit le plus faible car le rapport  $q/\sigma$  des voies possédant un bruit télégraphique est plus faible que la valeur de référence. L'impact de la compression sur ces voies n'est donc jamais plus important que celui qu'on aurait avec le rapport  $q/\sigma$  de référence. Par ailleurs, on peut voir une variation périodique plus prononcée sur les voies à haute fréquence. Elle est produite par les fortes variations du signal galactique dans la zone du centre galactique qui modifient la répartition des échantillons dans les intervalles de quantification. Enfin, l'évolution pour les thermomètres montre également une relative stabilité. On observe cependant sur les thermomètres des PID 1R, 2N et 2R ainsi que sur les thermomètres de l'étage à 0.1K une variation qui semble

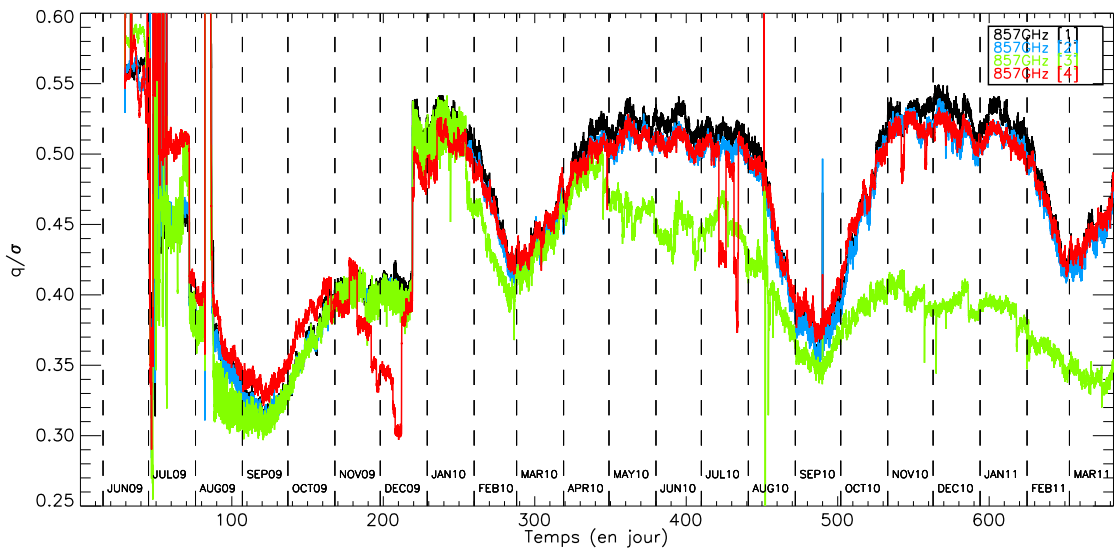
(a) Évolution du rapport  $q/\sigma$  pour les voies à 100GHz.(b) Évolution du rapport  $q/\sigma$  pour les voies à 143GHz.(c) Évolution du rapport  $q/\sigma$  pour les voies à 217GHz.



(d) Évolution du rapport  $q/\sigma$  pour les voies à 353GHz.



(e) Évolution du rapport  $q/\sigma$  pour les voies à 545GHz.



(f) Évolution du rapport  $q/\sigma$  pour les voies à 857GHz.

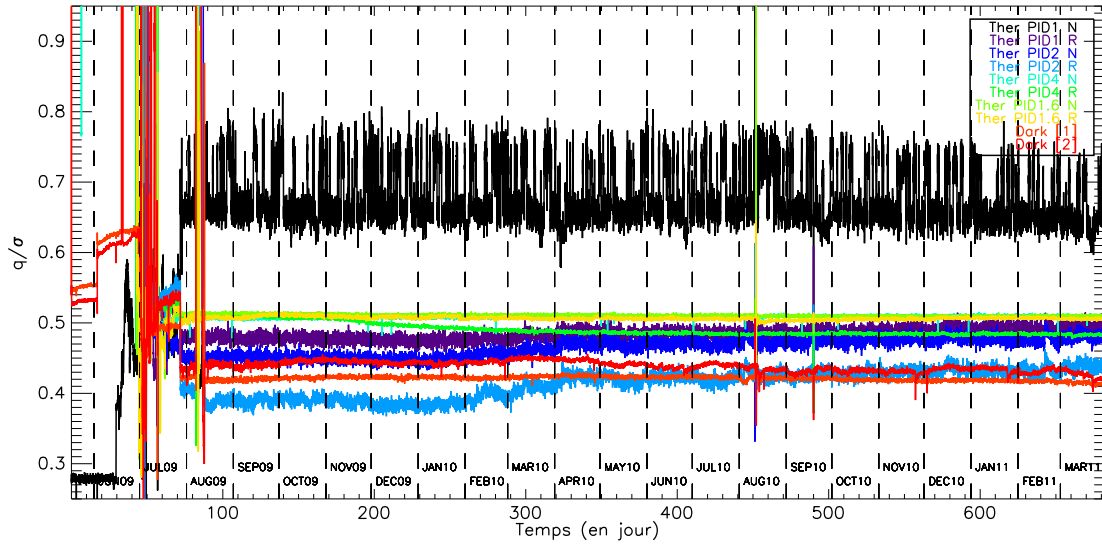
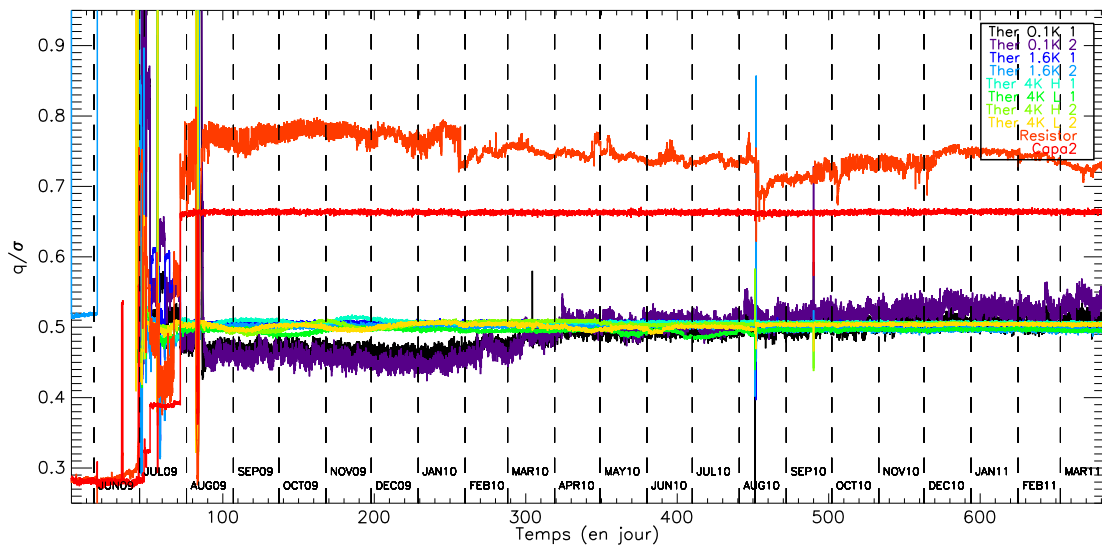
(g) Évolution du rapport  $q/\sigma$  pour les thermomètres.(h) Évolution du rapport  $q/\sigma$  pour les thermomètres.

FIGURE 5.9 – **Évolution du rapport  $q/\sigma$  au cours du temps.** Elle est tracée du 16 mai 2009 au 29 mars 2011. Chaque point est calculé comme la moyenne sur 60 minutes. L'évolution du rapport  $q/\sigma$  au cours du temps est relativement constante, après la phase initiale de réglages qui s'est terminée le 29 août 2009. On observe, un comportement particulier sur les voies possédant du bruit télégraphique. Pour ces voies le niveau de bruit oscille entre plusieurs états, ce qui entraîne un changement du rapport  $q/\sigma$ . On peut également voir l'effet produit par l'arrêt du compresseur à 4K début août 2009, ainsi que le changement du refroidisseur à adsorption le 10 août 2010 qui ont entraîné dans les deux cas un réchauffement global.

être corrélée. Sur la figure 5.10 on compare l'opposée de leurs variations, normalisées à 1 sur les 1000 premiers points, au taux de rayons cosmiques, compté par les SREM. On observe une forte corrélation, ce qui signifie qu'il existe une anti-corrélation entre le taux de rayons cosmiques et le rapport  $q/\sigma$  de ces voies. Elle s'explique par la variation du bruit effectif sur les thermomètres. Le bruit est plus important lorsque le taux de rayons cosmiques est plus élevé, ce qui diminue le rapport  $q/\sigma$  et explique l'anti-corrélation.

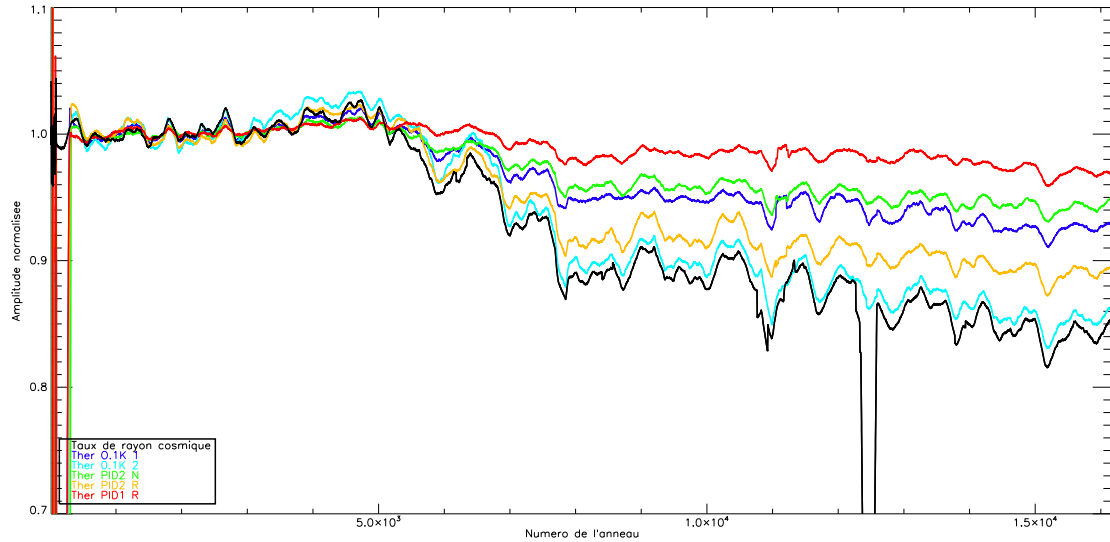


FIGURE 5.10 – **Corrélation entre le rapport  $q/\sigma$  de thermomètres et le taux de rayons cosmiques.**

L'impact des rayons cosmiques dépose localement une grande quantité d'énergie. L'évolution du taux de rayons cosmiques est tracée en noir. Elle est mesurée par le SREM. Nous traçons ici, l'évolution donnée par l'objet TC1\_countrate qui contient le taux de particules d'énergie supérieure à 0.085Mev. Les courbes en couleur représentent l'opposée des variations du rapport  $q/\sigma$  pour les thermomètres des PID et ceux de l'étage à 1.8K. Pour l'ensemble des courbes, les valeurs sont moyennées sur chaque anneau et sont lissées avec une fenêtre glissante de largeur 100. L'ensemble des courbes est normalisé à 1 sur les 1000 premiers points.

## 5.5 Suivi du débit

Une fois les réglages de la compression fixés, le débit devrait être constant si la distribution en amplitude du signal ne changeait pas. Pendant toute la durée de la mission, il est donc intéressant de faire un suivi du débit afin de voir l'évolution de l'état des détecteurs. Par ailleurs, si la limite sur le débit imposée à l'instrument HFI est atteinte, cela entraîne des pertes de données. Le suivi quotidien permet donc de détecter toute variation anormale afin de prendre les mesures nécessaires en cas de problèmes pour ne pas perdre de données. Enfin, l'efficacité de la compression est directement liée au niveau de bruit de chaque voie (voir section 5.2). Ainsi, le suivi du débit permet de détecter des variations singulières du niveau de bruit qui peuvent être causées par exemple par une variation de la température, des anomalies dans le fonctionnement des bolomètres ou de leur électronique de lecture, une variation du taux de rayons cosmiques...

### 5.5.1 Débit d'un signal purement gaussien

Il est possible de calculer le débit après compression d'un signal purement gaussien d'écart type  $\sigma$ , en fonction de la taille du pas de quantification  $q$  de la compression. La densité de probabilité d'un signal gaussien de moyenne nulle et d'écart type  $\sigma$  est donnée par :

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad (5.18)$$

On peut alors calculer la probabilité qu'un échantillon tombe dans l'intervalle de quantification  $i$  en injectant l'équation 5.18 dans l'équation 5.3. Puis, connaissant le nombre de bits  $b_i$  utilisé par le mot codant pour l'intervalle  $i$ , on peut finalement en déduire le nombre moyen de bits par échantillons, après compression d'un signal gaussien. Il est donné par la formule :

$$d = \sum_i p_i b_i \quad (5.19)$$

D'après le tableau 5.1 donnant le nombre de bits par intervalle, on trouve un débit moyen après compression de 4.03 bits par échantillon pour un rapport  $q/\sigma$  de 0.4 et de 3.78 bits par échantillon pour un rapport de 0.5.

Finalement, on peut calculer le débit nécessaire pour transmettre les données compressées, provenant de l'ensemble des voies de l'instrument HFI (54 voie avec  $q/\sigma = 0.4$  et 18 avec  $q/\sigma = 0.5$ ). Étant donnée la fréquence d'échantillonnage de  $F_{s\text{amp}} = 180.38\text{Hz}$ , on obtient un débit moyen de  $(54 \times 4.027 + 18 \times 3.777) \times F_{s\text{amp}}/1024$  soit 50.28 kbits/s. Il reste donc une marge de 17kbits/s pour atteindre la limite de débit disponible pour les données. Il faut garder à l'esprit que ces valeurs du débit par voies et du débit total, sont des bornes inférieures. Dans le cas réel, certains événements peuvent produire des échantillons qui peuplent les intervalles éloignés de la moyenne et qui sont codés sur un plus grand nombre de bits. Le débit réel sera donc plus important, comme on le verra dans les sections suivantes.

### 5.5.2 Évolution réelle du débit

Les figures 5.11(a) à 5.11(f) montrent l'évolution du débit, depuis le début de la mission, pour l'ensemble des voies, regroupées par fréquence d'observation. On peut voir que le débit est systématiquement supérieur au débit théorique attendu pour un signal purement gaussien. Le débit additionnel observé sur l'ensemble des voies est produit en grande partie, par les rayons cosmiques (voir section 5.6.3) et dans une moindre mesure par la variation de l'émission de la poussière dans le plan galactique (voir section 5.5.3).

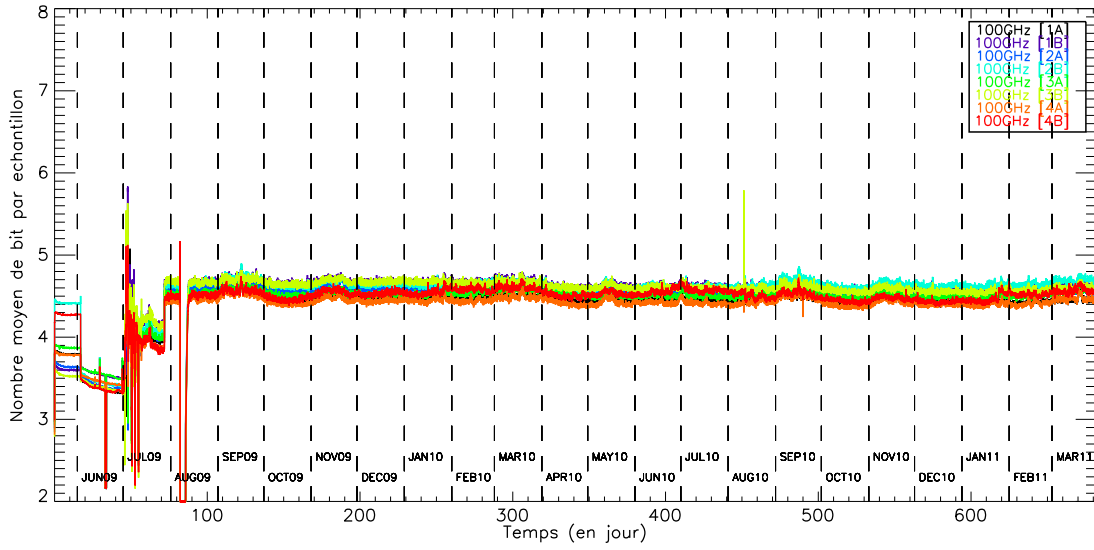
On observe un grand nombre d'impacts de rayons cosmiques sur l'ensemble des voies (plusieurs dizaines par minutes). Lorsqu'ils frappent un bolomètre, ils peuvent déposer localement une grande quantité d'énergie. Cela entraîne une brutale augmentation de la température du bolomètre suivie d'une décroissance exponentielle qui dépend des caractéristiques fines du bolomètre. Cela se traduit donc par une modification de la répartition en amplitude des échantillons. En particulier la forte augmentation de l'amplitude entraîne un peuplement des intervalles de quantification dans les régions éloignées de la moyenne, qui sont codés sur un nombre plus important de bits, ce qui conduit à une augmentation du débit moyen. Le taux d'impact de rayons cosmiques dépend essentiellement de leur énergie. Il varie peu avec le temps, c'est pourquoi le débit moyen par voie reste stable depuis le début des observations.

### 5.5.3 Effet de l'émission de la poussière galactique

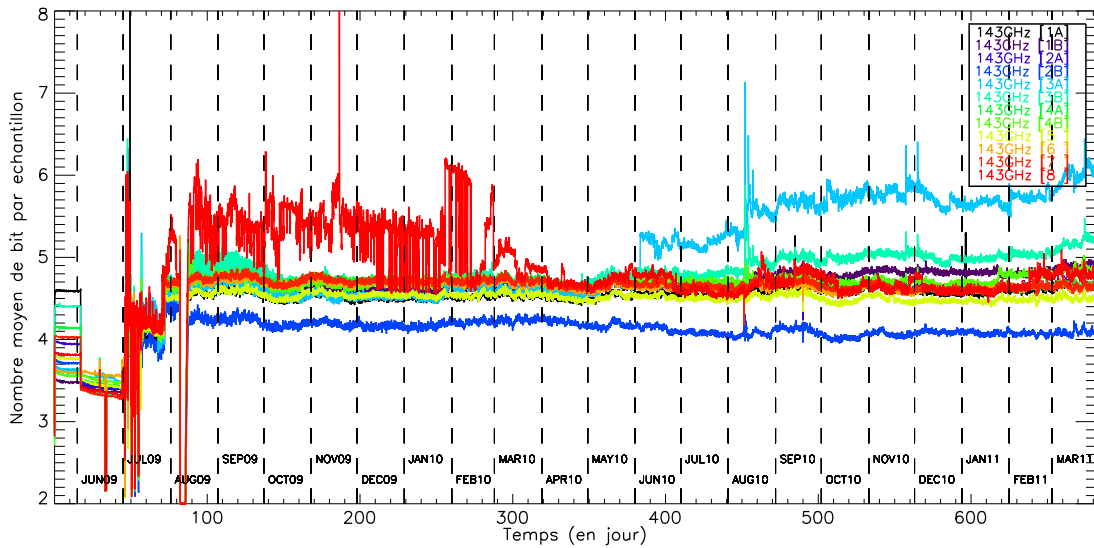
Si on observe l'évolution du débit sur de courtes périodes, on remarque une augmentation périodique systématique du nombre de bits moyen par échantillon sur certaines tranches de compression. Cette variation est produite par l'observation du plan galactique lors de chaque cercle. On croise ainsi deux fois le plan galactique par minute environ. Dans cette région l'émission de la poussière varie brutalement et l'hypothèse d'un signal constant sur une tranche de compression n'est plus vraie. La forte variation d'amplitude sur l'échelle d'une tranche de compression peuple des intervalles de quantification éloignés de la moyenne, ce qui conduit à une augmentation du débit.

Sur la figure 5.13 on a tracé l'évolution du nombre de bits moyen par échantillon pour chaque tranche de compression, parallèlement à l'évolution du signal, sur une période de dix minutes, pour une voie à 217GHz et une voie à 857GHz. À basse fréquence, la variation principale du signal, est produite par le dipôle cosmologique. Comme on peut le voir l'échelle de variation est suffisamment lente pour être négligeable sur une tranche de compression. On remarque également des variations courtes et brutales qui sont provoquées par l'observation du plan galactique. Les tranches de compression dans lesquelles ces variations tombent, possèdent alors un nombre de bits par échantillon plus élevé. Pour la voie à 857GHz, les variations sont largement dominées par l'émission de la poussière. L'effet sur le débit est encore plus marqué.

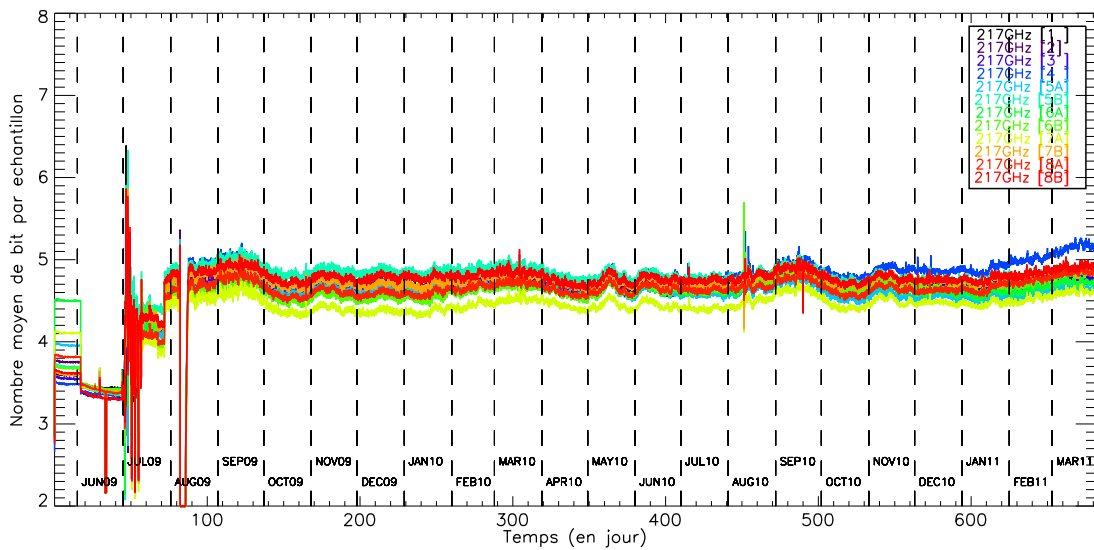




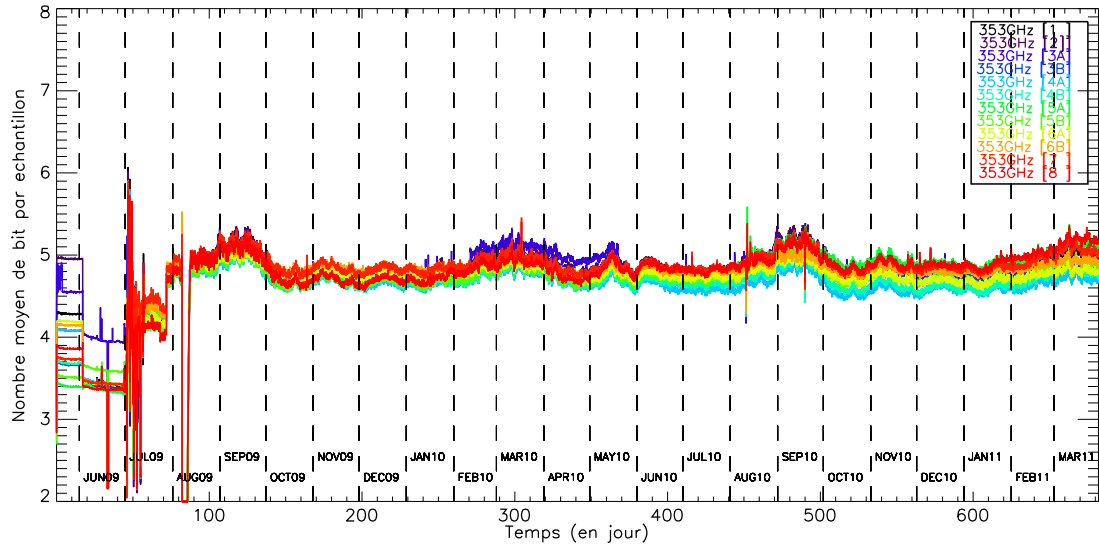
(a) Débit des voies à 100GHz.



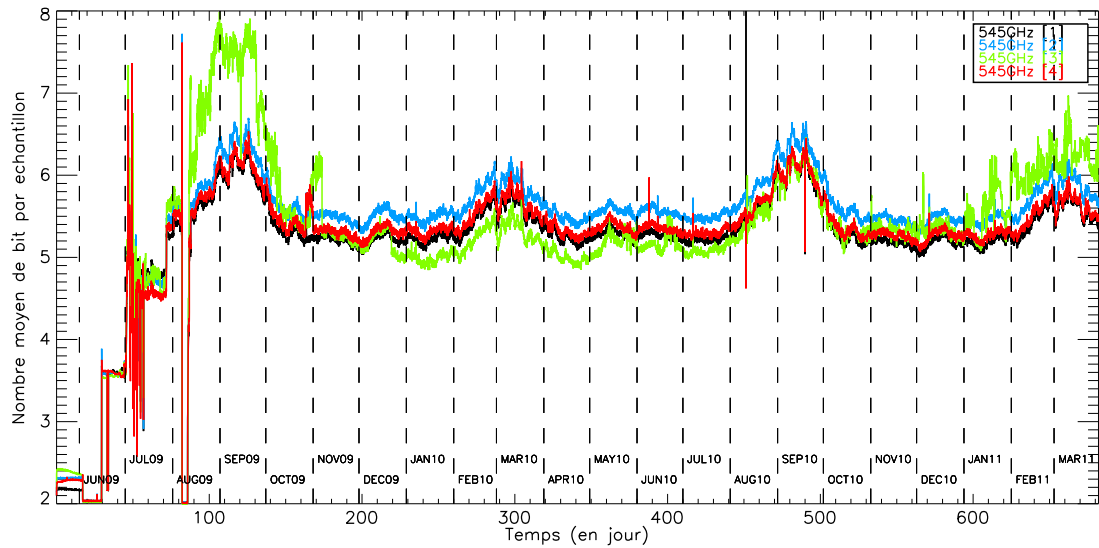
(b) Débit des voies à 143GHz.



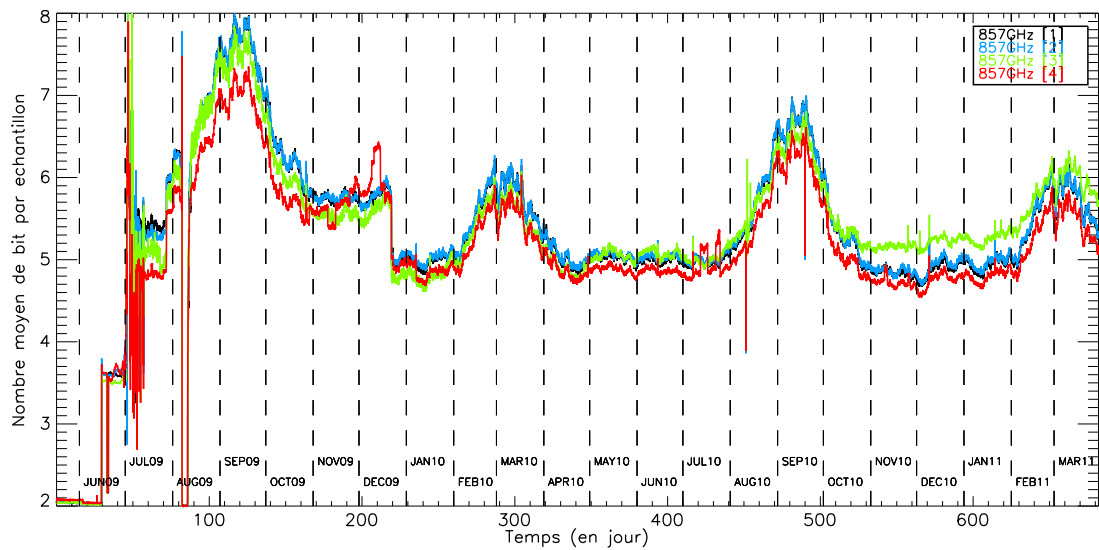
(c) Débit des voies à 217GHz.



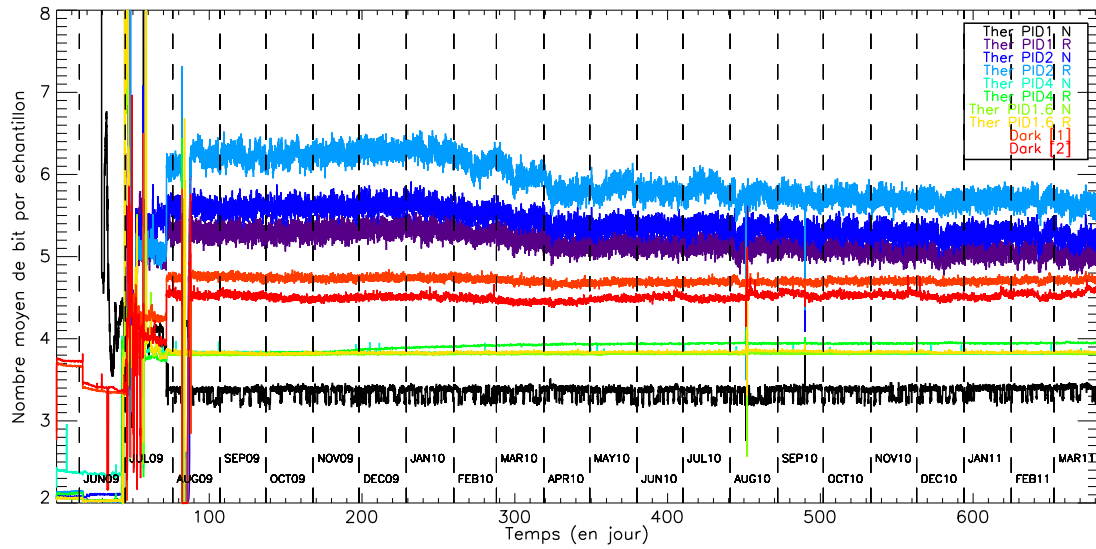
(d) Débit des voies à 353GHz.



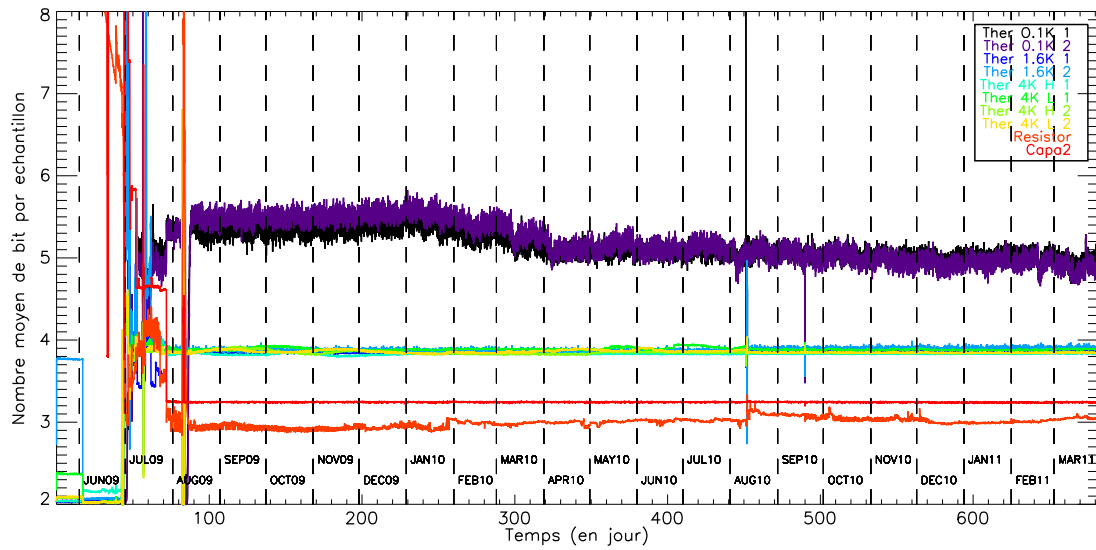
(e) Débit des voies à 545GHz.



(f) Débit des voies à 857GHz.



(g) Débit des thermomètres.



(h) Débit des thermomètres.

FIGURE 5.11 – Évolution du débit moyen par échantillon pour chaque voie. Elle est tracée, du 16 mai 2009 au 29 mars 2011. Chaque point représente la moyenne sur 60 minutes. Les voies à la même fréquence sont regroupées sur les différents graphiques.

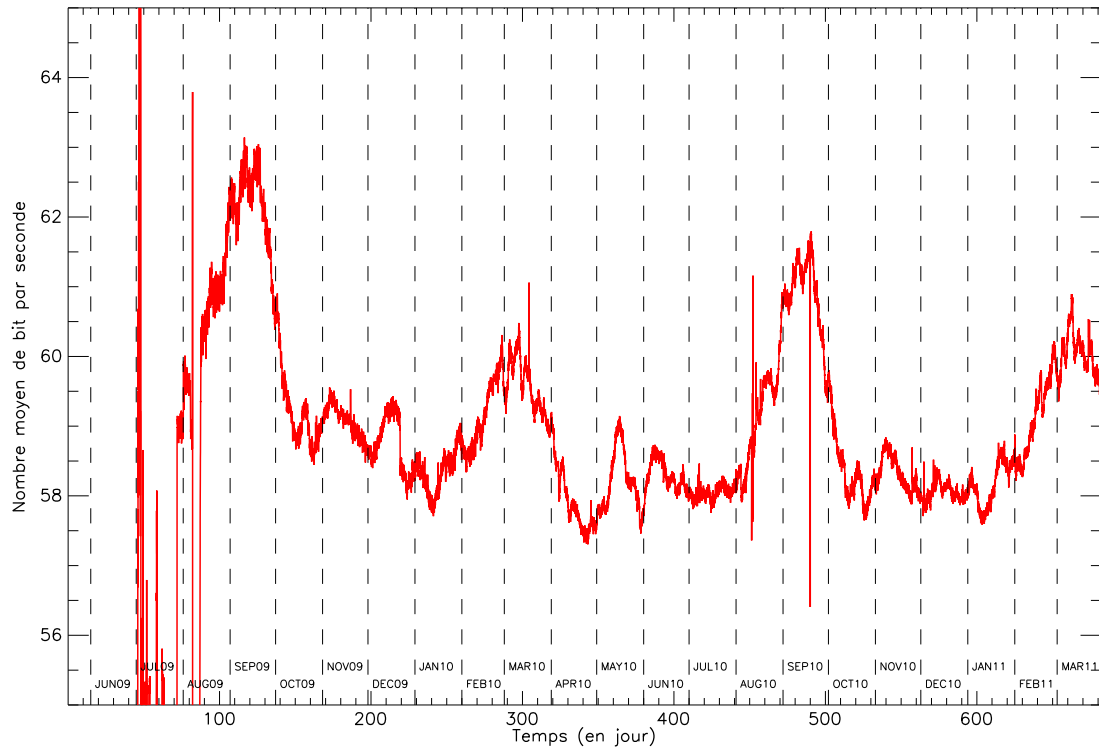


FIGURE 5.12 – **Évolution du débit total.** Chaque point représente la moyenne, sur 60 minutes, de la somme des débits de l'ensemble des 72 voies de l'instrument HFI. L'évolution est tracée du 16 mai 2009 au 29 mars 2010. Les variations sont dominées par l'évolution du débit des voies à haute fréquence. Pour ces voies, le débit varie suivant la position et l'orientation avec laquelle on croise le plan galactique. On peut voir qu'on n'a cependant jamais dépassé les 65kb/s disponibles pour les données. Le débit calculé ne prend pas en compte le débit additionnel nécessaire pour les en-têtes dont la taille est fixe.

De façon plus globale, on observe que le débit est très stable sur l'ensemble des voies. On peut remarquer cependant des variations notables sur plusieurs semaines sur les voies à 857GHz et dans une moindre mesure sur les voies à 545GHz. À ces longueurs d'onde, le signal est dominé par l'émission de la poussière qui varie fortement avec la direction d'observation. L'émission de la poussière est plus intense dans le plan galactique et diminue à haute latitude galactique. Elle atteint un maximum au niveau du centre galactique qui a été observé pour la première fois entre le 12 et le 13 septembre 2009. C'est à cette date que le débit total a atteint son niveau maximum. L'augmentation du débit est causée par les fortes variations de l'émission de la poussière, à l'échelle d'une tranche de compression, dans la zone proche du centre galactique. Dans ces régions, la répartition en amplitude des échantillons, couvre une large gamme dynamique. L'approximation du signal par un offset environ constant auquel s'ajoute un bruit blanc gaussien, sur laquelle se base l'algorithme de compression n'est plus valable. Les fortes variations d'amplitude à l'intérieur d'une tranche de compression donne lieu à un grand nombre d'échantillons se situant loin de la moyenne et qui peuplent donc les intervalles de quantification moins probables, utilisant un plus grand nombre de bits. Cela se traduit par une augmentation du débit. Il faut également noter que la stratégie de pointage de PLANCK est particulièrement défavorable pour le débit dans ces régions. En effet, en plus de croiser la région du centre galactique qui possède la plus forte intensité de l'émission de la poussière, le balayage du ciel est relativement parallèle au plan galactique à cette époque d'observation. La direction d'observation tombe donc en majeure partie dans les régions de forte émission galactique.

#### 5.5.4 Effet du transpondeur

Si on regarde en détail, l'évolution du débit dans les premiers mois de la mission, on peut apercevoir des variations périodiques corrélées sur toutes les voies (voir figure 5.14). Le début de

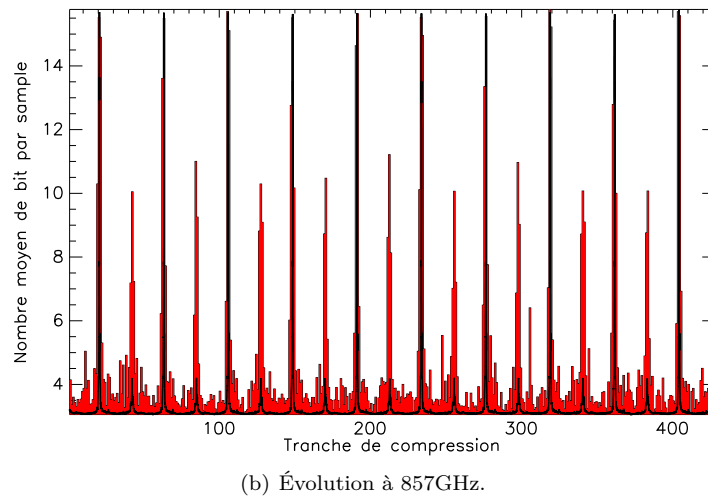
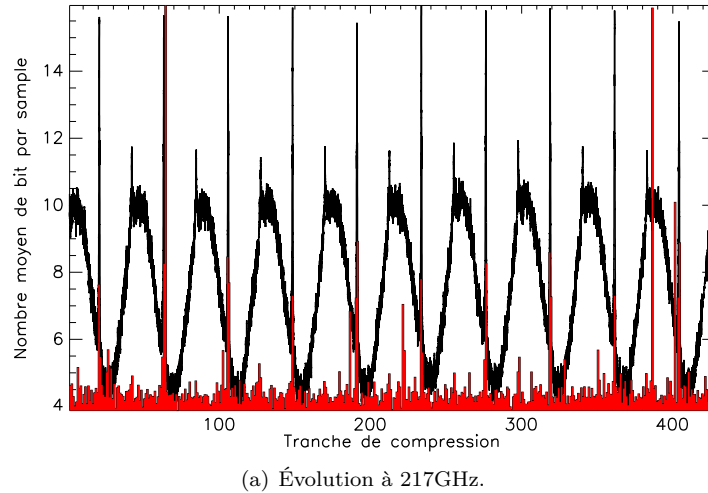


FIGURE 5.13 – **Impact de l'émission de la galaxie sur le débit.** L'intensité de l'émission de la poussière est plus élevée dans la plan galactique. Cela se traduit par une variation rapide du signal à l'échelle d'une tranche de compression. L'efficacité de la compression est alors affectée car de nombreux échantillons peuplent des intervalles de quantification éloignés de la moyenne. En noir on peut voir l'évolution du signal sur une échelle arbitraire. Parallèlement, l'évolution du nombre de bits moyen par échantillon, pour chaque tranche de compression est représentée par les rectangles rouges. La figure du haut représente l'évolution pour la voie 5A à 217GHz, la figure du bas celle de la voie 2 à 857GHz. On remarque qu'à haute fréquence, l'émission de la galaxie domine largement les variations tandis qu'à 217GHz la variation principale observée est celle du dipôle. L'effet sur le débit est plus important sur ces voies. L'évolution est tracée du 11 décembre 2010 à 13h00 au 11 décembre 2010 à 13h10.

l'augmentation observée, coïncide avec la période quotidienne de transfert des données qui dure près de trois heures. Elle est produite par la mise sous tension du transpondeur, chargé de transmettre les données vers la Terre, dont l'électronique réchauffe le satellite, ce qui modifie le fonctionnement thermique de l'étage cryogénique à 4K. Ces variations thermiques produisent une modification de l'intensité des raies parasites générées par le mouvement mécanique du piston du compresseur de l'étage cryogénique à 4K, ce qui fait varier le bruit des mesures et change l'efficacité de la compression, conduisant finalement, à une variation du débit. À partir du 25 janvier 2010, le transpondeur a été laissé allumé en permanence afin de permettre à la température de se stabiliser et d'éviter les variations quotidiennes. À partir de cette date, les variations ne sont plus visibles, comme on peut l'observer sur la figure 5.14 qui compare le débit total à un an d'intervalle avant et après la mise sous tension permanente du transpondeur.

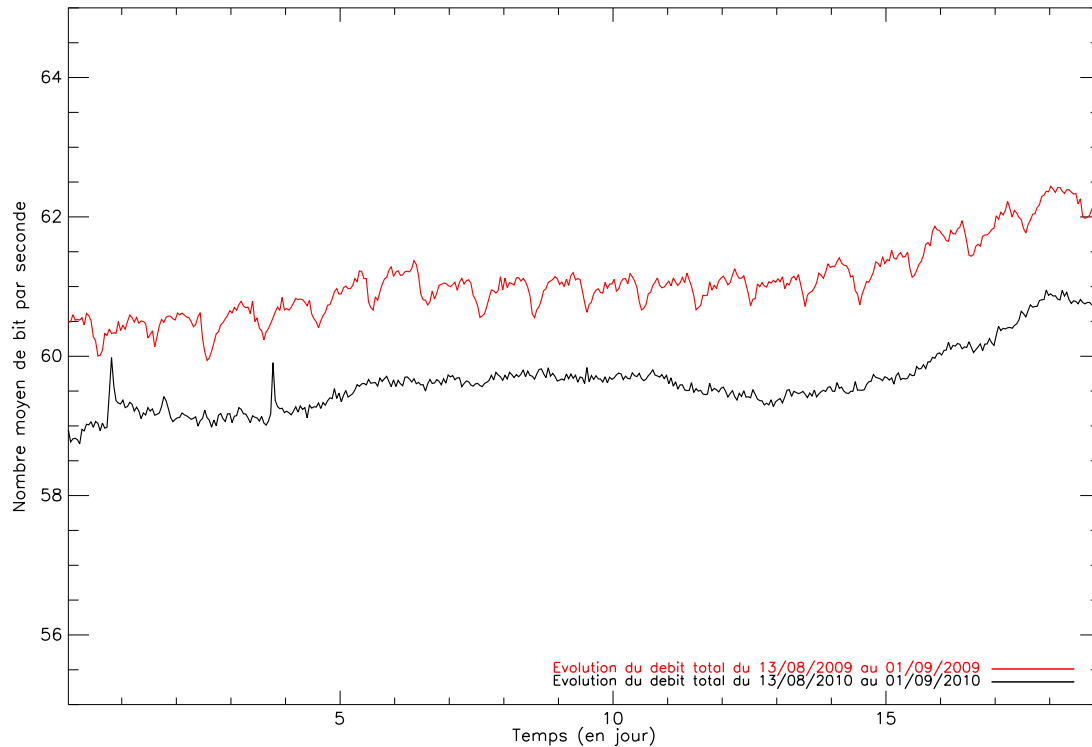


FIGURE 5.14 – **Effet sur le débit de la mise sous tension du transpondeur.** La courbe rouge représente la variation du débit total du 16 août 2009 au 01 septembre 2009, la courbe en noir représente l'évolution un an plus tard du 16 août 2010 au 01 septembre 2010. Les points de mesure correspondent à la moyenne sur des tranches d'une heure. On peut voir sur la courbe d'août 2009, l'effet de la mise sous tension quotidienne du transpondeur, pendant la période de communication avec la Terre. Cette mise sous tension réchauffe le satellite et conduit à une variation du bruit de mesure entraînant finalement une variation du débit des données. Afin de limiter ces variations le transpondeur a été mis sous tension permanente depuis janvier 2010, ce qui supprime totalement ces variations comme on peut le voir sur la courbe du débit total en août 2010 (après la mise sous tension permanente).

### 5.5.5 Effet du bruit télégraphique

On constate que d'une façon générale, les variations du débit sont similaires sur l'ensemble des voies à une fréquence donnée. Quelques exceptions ont tout de même été mises en évidence. Sur la huitième voie à 143GHz, par exemple, on observe des variations brutales du débit que l'on ne retrouve pas sur les autres voies à cette fréquence. Ces variations sont causées par un comportement anormal de ce détecteur. Il possède un bruit particulier appelé bruit télégraphique, qui se traduit par des sauts rapides du niveau de bruit, qui oscille entre deux états (voire plus). L'efficacité de la compression est affectée par ce comportement car un grand nombre d'échantillons vont périodiquement peupler des intervalles différents lorsque le bruit change d'état. Une étude approfondie montre que d'autres voies possèdent du bruit télégraphique, dans une moindre mesure. On peut observer par exemple que la troisième voie à 545GHz a un comportement particulier comparée aux autres voies à la même fréquence. Enfin, la quatrième voie à 857GHz ainsi que le thermomètre du PID 1R possèdent également du bruit télégraphique.

La sensibilité de l'efficacité de la compression au bruit télégraphique, permet d'utiliser la mesure de l'évolution du débit ou des paramètres de compression pour tracer la présence de ce bruit. Ce n'est pas une mesure très précise mais elle a été utilisée pour confirmer des analyses dédiées ou pour mettre en évidence l'apparition de ce bruit.

### 5.5.6 Événements particuliers

Plusieurs événements particuliers ont eu lieu depuis le début des observations. Le 6 août 2009, suite au déclenchement de plusieurs alertes de sécurité automatiques, le compresseur à 4k s'est mis hors tension. Cet événement a eu lieu juste après la fin de la transmission avec la Terre et n'a donc été détecté que plus de 20 heures plus tard. L'ensemble de l'étage à 4K s'est alors fortement réchauffé suivi par les sous-étages à 1.5k, 100mk... La forte augmentation de température a entraîné une augmentation du bruit puis du débit ce qui a conduit à la perte de données. On peut voir, sur les figures 5.11(a) à 5.11(f) que le débit moyen a chuté à zéro pendant plusieurs jours en attendant le retour à une température normale. Suite à la détection de l'arrêt du compresseur il a été remis sous tension : les causes du déclenchement des alertes de sécurité ont été étudiées et une solution a été trouvée afin d'empêcher un renouvellement de la panne.

Par ailleurs, le 11 août 2010, le refroidisseur à adsorption utilisé depuis le début de la mission a été remplacé par le second refroidisseur présent sur la satellite. C'est une opération qui était prévue et qui a été réalisée avec succès. Cependant, pendant une courte période, ce changement a entraîné une augmentation de la température, qui a conduit à une augmentation du niveau de bruit de l'ensemble des voies, puis du débit. On peut observer le pic correspondant sur les figures 5.11(a) à 5.11(f), visible sur l'ensemble des voies.

## 5.6 Suivi des erreurs de compression

Dans la section 5.3.3, on a vu que dans certaines situations particulières, le DPU pouvait générer des erreurs de compression. Quand plus de 64 échantillons ont été codés sur 22bits, les données suivantes qui devraient être codées sur 22bits provoquent une erreur de compression. Si l'amplitude d'un échantillon ne peut pas être codée sur 22bits, c'est-à-dire si sa valeur (en unité ADU) après démodulation et soustraction de l'offset, est supérieure à  $66584q$  ou inférieure à  $-66584q$ , le DPU génère également une erreur de compression. Par ailleurs, quand l'algorithme de compression n'est pas assez efficace et que le débit moyen intégré est trop élevé, le DPU déclenche une erreur de type EOS. Enfin, dans le pire des cas, quand la compression est tellement inefficace que trop de paquets sont produits à la fois, certains paquets peuvent être perdus, lorsque le tampon en mémoire du DPU est totalement rempli, ce qui provoque des trous dans les données.

Pour s'assurer du bon fonctionnement de l'instrument HFI, il est nécessaire de faire un suivi constant des erreurs de compression et de la perte de données. Comme on l'a montré dans la section 5.3, l'algorithme ne doit pas produire d'erreurs de compression si les données en entrée n'ont pas subi de perturbations particulières. J'ai donc développé un ensemble de programmes permettant de surveiller la présence d'erreurs et j'ai effectué un suivi systématique depuis le lancement de la mission.

### 5.6.1 Recherche de trous dans les données provoqués par la saturation du tampon mémoire

Au niveau L1, les trous dans les données sont remplacés par des 0 dans les TOIs du signal brut non démodulé. La méthode de détection des trous est alors relativement simple. Il suffit de lire ces TOIs pour chaque voie et rechercher les valeurs à 0. En principe, dans le cas général, d'autres situations peuvent conduire le traitement en L1 à produire des échantillons à 0, comme par exemple les erreurs de compression. Cependant, comme plusieurs voies sont regroupées dans un paquet, la perte de données provoquée par la saturation du tampon en mémoire du DPU, est visible sur plusieurs voies simultanément, en touchant un nombre d'échantillons multiple de 254, sur les voies dont l'ensemble des données de la tranche de compression sont contenues dans le paquet. Cependant, d'autres causes peuvent conduire à la perte de paquets comme les problèmes de transmission, ou de traitement au niveau du MOC. Il n'est pas possible de distinguer ces cas, à partir des informations contenues dans la base de données de PLANCK car aucun marquage particulier n'est effectué, au moment de l'injection des données par L1, pour distinguer les différentes origines des erreurs. Il est donc nécessaire de vérifier la présence de ce type de problèmes auprès de l'équipe de L1.

Depuis le début de la mission, plusieurs milliers de trous ont été détectés. La majorité ne touchent que quelques échantillons et a été provoqué par des erreurs de compression. Pour les autres, car-

respondant à une perte sur un nombre d'échantillons multiple de 254, une remontée systématique a été effectuée auprès des responsables du traitement en L1. Aucun trou dans les données, lié à la saturation du DPU, n'a été mis en évidence. La plupart des trous ont été observés avant le 23 septembre 2009 et ont pu être expliqués par des erreurs de traitement en L1, identifiées et corrigées par l'équipe responsable, à partir du 23 septembre 2009. Un des trous a été identifié comme le déclenchement d'un EOS (voir section suivante). Depuis le 23 septembre 2009, quatre trous dans les données supplémentaires ont été détectés, certains dépassant les 3000 secondes. Ils sont tous liés à des problèmes de transmission pour lesquels, les données ne sont pas perdues mais n'ont pas pu être injectées dans la base de données. Elles seront présentes dans la prochaine version de la base de données.

### 5.6.2 Recherche des EOS

Le déclenchement d'un EOS entraîne l'arrêt de l'envoi des « paquets science » par le DPU. Cela se traduit par une perte de données sur l'ensemble des voies simultanément. Dans la base de données, un objet permet de savoir si le signal d'au moins une des voies est disponible pour un temps donné. Pour détecter les EOS, nous avons donc utilisé cet objet, appelé ONE\_OR\_MORE\_BOLO, qui vaut 0 lorsqu'aucun des échantillons n'a été reconstruit. Il n'existe aucun autre moyen à partir des informations contenues dans la base de données de détecter ce type d'erreur. Elle ne sont pas marquées de façon distincte par le niveau L1. Comme d'autres causes peuvent entraîner la perte d'un ou plusieurs paquets, conduisant dans certains cas à la perte de données simultanément sur l'ensemble des voies, il est possible d'avoir une valeur de 0 pour l'objet ONE\_OR\_MORE\_BOLO, même si aucun EOS n'a eu lieu.

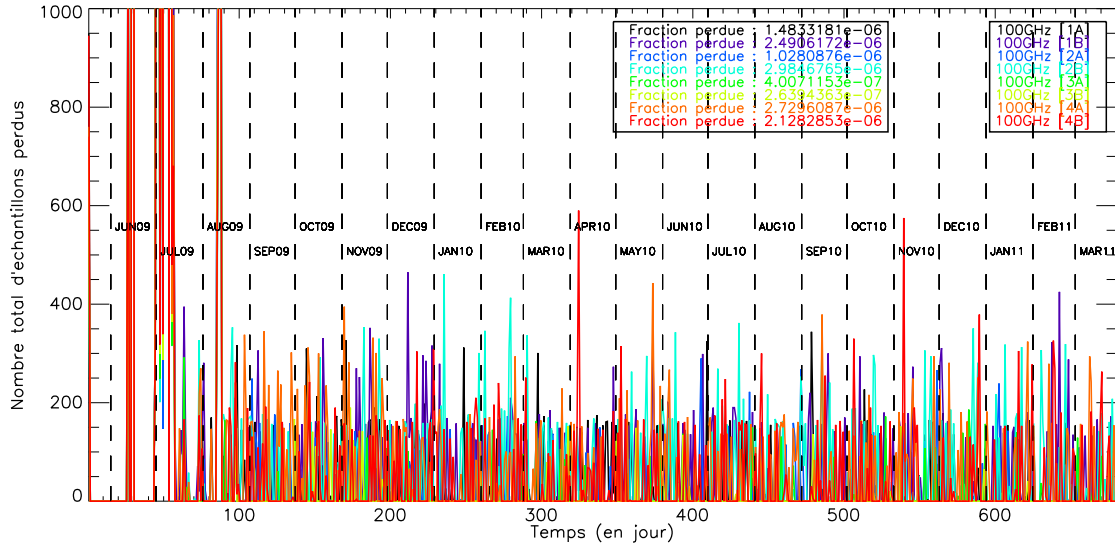
Depuis le début des observations régulières, la recherche systématique a permis de détecter 40 périodes, où on observe une perte totale des données simultanément sur l'ensemble des voies. Une seule est attribuée au déclenchement d'un EOS. Elle s'est produite, début août 2009, suite à l'arrêt de l'étage cryogénique à 4K, conduisant à un réchauffement global de l'instrument. Toutes les autres pertes de données détectées sont attribuées à des erreurs de traitement au niveau L1 conduisant à la perte de paquets. Elles sont survenues pour la plupart, avant le 23 septembre 2009, date à laquelle une correction dans le traitement a été implémentée en L1.

### 5.6.3 Recherche des erreurs de compression

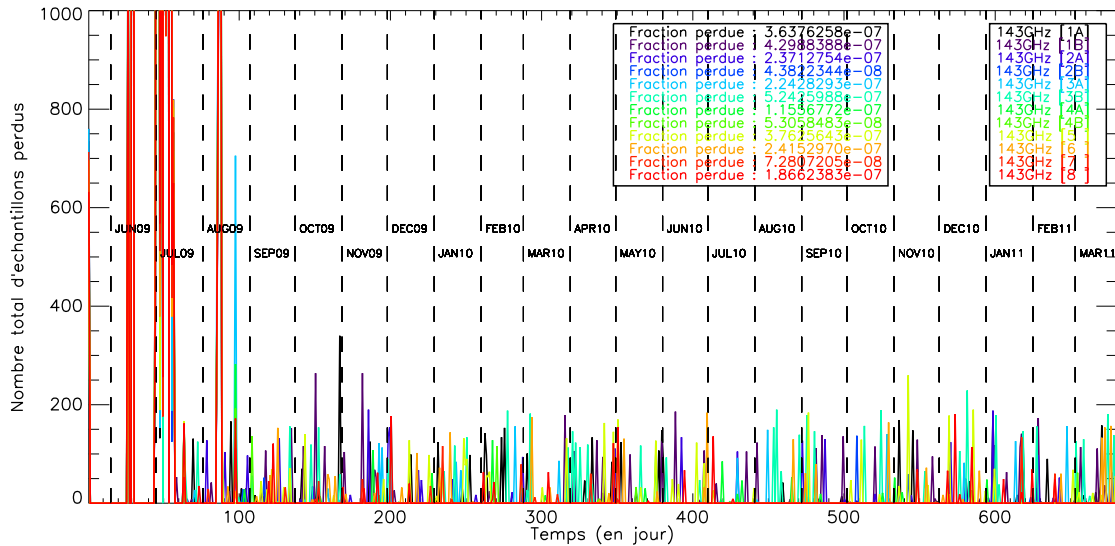
Dans la base de données de PLANCK, il est possible d'associer des marquages particuliers à chaque échantillon. Par exemple, lors du traitement des données, l'impact des rayons cosmiques, ou le passage d'une planète sont repérés, afin de distinguer les échantillons affectés et d'appliquer un traitement adapté. Malheureusement, au début de la mission, le niveau L1 ne construisait aucun marquage particulier pour les échantillons issus d'une erreur de compression. Ils étaient simplement mis à zéros dans la TOI du signal brut puis étaient considérés comme des échantillons valides. Suite à des problèmes rencontrés au cours du traitement des données, à cause des échantillons issus d'erreurs de compression qui n'étaient pas distingués des échantillons valides, il a été décidé de marquer ces échantillons comme non disponibles (NOT WRITTEN dans la base de données). Cette amélioration permet de résoudre le problème du traitement des données qui est prévu pour ne pas prendre en compte les échantillons non valides. Cependant, aucune autre information n'est disponible permettant de distinguer directement les échantillons provenant d'une erreur de compression, des autres types d'erreurs. Nous avons alors utilisé une méthode de détection en deux étapes. Dans un premier temps, nous recherchons les tranches de compression contenant une valeur non nulle pour la fréquence des mots codant pour des erreurs de compression, grâce au groupe contenant les statistiques corrigées des intervalles de quantification décrit dans la section 5.4.6. On utilise pour cela les objets POSCOMPR\_ERROR et NEGCOMPR\_ERROR. Dans un second temps on recherche le signal de la tranche de compression correspondante dans la TOI des données et on trace son évolution au cours du temps, en mettant en évidence les erreurs de compression. Ces dernières correspondent aux échantillons marqués « NOT WRITTEN » et dont la valeur vaut 0.

Plusieurs milliers d'erreurs de compression ont ainsi été détectées depuis le début de la phase d'observation régulière. Il est difficile de déterminer leur origine de façon automatisée car on ne connaît pas forcément toutes les causes qui peuvent produire des erreurs d'une part, mais également, car les erreurs produites par des phénomènes différents laissent une trace similaire. Pour déterminer

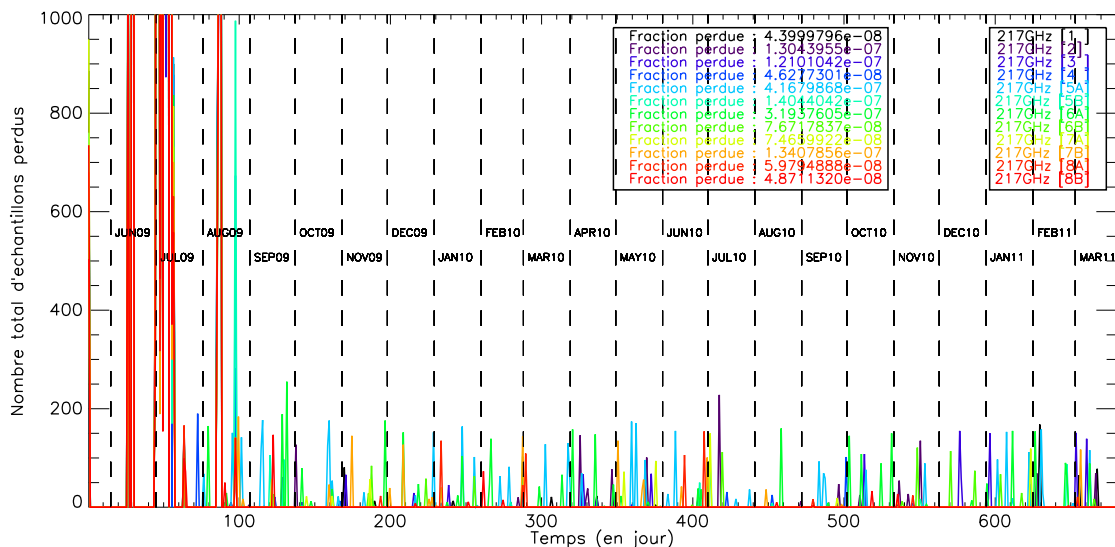




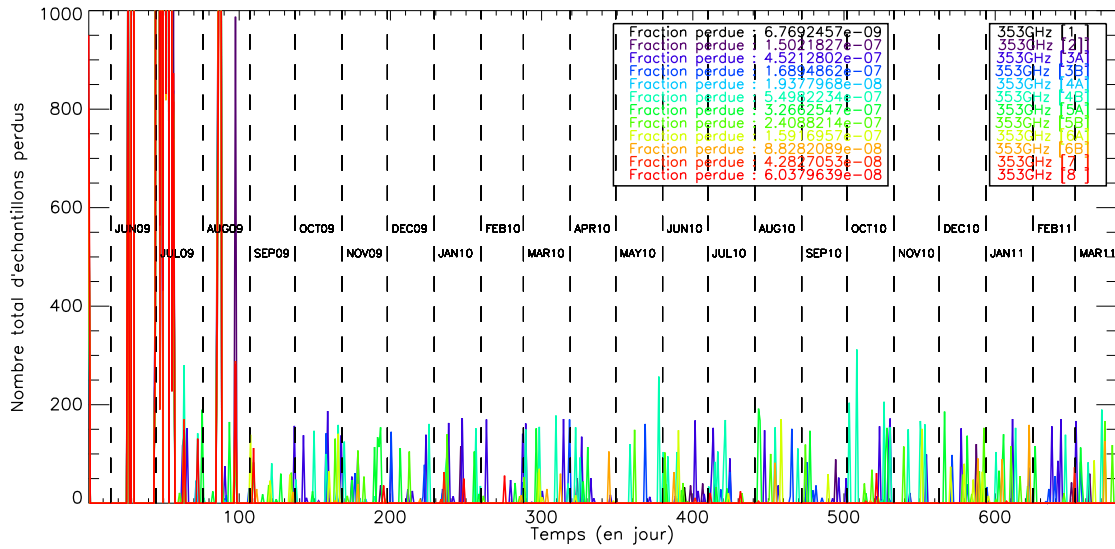
(a) Nombre d'échantillons perdus par jour pour les voies à 100GHz.



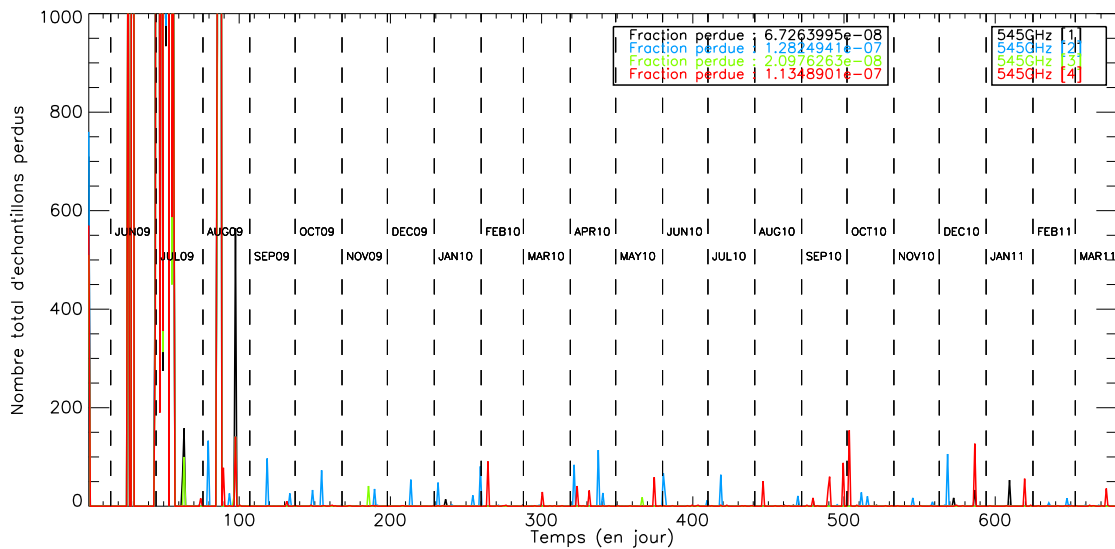
(b) Nombre d'échantillons perdus par jour pour les voies à 143GHz.



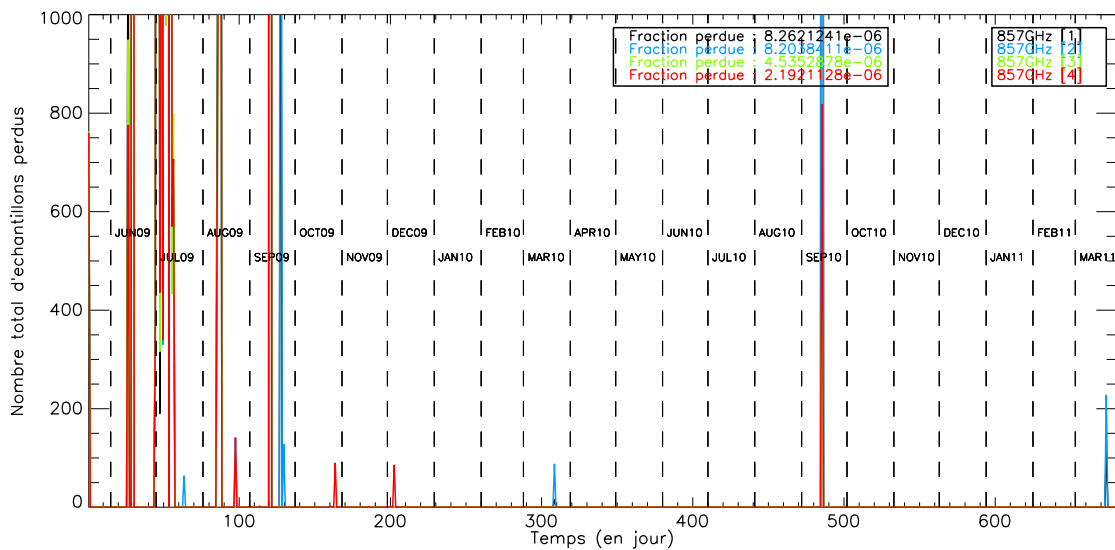
(c) Nombre d'échantillons perdus par jour pour les voies à 217GHz.



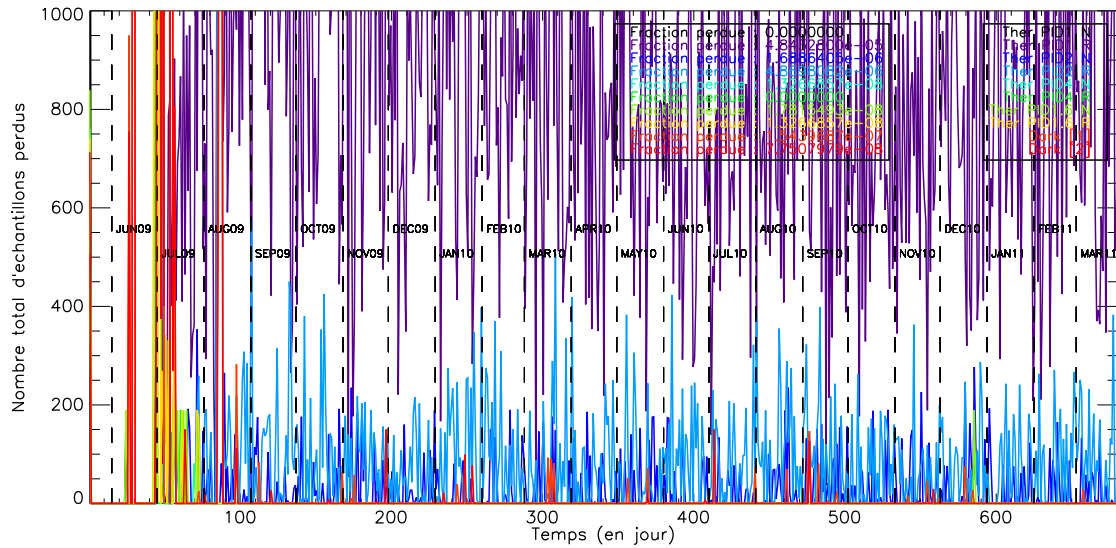
(d) Nombre d'échantillons perdus par jour pour les voies à 353GHz.



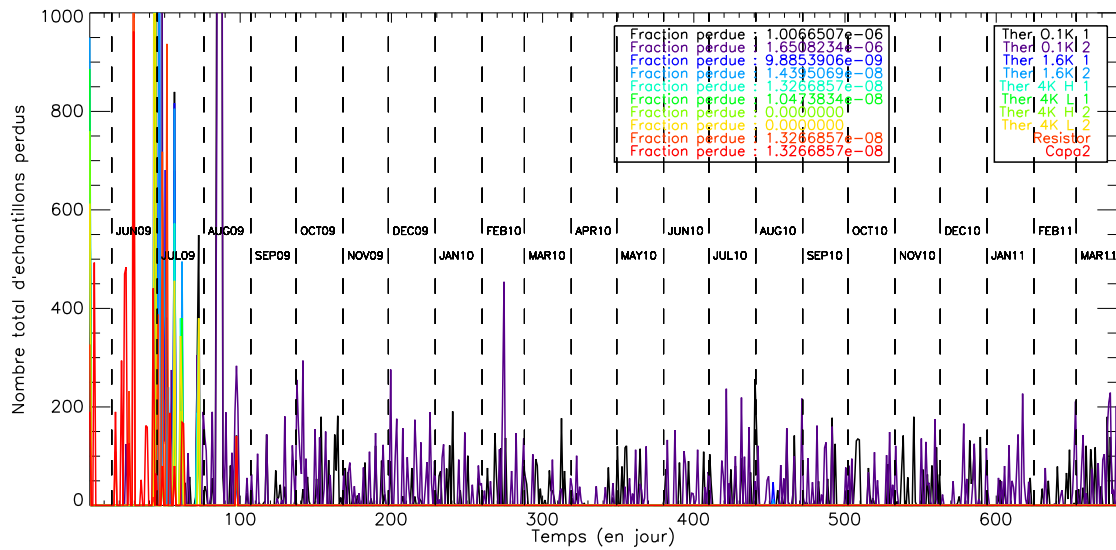
(e) Nombre d'échantillons perdus par jour pour les voies à 545GHz.



(f) Nombre d'échantillons perdus par jour pour les voies à 857GHz.



(g) Nombre d'échantillons perdus par jour pour les thermomètres.



(h) Nombre d'échantillons perdus par jour pour les thermomètres.

FIGURE 5.15 – Évolution de nombre d'échantillons perdus à cause des erreurs de compression. Chaque point est la moyenne sur une journée. Depuis le début des observations en août 2009 le taux d'échantillon perdu est stable. Il est dominé par les erreurs provoquées par les plus forts impacts des rayons cosmiques, dont le taux varie peu.

l'origine des erreurs, depuis le début des observations, une surveillance visuelle a été réalisée pour chaque erreur rencontrée. Toutes les erreurs ont été systématiquement étudiées et leurs causes sont maintenant bien comprises. On peut les classer en trois catégories : les erreurs produites par l'impact des rayons cosmiques, qui sont de loin les plus courantes, celle produites par l'émission galactique et celle produites par l'émission des planètes. Le nombre d'échantillons perdus à cause d'erreurs de compression, par jour et par voies, est résumé sur la figure 5.15. On compte en général, entre zéro et une dizaine de tranches de compression contenant des erreurs de compression et conduisant au maximum à la perte d'environ 200 échantillons par jour. Ces valeurs sont stables depuis le début des observations car elles sont largement dominées par les erreurs produites par l'impact des rayons cosmiques dont le taux varie peu. La fraction de données perdues depuis le début des observations, pour chaque voie, est donnée sur la figure 5.15 et varie entre  $10^{-6}$  et  $10^{-9}$ , ce qui est totalement négligeable pour la qualité des observations.

### Erreurs provoquées par les rayons cosmiques

Parmi les erreurs de compression rencontrées depuis le début des observations, les plus courantes sont dues à des rayons cosmiques de haute énergie. Comme on peut le voir sur la figure 5.16 la variation du signal engendrée par leur impact peut être très importante à l'échelle d'une tranche de compression. Dans les cas les plus extrêmes, l'estimation de la ligne de base pour la démodulation est trop fortement biaisée, ce qui produit un signal oscillant avec une grande amplitude et une estimation de l'offset du signal très éloignée de sa valeur réelle. Un grand nombre d'échantillons se trouve alors éloigné de l'intervalle central et doit donc être codé sur 22bits. Lorsque la limite de 64 est atteinte, l'ensemble des échantillons suivant devant être codé sur 22bits produit une erreur de compression.

La fréquence des rayons cosmiques diminuant lorsque leur énergie augmente, on n'observe heureusement qu'un faible taux d'événements de ce type. En moyenne suivant les voies cela varie entre 0 et 10 tranches de compression par jour, produisant des erreurs de compression. Comme on va le voir dans la suite, seuls deux autres événements ont conduit à des erreurs de compression. Depuis le début des observations, l'impact des rayons cosmiques est donc la source largement dominante d'erreurs de compression. Cependant, leur conséquence sur la qualité des observations est négligeable, de par le peu d'échantillons touchés, mais également grâce à la stratégie de pointage de PLANCK. En effet, comme on l'a vu dans la section 4.3.7, le satellite observe en moyenne 47 fois chaque cercle sur le ciel avant de changer sa direction d'observation. Ainsi si quelques échantillons ont généré des erreurs de compression, les passages redondants assurent dans tous les cas que la zone du ciel a été observée un grand nombre de fois. Par ailleurs, la nature aléatoire de l'impact des rayons cosmiques évite que ce soit systématiquement la même zone du ciel qui soit touchée.

### Erreurs provoquées par l'émission de la poussière galactique

La seconde source d'erreur de compression est la brutale variation de l'intensité de l'émission de la poussière au niveau du centre galactique. L'émission de la poussière augmente fortement avec la fréquence comme on peut le voir par exemple sur les cartes du ciel de PLANCK de la figure 8.3. Dans les voies de l'instrument HFI faisant leurs observations aux plus hautes fréquences (545GHz et 857GHz), le signal est dominé par cette composante, en particulier dans les régions proches du plan galactique. Au niveau du centre galactique, on trouve une zone où l'émission atteint son niveau maximum. Lors du premier balayage de cette zone du ciel par le satellite, entre le 14 et le 15 septembre 2009, on a observé de façon presque systématique, des erreurs de compression sur l'ensemble des voies à 857GHz. Sur la figure 5.17 on peut voir le signal d'une des voies à 857GHz lors du passage dans cette partie du ciel. On constate une brutale variation du signal supérieure à 10000 fois le pas de quantification sur une échelle de temps inférieure à celle d'une tranche de compression. Cela produit donc un effet similaire à celui des rayons cosmiques : une mauvaise estimation de la ligne de base, puis une mauvaise estimation de l'offset, ce qui conduit ensuite à coder un grand nombre d'échantillons sur 22bits et si la limite de 64 est atteinte, produit des erreurs de compression. Dans ce cas particulier, comme la plupart des échantillons sont trop loin de la moyenne, une bonne démodulation et une estimation correcte de la moyenne donnerait quand même trop de mesures sur 16bits. Cet effet n'est donc pas totalement identique à celui d'un rayon cosmique qui n'affecte qu'un plus petit nombre d'échantillons. Au cours de cet événement, un nombre important de tranches de

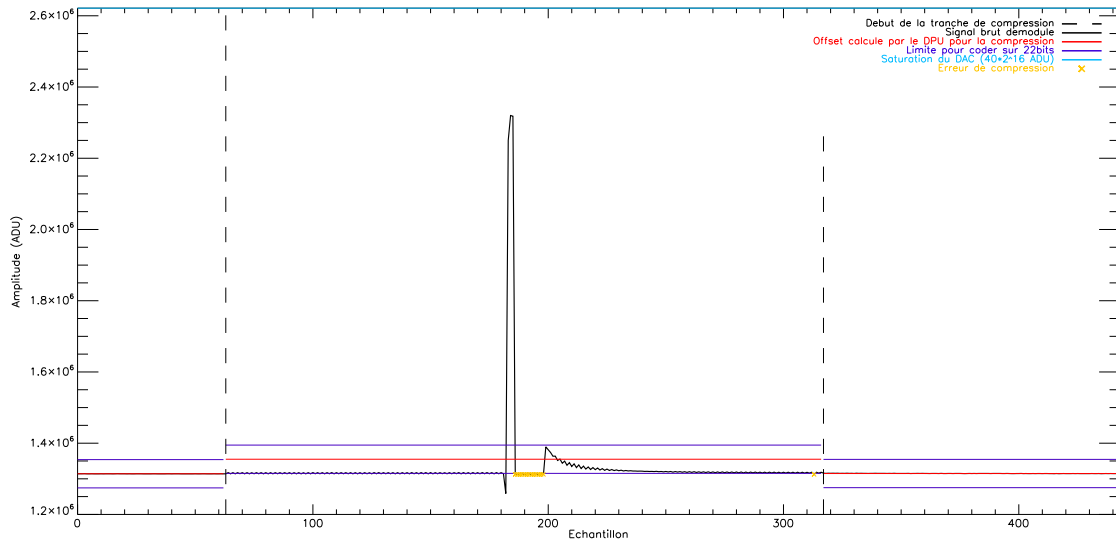


FIGURE 5.16 – **Erreur de compression produite par l'impact d'un rayon cosmique.** Les rayons cosmiques de haute énergie, peuvent déposer localement une grande quantité d'énergie. Cela induit une brutale augmentation de la température dans les bolomètres, suivie d'une décroissance exponentielle caractéristique du bolomètre. Lorsque l'amplitude est trop élevée, cela biaise trop fortement l'estimation de la ligne de base pour la démodulation, puis de l'offset, sur une tranche de compression. Beaucoup d'échantillons vont donc devoir être codés sur 22bits par l'algorithme de compression. Pour éviter d'avoir un débit trop important sur une tranche de compression, une limite de 64 échantillons pouvant être codés sur 22bits a été introduite. Tous les échantillons suivants de la tranche de compression, devant être codés sur 22bits, génèrent alors une erreur de compression, représentés par des croix jaunes sur la figure ci-dessus. Pour donner un ordre d'idée de l'amplitude de l'impact du rayon cosmique la ligne bleue horizontale, représente la limite au-delà de laquelle les échantillons doivent être codés sur 22bits. En rouge on peut voir le biais sur l'offset introduit par l'impact du rayon cosmique considéré. L'amplitude du signal produit par un rayon cosmique peut atteindre des valeurs jusqu'à  $10^5$  fois plus élevées que la normale et arrive même dans certain cas à saturer les détecteurs. Le figure ci-dessus correspond à un impact sur le bolomètre 1B à 100GHz. L'amplitude des échantillons ayant subi une erreur de compression a été choisi arbitrairement à la valeur moyenne afin qu'ils soient visibles sur la courbe.

compression ont produit des erreurs. Cependant grâce à la stratégie de pointage du satellite, qui inclue une forte redondance sur chaque portion du ciel et compte tenu du fait que les tranches de compression ne tombent pas toujours exactement sur la même zone du ciel, les erreurs n'ont pas eu d'effet important sur les données. La seule conséquence est une faible diminution du rapport signal sur bruit de certains pixels qui ont été reconstruits avec quelques échantillons en moins.

### Erreurs provoquées par l'émission des planètes

Finalement, l'observation de la planète Jupiter a également généré des erreurs de compression sur la quatrième voie à 857GHz. La planète Jupiter est le corps dont l'amplitude est la plus élevée aux longueurs d'onde de PLANCK. Étant donnée la résolution de l'instrument HFI, la taille de Jupiter et sa distance du satellite, elle est vue comme un objet quasi-punctuel. La variation de l'amplitude induite par son passage devant les détecteurs est donc très brutale et très brève. Les autres planètes du système solaire situées sur des orbites plus éloignées du Soleil que la Terre sont aussi observées par PLANCK et sont vues comme des objets punctuels. L'observation des planètes joue un rôle essentiel pour l'instrument HFI. Grâce à leur forte amplitude elle peuvent être observées avec un rapport signal sur bruit élevé. Leur caractère punctuel est alors utilisé pour déterminer de façon précise le profil du lobe de l'instrument. Par ailleurs, la brusque variation d'amplitude sur un temps très court permet d'étudier la réponse des bolomètres et en particulier les composantes de cette réponse possédant de longues constantes de temps. Lors de la première observation de Jupiter (voir

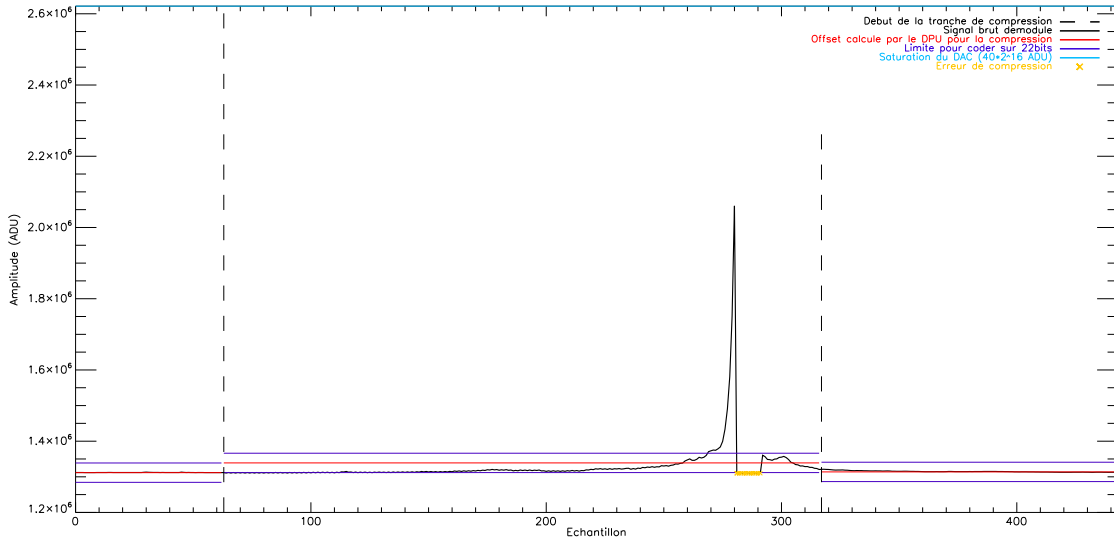


FIGURE 5.17 – **Erreur de compression lors du balayage du centre galactique sur les voies à haute fréquence de l'instrument HFI.** Lors du balayage du centre galactique, la variation de l'émission de la poussière à 857GHz, à l'échelle d'une tranche de compression, a été trop brutale. Cela s'est traduit par un trop grand nombre d'échantillons à coder sur 22bits. La limite de 64 a été atteinte et les échantillons suivants qui auraient dû être codés sur 22bits, ont généré des erreurs de compression (croix jaune dans la figure ci-dessus). L'ensemble des voies à 857GHz a été impacté. Lors de la seconde couverture du ciel les paramètres de compression pour les voies à haute fréquence ont été modifiés et depuis aucune erreur n'a été détectée. L'erreur représenté est une des erreurs du bolomètre 1 à 857GHz. L'amplitude des échantillons ayant subi une erreur de compression a été choisi arbitrairement à la valeur moyenne afin qu'ils soient visibles sur la courbe.

figure 5.18), le 29 octobre 2009 à 00h15 (UT), sur la quatrième voie à 857GHz, l'amplitude du signal était trop élevée et cela a provoqué une saturation de la compression qui a produit une erreur de compression sur chaque cercle du *ring* concerné. Parmi l'ensemble des erreurs de compression rencontrées, celles observées sur Jupiter sont celles qui auraient pu avoir le plus de conséquences. D'un part car elles sont présentes de façon systématique et d'autre part, car l'observation des planètes est à la base de la mesure des profils des lobes vus par les différentes voies et sert également à mesurer la réponse lente des bolomètres. Cependant, les planètes Mars, Uranus et Neptune ont également un rapport signal sur bruit important et sont aussi utilisées pour déterminer le profil du lobe et la réponse des bolomètres. Ces mesures complémentaires limitent donc l'impact des erreurs rencontrées. Par ailleurs, grâce aux changement des paramètres de compression sur les voies à 857GHz (voir section 5.7), les observations suivantes de Jupiter n'ont pas produit d'erreur et elles ont permis de faire les analyses nécessaires. Ces problèmes n'ont donc pas eu de conséquence sur la qualité des observations.

## 5.7 Ajustement des paramètres de compression pour la seconde couverture du ciel

Lors de la première couverture du ciel on a observé que le débit des thermomètres et des bolomètres aux fréquences de 100GHz à 345GHz, était très stable. Un faible nombre d'erreurs de compression, produit par l'impact de rayons cosmiques déposant une trop grande quantité d'énergie, a été mis en évidence sur l'ensemble des voies. Par la nature de leur source, ces erreurs se produisent de façon aléatoire. Elles affectent moins d'un dizaine de tranches de compression par jour et par voie, entraînant la perte d'un faible nombre d'échantillons. L'impact sur les données de ce type d'erreurs est totalement négligeable, grâce à la redondance des observations au niveau des *rings*, aux quatre couvertures du ciel prévues, aux faibles nombre d'échantillons touchés et à leur distribution aléatoire.

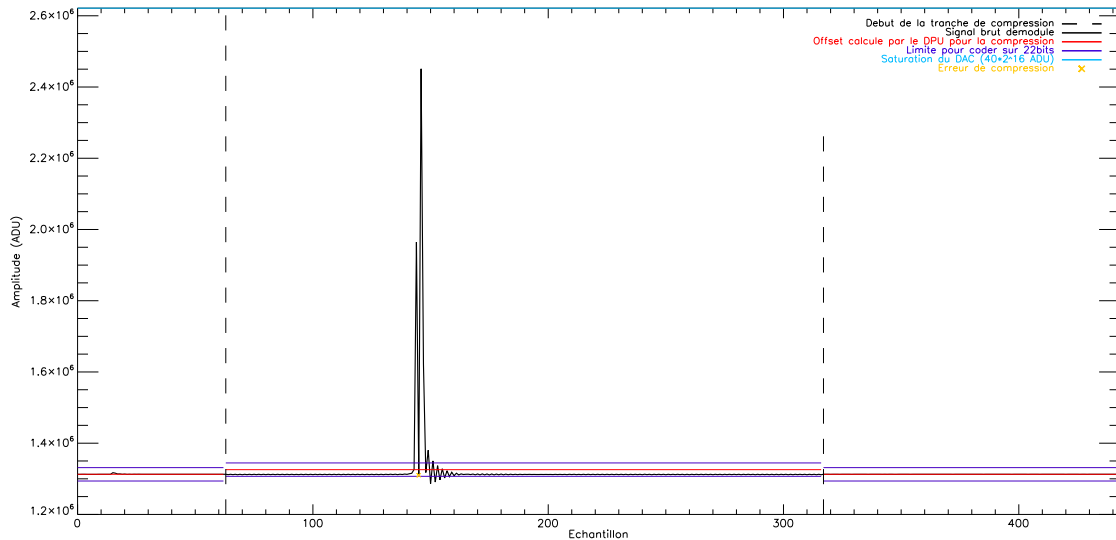


FIGURE 5.18 – **Erreur de compression sur Jupiter pour la quatrième voie à 857GHz.** L’amplitude du signal produit par la première observation de la planète Jupiter sur la quatrième voie d’observation à 857GHz de l’instrument HFI était trop importante et la valeur de son maximum a dépassé la valeur maximale possible pour l’algorithme de compression ce qui a généré une erreur de compression (la croix jaune). L’erreur correspond à l’erreur produite sur le vingtième cercle d’observation de Jupiter pour le bolomètre 4 à 857GHz. L’amplitude des échantillons ayant subi une erreur de compression a été choisi arbitrairement à la valeur moyenne afin qu’ils soient visibles sur la courbe.

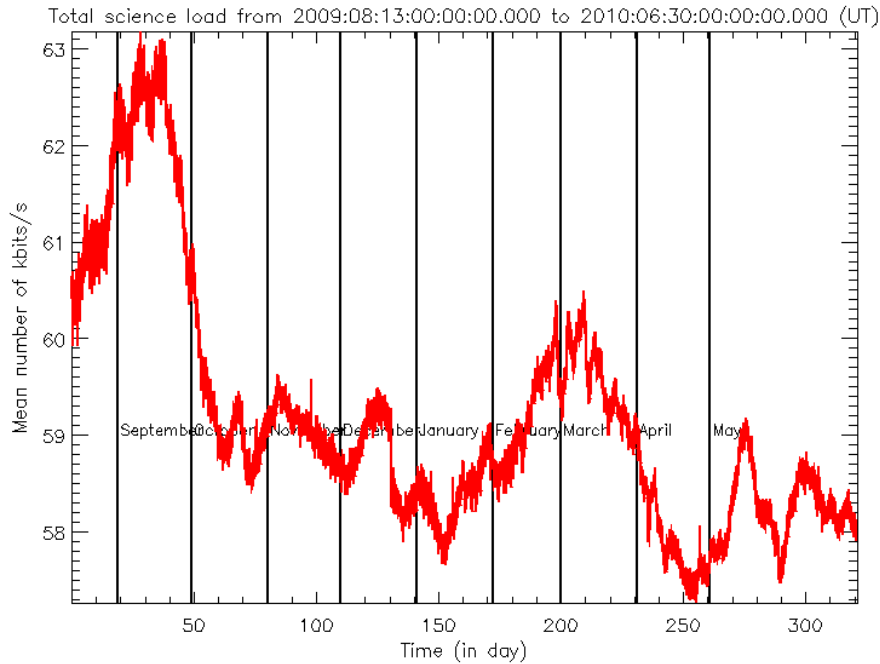
Sur les voies à 545GHz et 857GHz, une progressive augmentation du débit sur plusieurs semaines, atteignant plus de 2kbits sur certaines voies, a été observée (voir les figures 5.11(a) à 5.11(f)). Elle est causée par la forte variation de l’amplitude de l’émission de la poussière, lorsqu’on s’approche du centre galactique, qui modifie la répartition en amplitude des échantillons à l’intérieur d’une tranche de compression. Un plus grand nombre d’échantillons se trouvent loin de la moyenne et l’algorithme de compression devient alors moins efficace. Le maximum a été atteint autour du 12 septembre. La limite de débit imposée au DPU n’a pas été atteinte mais la marge restante a été jugée insuffisante.

En plus de l’augmentation du débit, le passage du centre galactique a produit des erreurs de compression sur l’ensemble des canaux à 857GHz, entre le 12 et le 13 septembre 2009, quand les détecteurs ont croisé la zone où l’émission de la poussière devient la plus intense (voir figure 5.17). Dans cette zone, la brusque variation du signal à l’échelle d’une tranche de compression est trop importante. Un trop grand nombre d’échantillon se retrouvent loin de l’intervalle central de quantification et doivent être codés sur 22bits. La limite du nombre d’échantillons codés sur 22bits a été systématiquement atteinte sur les quatre voies. Encore une fois, les conséquences sur les données ont été négligeables grâce à la redondance des observations au niveau du *ring*. Finalement, une erreur de compression a aussi été observée lors de l’observation de la planète Jupiter par la quatrième voie à 857GHz. Elle a également été causée par la trop forte variation de l’intensité sur la tranche de compression.

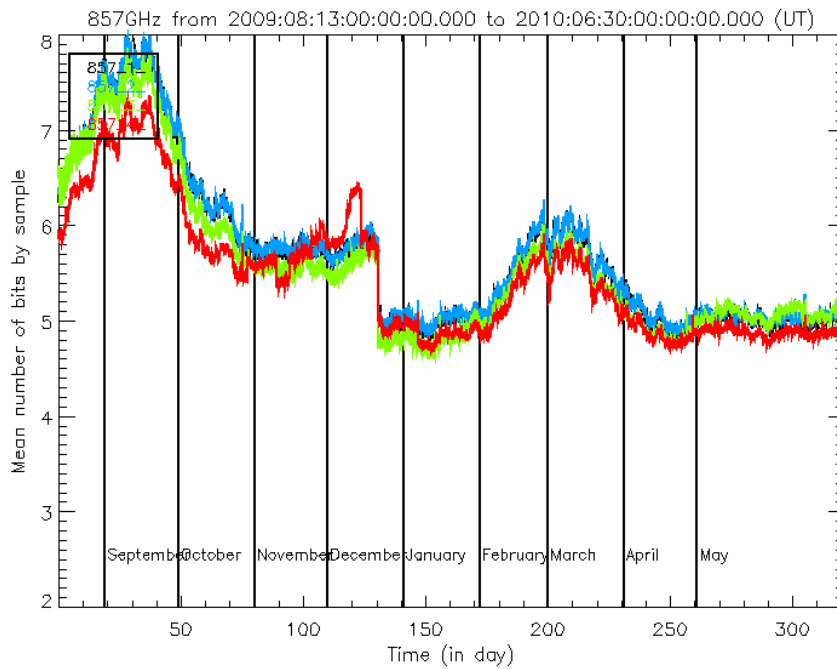
L’ensemble de ces problèmes, rencontrés lors de la première couverture du ciel, a pesé en faveur d’un réajustement des paramètres de compression pour les voies à haute fréquence. Il a été décidé de passer d’une valeur du rapport  $q/\sigma$  de 0.4 à une valeur de 0.5 pour les quatre canaux à 857GHz. Des simulations ont montré que cela permet de diminuer le débit d’environ 2kbits/s ce qui donne une marge suffisante avant d’atteindre le débit maximum. Il est difficile de montrer que ces changements suppriment tous les risques d’obtenir une erreur de compression proche du centre galactique car cela dépend des détails des variations de l’émission, de la direction exacte dans laquelle on croise le maximum et de la position de ce maximum dans la tranche de compression.

Le changement des paramètres a été effectué le 21 décembre 2009. On peut voir sur les figures 5.19(b) et 5.19(a), la diminution du débit sur les voies à haute fréquence et sur le débit total respectivement. Grâce à ce changement, les observations suivantes du centre galactique n’ont pas

conduit à une augmentation du débit aussi élevée. On est ainsi resté en dessous de la limite imposée.



(a) Débit total



(b) Débit des voies à 857GHz

FIGURE 5.19 – **Effet du changement des paramètres de compression sur le débit.** Le 21 décembre 2009, les paramètres de compression des quatre voies à 857GHz ont été ajustés. Le rapport  $q/\sigma$ , initialement fixé à 0.4 a été augmenté à 0.5. Cela permet de diminuer le débit des voies à 857GHz, ce qui est bien visible sur la figure 5.19(b). On peut voir sur la figure 5.19(a) que le débit total a diminué d'environ 2kbits/s, ce qui assure une marge suffisante lors du passage proche du centre galactique pour ne pas atteindre la limite maximale du débit, ce qui déclencherait un EOS. Par ailleurs, ce changement permet également de limiter également le nombre d'erreurs de compression provoquées par la forte variations du signal au passage du centre galactique.



## 5.8 Bilan

j'ai effectué un suivi permanent de la compression depuis le début des observations. Cette analyse régulière a permis de s'assurer du bon fonctionnement de la compression et du respect des contraintes sur le débit de données de l'instrument HFI. Depuis le début des observations, aucune perte de paquets causée par la saturation du tampon en mémoire du DPU n'a été détectée. Le déclenchement d'un EOS s'est produit début août 2009, suite à l'arrêt du compresseur à 4K qui a conduit à un réchauffement global de l'instrument. Enfin, de façon régulière, un peu moins d'une dizaine de tranches de compression contenant des erreurs de compression, par jour et par voie, sont observées depuis le début des observations. Ces erreurs sont produites majoritairement par les plus forts impacts des rayons cosmiques et conduisent à la perte de quelques secondes de données par jours, dans le pire des cas. Au total la fraction de données perdue à cause des erreurs de compression ne dépasse jamais  $10^{-6}$  sur les bolomètres, ce qui est totalement négligeable pour la qualité des observations.

Le suivi des réglages de la compression a montré qu'ils étaient stables pour la plupart des voies, à l'exception de celles affectées par du bruit télégraphique. Pour ces dernières les variations observées ne peuvent pas conduire à une dégradation de la qualité des données. Finalement le débit des données, a toujours été inférieur à la limite maximale. Des variations importantes, sur les voies à haute fréquence, provoquées par les fortes variations de l'émission de la poussière dans le plan galactique (en particulier au niveau du centre galactique), ont conduit à un ajustement des paramètres de compression pour les bolomètres à 857GHz. Ces changements ont permis de s'assurer d'une marge suffisante pour le débit. Elles ont également permis de s'affranchir de la plupart des erreurs de compression observées sur ces canaux, rencontrées au cours de la première couverture du ciel, de façon presque systématique, sur les tranches de compression coïncidant avec l'observation du maximum galactique.

## Troisième partie

# Estimation du spectre de déflexion dans les données de HFI



## Chapitre 6

# Simulation et analyse de signaux pour Planck

Ce chapitre présente les outils et les méthodes utilisés ou développés, pour la génération d'un pipeline d'analyse des cartes du ciel complet, par petites portions (dénommées « patches » dans la suite), dédié à l'étude de l'effet de lentille gravitationnelle, dans les données de PLANCK. Cette étude nécessite de maîtriser à la fois la simulation des effets de lentille gravitationnelle et l'estimation du spectre de puissance de la déflexion qui est le but de l'analyse.

Générer un ensemble de codes de simulation et d'estimation peut parfois prendre une allure de serpent qui se mord la queue. En effet, pour valider un code de simulation un des tests essentiel est de contrôler la justesse du spectre de puissance angulaire. Pour cela il faut avoir à disposition un code d'estimation du spectre. Cependant pour valider le code d'estimation il faut avoir à disposition un ensemble de simulations. Cela à particulièrement été vrai dans le cas des codes par patch. Heureusement, la simulation de cartes gaussiennes et l'estimation de spectres de puissance sont des problèmes déjà bien connus. Dans la mesure du possible nous avons essayé de ne pas réinventer la roue et nous avons réutilisé au maximum les programmes déjà existants qui correspondent principalement aux programmes de simulation et d'estimation sur la sphère.

Dans un premier temps nous montrerons comment on simule les spectres de puissance, les champs gaussiens, l'effet de lentille gravitationnelle et les effets instrumentaux dominants, puis nous détaillerons le processus d'estimation des spectres de puissance de température et de déflexion. Nous présenterons dans chaque cas l'approche dans le plan et sur la sphère.

## 6.1 Simulation des spectres de puissance

### 6.1.1 Paramètres cosmologiques

Dans notre étude, nous avons besoin de générer des simulations des cartes du CMB et du potentiel gravitationnel intégré afin de pouvoir faire une série de tests sur la méthode d'analyse que nous avons développé. Le spectre de puissance de ces champs dépend du modèle d'Univers choisi et de la valeur des paramètres cosmologiques qui le contraignent. Les observations actuelles conjointes du CMB, du fond diffus infrarouge, des supernovae, des oscillations baryoniques... de par leur qualité, permettent de donner de fortes contraintes sur les modèles d'Univers. Un certain nombre de paramètres sont nécessaires pour décrire les modèles. Leurs valeurs sont ajustées pour refléter au mieux les observations. Suivant le modèle choisi et la valeur des paramètres, on obtient des spectres de puissance différents pour les anisotropies de température et de polarisation du CMB que l'on confronte alors aux observations afin d'en déduire la meilleure estimation des paramètres.

Il a été montré que les observations actuelles sont bien décrites par un modèle d'Univers plat de type  $\Lambda$ CDM avec des inhomogénéités issues de fluctuations adiabatiques gaussiennes et des modes tensoriels négligeables. Nous cherchons à générer des simulations réalistes qui reproduisent bien les observations actuelles. Pour cela nous nous basons sur l'estimation des paramètres cosmologiques obtenue par l'équipe de WMAP, intégrant 7 ans de données Larson et al. (2010).

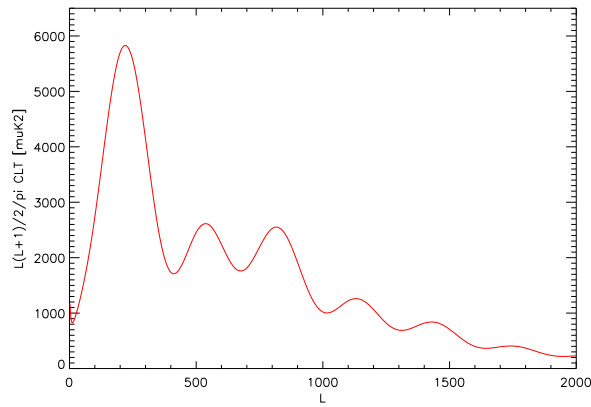
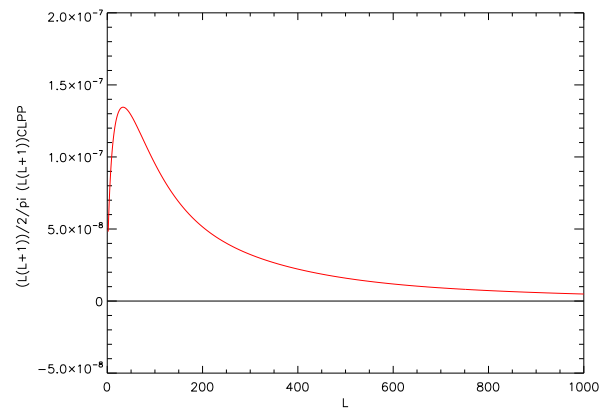

 (a)  $Cl^{TT}$ 

 (b)  $Cl^{dd}$ 

FIGURE 6.1 – **Spectres de puissance angulaire théoriques des anisotropies de température et de la déflexion.** Le spectre de puissance des anisotropies de température du CMB, visible sur la figure (a), est calculé avec CAMB dans un modèle  $\Lambda$ CDM à partir des paramètres cosmologiques déterminés par l'équipe de WMAP intégrant sept années d'observations. La figure (b) montre quand à elle, le spectre de puissance de la déflexion correspondant.

### 6.1.2 Simulation avec CAMB

Pour calculer les spectres de puissance des anisotropies de température et de polarisation du CMB, il faut faire évoluer linéairement les perturbations primordiales dans un modèle d'Univers et en déduire la répartition statistique de la densité lors de l'époque de recombinaison, ce qui permet finalement d'estimer les spectres de puissance angulaire du CMB. Au cours de leur parcours, depuis la surface de dernière diffusion jusqu'au détecteur, les photons du CMB subissent des perturbations, notamment la déflexion par effet de lentille gravitationnelle produit par les grandes structures. Pour tenir compte de l'impact de cet effet sur les spectres de puissance, il faut finalement propager les photons de la surface de dernière diffusion jusqu'au détecteur.

Plusieurs codes ont été développés afin d'estimer le spectre de puissance des anisotropies et de la polarisation du CMB à partir des paramètres cosmologiques. Nous utilisons le programme CAMB (Code for Anisotropies in the Microwave Background), développé par Antony Lewis et Anthony Challinor, dont les sources sont disponibles sur le site <http://www.camb.info>. Il permet d'estimer les spectres de puissance des anisotropies de température et de la polarisation du CMB, avec une précision de l'ordre de 0.3%, mais également d'estimer les spectres de puissance lentillés, le spectre de puissance du potentiel gravitationnel intégré ainsi que leurs spectres croisés, dans une large gamme de modèles d'Univers. Nous avons utilisé la version de février 2009. Une nouvelle version a été publiée depuis, en janvier 2010 mais les améliorations apportées n'ont pas d'influence sur notre

étude.

Pour simuler les spectres, nous prenons en compte les corrections non linéaires dans CAMB. Nous n'incluons pas de modes tensoriels. Sur la figure 6.1 on peut voir les spectres des anisotropies de température et de la déflexion obtenus.

## 6.2 Simulation de cartes gaussiennes

Le CMB est un champ gaussien isotrope (voir section 2). Ses fluctuations sont donc entièrement décrites par leur spectre de puissance. La méthode de simulation des cartes repose sur cette hypothèse et sur les propriétés d'un champ gaussien. Elle se base sur un tirage aléatoire contraint.

### 6.2.1 Simulation sur la sphère

#### Pixelisation de la sphère avec HEALPix

La représentation d'un signal sur tout le ciel, nécessite de définir une pixelisation de la sphère. Les premières cartes du CMB du ciel complet, produites par COBE utilisaient une pixelisation carrée d'un ensemble de six plans, issus de la projection de la sphère sur les faces d'un cube. Depuis, de nombreuses méthodes plus adaptées ont été développées pour pixeliser la sphère, dans le but d'être utilisées par les futures expériences de mesure du CMB, notamment, HEALPIX, GLESP, Igloo, Icosahedron, ECP. La pixelisation HEALPIX tend à s'imposer dans un nombre croissant d'expériences, ces dernières années. Son succès est certainement dû au développement d'un ensemble performant de programmes et bibliothèques, permettant de travailler avec cette pixelisation, librement accessibles et disponibles dans plusieurs langages (C, C++, FORTRAN, IDL, python).

Les principales propriétés de cette pixelisation sont : une indexation hiérarchique des pixels qui facilite l'accès rapide aux pixels ou la recherche de plus proches voisins, une surface identique pour tous les pixels qui permet d'avoir un bruit constant par pixel pour une observation homogène et enfin une répartition des centres des pixels sur des cercles d'isolatitude qui permet de réaliser des transformations en harmonique sphérique basées sur des FFT, ce qui allège grandement les calculs. Les pixels HEALPIX sont issus de la subdivision d'un ensemble de 12 pixels de base (voir figure 6.2). Le nombre de subdivisions est contraint par le paramètre *nside*. Ainsi, une carte possède au total  $12nside^2$  pixels.

Les cartes du ciel produites par le consortium PLANCK utilisent la pixelisation HEALPIX<sup>1</sup>. Étant donnée la résolution des instruments de PLANCK de quelques minutes d'arc, il a été décidé de produire l'ensemble des cartes avec un paramètre *nside* de 1024 pour les cartes de LFI et un *nside* de 2048 pour les cartes de HFI. Un simple calcul montre que les cartes de HFI comportent plus de 50 millions de pixels, avec une taille angulaire moyenne entre leurs centres de 1.7 minutes d'arc. Pour avoir une résolution comparable à celle des données de PLANCK, nous avons donc produit l'ensemble des simulations avec un paramètre *nside* de 2048.

#### Fonction fenêtre des pixels

Le signal observé sur le ciel est un signal continu. Cependant, les observations sont représentées sur des cartes pixelisées et une valeur unique est affectée à chaque pixel. On peut écrire l'effet de la pixelisation par la formule suivante (Hivon et al. (2002)) :

$$T_p = \int W^p(\theta, \phi) T(\theta, \phi) d\Omega \quad (6.1)$$

où,  $T_p$  et  $W^p(\theta, \phi)$  sont respectivement, la valeur affectée au pixel  $p$  et sa fonction fenêtre.  $T(\theta, \phi)$  représente la valeur du signal observé, dans une direction sur la sphère, définie par  $(\theta, \phi)$ . La plupart des algorithmes de génération des cartes à partir des observations sur le ciel, effectuent une moyenne des échantillons tombant dans chaque pixel. Dans ce cas, on peut définir la fonction fenêtre des pixels par :

$$W^p(\theta, \phi) = \begin{cases} \frac{1}{A_p} & \text{à l'intérieur du pixel} \\ 0 & \text{sinon} \end{cases} \quad (6.2)$$

1. Voir le site web <http://healpix.jpl.nasa.gov/>

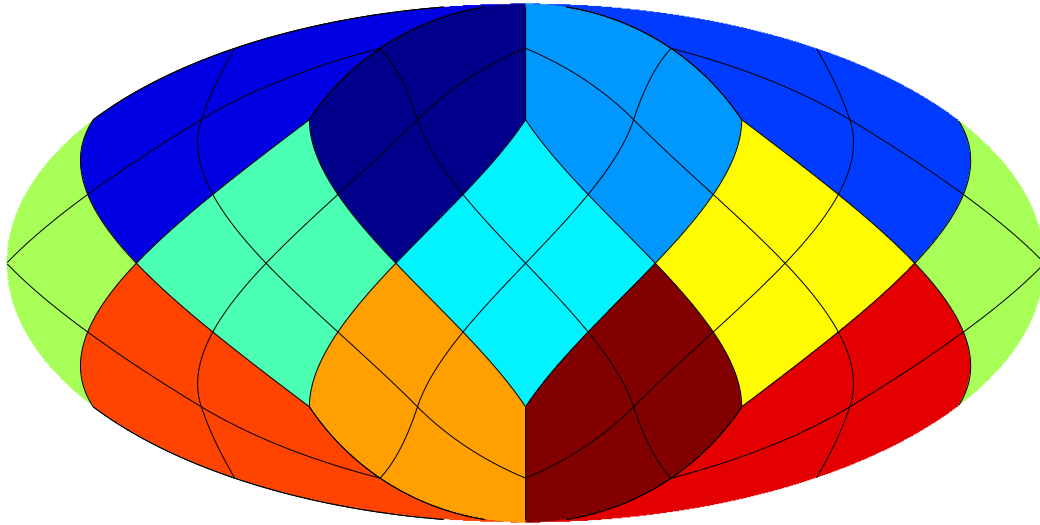


FIGURE 6.2 – **Pixelisation HEALPix**. On obtient une carte de résolution arbitraire en subdivisant chaque pixel en quatre sous-pixels à partir d’un ensemble de douze pixels de base. La résolution est contrôlée par le paramètre  $n_{side}$  qui définit le nombre de subdivisions par rapport aux pixels de base. Chaque carte comporte ainsi  $12n_{side}^2$  pixels. La figure ci-dessus montre la projection de Mollweide des 12 pixels de base (visible en couleur). On peut également voir en noir, les contours des pixels issus de la première subdivision ( $n_{side} = 2$ )

où  $A_p$  est la surface du pixel  $p$ , qui est constante pour la pixelisation HEALPIX ( $A_p = A$ ). La pixelisation lisse toutes fluctuations du signal à des échelles plus petites que la plus large dimension d’un pixel (Crittenden & Turok (1998)). Cela se traduit par une perte de puissance à grand  $\ell$  dans le spectre de puissance angulaire du signal. C’est un effet systématique purement géométrique et qui a été estimé avec précision pour les pixels HEALPIX, pour des cartes de paramètre  $n_{side}$  pouvant aller jusqu’à 8192.

### Simulation

Dans la section 2 on a vu que dans l’état actuel des connaissances, le CMB peut être considéré dans une bonne approximation comme un champ gaussien statistiquement isotrope. Ses propriétés statistiques sont donc totalement décrites par son spectre de puissance angulaire. Si on décompose une carte sur la base des harmoniques sphériques, les coefficients  $a_{\ell m}$  de la décomposition, pour un  $\ell$  donné, sont des variables aléatoires gaussiennes indépendantes :

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

où  $C_\ell$  est le spectre de puissance angulaire de la carte.

Pour simuler une carte gaussienne il faut donc effectuer un tirage d’ $a_{\ell m}$  gaussien de variance  $C_\ell$  pour un  $\ell$  donné. Pour prendre en compte l’effet de la fonction fenêtre des pixels, il faut moduler le tirage par la fonction correspondante qui dépend de la géométrie des pixels comme décrit dans la section 6.2.1.

Plusieurs codes ont été développés pour générer des cartes du ciel avec une pixelisation HEALPIX à partir d’un spectre de puissance. Dans notre analyse, nous avons utilisé en particulier la version c++ des bibliothèques HEALPIX ainsi que le code LENSPIX<sup>2</sup> qui permet en plus de simuler l’effet de lentille gravitationnelle (voir section 6.3.1) mais qui utilise les routines fortran de HEALPIX pour simuler les cartes gaussiennes. Seul le code simfast permet de prendre en compte l’effet de la fonction

2. Voir le site web <http://cosmologist.info/lenspix>

fenêtre des pixels. Une exemple de carte du ciel du CMB, simulée avec la méthode précédente est visible sur la figure 6.3.

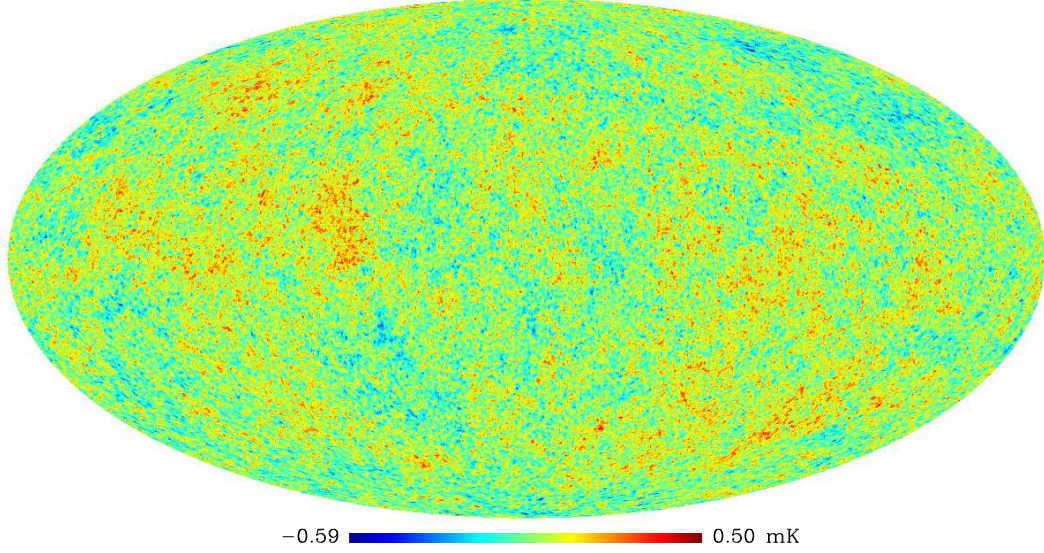


FIGURE 6.3 – **Carte du ciel gaussienne.** La simulation est effectuée avec LENSPIX à partir du spectre de puissance du CMB issu de CAMB et basé sur l'estimation des paramètres cosmologiques publiée par WMAP.

## 6.2.2 Simulation sur le plan

Le principe de la simulation d'un champ gaussien dans un plan, est similaire à celui sur la sphère. Dans le plan, l'espace de Fourier joue un rôle analogue à celui des harmoniques sphériques sur la sphère (voir section 2.3.3).

Pour simuler une réalisation d'un champ gaussien de spectre de puissance  $C(\ell)$ , de taille angulaire  $L_x$  par  $L_y$ , nous commençons par générer une carte  $G_{ij}$  à partir d'un tirage de bruit blanc gaussien de variance 1 et de moyenne nulle. Cette carte vérifie donc la relation :

$$\langle G_{ij} G_{i'j'} \rangle = \delta_{ii'} \delta_{jj'} \quad (6.3)$$

Il suffit ensuite de calculer sa transformée de Fourier et de moduler ses coefficients  $a(\mathbf{k}_{mn})$  de la façon suivante :

$$\hat{a}(\mathbf{k}_{mn}) = a(\mathbf{k}_{mn}) \sqrt{C(|\mathbf{k}_{mn}|)} \quad (6.4)$$

avec le vecteur d'onde  $\mathbf{k}_{mn}$  donné par :

$$\mathbf{k}_{mn} = \begin{pmatrix} \frac{2m\pi}{L_x} \\ \frac{2n\pi}{L_y} \end{pmatrix} \quad (6.5)$$

Les coefficients de Fourier du signal dans le plan, s'obtiennent donc à partir du spectre de puissance angulaire sur la sphère  $C_\ell$ . Pour les valeurs entières de  $|\mathbf{k}_{mn}|$ , on a la relation (voir White et al. (1999)) :

$$C(\ell) = C(|\mathbf{k}_{mn}| = \ell) \quad (6.6)$$

Pour les autres valeurs, il est nécessaire d'interpoler le spectre de puissance angulaire  $C(\ell)$  en  $|\mathbf{k}_{mn}|$ . Pour cela nous réalisons une simple interpolation linéaire qui donne une bonne approximation car le spectre de puissance du CMB est suffisamment lisse. On peut voir un exemple sur la figure 6.4.



## 6.3 Simulation de l'effet de lentille gravitationnelle

Au cours de leur trajet depuis la surface de dernière diffusion jusqu'au détecteur, les photons du CMB sont déviés par l'effet de lentille gravitationnelle produit par les grandes structures (voir figure 3.2). Le signal du CMB mesuré dans une direction correspond donc au signal provenant d'une position légèrement différente au niveau de la surface de dernière diffusion (équation 3.12). Il n'est ainsi jamais possible d'observer directement le signal du CMB non lentillé tel qu'il apparaît lors du découplage. Le déplacement total des photons détectés, dépend du déplacement produit par le potentiel gravitationnel intégré sur l'ensemble du parcours (voir équation 3.3). Simuler l'effet de lentille gravitationnelle nécessite alors d'estimer le signal du CMB en chaque point du ciel, après la déviations des photons par les grandes structures. La technique consiste à modéliser la déflexion des photons et à en déduire le signal lentillé, à partir d'une carte du CMB non lentillée, sensée représenter le signal au niveau de la surface de dernière diffusion.

Pour estimer la déflexion dans chaque direction, il est possible de procéder par lancé de rayon. C'est une technique qui consiste à suivre la trajectoire complète des rayons lumineux tombant dans chaque pixel de la carte et à modéliser l'ensemble des interactions qu'ils ont subi, afin d'en déduire la carte lentillée. C'est une méthode très lourde car elle nécessite de simuler l'ensemble des structures gravitationnelles de l'espace ainsi que leur évolution dans le temps, jusqu'à des échelles suffisamment fines, ce qui nécessite des simulations N-corps précises, très gourmandes en temps de calcul. Cette technique présente toutefois des avantages pour certaines études. Elle permet notamment de prendre en compte les effets non linéaires présents à petite échelle, ou encore les corrélations avec les amas SZ. Elle a récemment été employée pour réaliser des simulations précises pour l'expérience ACT Das & Bode (2008).

Il est possible de s'affranchir de la simulation des structures gravitationnelles. Il existe en effet une méthode alternative pour déterminer les déplacements en chaque point, qui ne nécessite pas de simulations N-corps, ce qui simplifie grandement la procédure. D'après l'équation 3.7, on peut écrire le champ de déflexion comme le gradient d'un champ de potentiel  $\phi$ , appelé champ de potentiel de lentille. C'est l'intégrale sur la ligne de visée du champ gravitationnel à trois dimensions qui est un champ gaussien dont le spectre de puissance peut être calculé avec CAMB pour un modèle d'univers donné (voir section 6.1.2). L'estimation du champ de déflexion nécessite alors simplement de générer une réalisation du champ gaussien  $\phi$  et de calculer son gradient. C'est cette méthode que nous avons utilisé.

On pourrait être tenté d'utiliser le développement limité donné par l'équation 3.15 et générer la carte lentillée à l'aide d'un produit du gradient de température par le gradient du potentiel. Cependant, cette approche ne conduit pas à des résultats suffisamment précis. La convergence de l'expression approchée est en effet très lente Hanson et al. (2010a).

### 6.3.1 Simulation sur la sphère

Pour simuler l'effet de lentille gravitationnelle sur la sphère nous utilisons le code LENSPIX, développé par Antony Lewis, qui est librement utilisable et disponible sur internet. On trouvera des détails sur son fonctionnement dans Lewis (2005) et Hamimeche & Lewis (2008). Il permet de simuler des cartes lentillées de température et de polarisation et possède de bonnes performances. Grâce à une architecture parallélisée, il est capable de générer des cartes HEALPIX en haute résolution (*nside* égal à 2048) en seulement quelques minutes. Pour obtenir ces performances, l'estimation du champ de déflexion est effectuée à partir du gradient du potentiel de lentille comme expliqué ci-dessus et les valeurs lentillées sont obtenues grâce à une interpolation du champ non lentillé. Récemment, une méthode de simulation de cartes lentillées a été développée dans Lavaux & Wandelt (2010). Elle utilise une méthode alternative d'interpolation : FLINTS (pour Fast and Lean Interpolation on the Sphere), qui est une méthode plus précise et rapide, qui tient compte des propriétés statistiques de la carte. Nous avons toutefois utilisé LENSPIX pour l'ensemble de nos simulations car le code FLINTS n'existait pas auparavant mais également car LENSPIX répondait déjà aux besoins de PLANCK.

La première étape de la procédure consiste à simuler des réalisations gaussiennes du CMB non lentillées et du potentiel gravitationnel intégré, à partir de leur spectre de puissance. Cette étape utilise les bibliothèques HEALPIX. La méthode utilisée est donc celle décrite dans la section 6.2.1. Toutefois, les effets de la fonction fenêtre des pixels (voir section 6.2.1) ne sont pas pris en compte. Il ne doivent de toute façon pas être intégrés car ils sont dus à la méthode de production des cartes

et ne correspondent pas à un effet physique. De manière plus générale, les cartes du CMB non lentillées ne doivent contenir aucun effet instrumental.

La seconde étape consiste à calculer le gradient  $\nabla\phi$  du potentiel de lentille  $\phi$  afin d'en déduire la carte de la déflexion  $\mathbf{d}$  (voir l'équation 3.7). Il faut faire particulièrement attention à ce calcul pour des simulations sur la sphère. En effet, les coordonnées du vecteur de déflexion ne sont pas données simplement par des dérivées spatiales (comme sur le plan) mais par des dérivées covariantes sur la sphère. Ainsi, le gradient doit être évalué sur la sphère le long des géodésiques. Les calculs sont effectués dans l'espace des harmoniques sphériques, ce qui simplifie grandement les équations. Les détails peuvent être trouvés dans Lewis (2005).

Connaissant la valeur de la déflexion en chaque direction, on peut finalement générer les cartes lentillées. Pour chaque pixel de la carte non lentillée pointant dans une direction  $\hat{\mathbf{n}}$ , il faut alors évaluer le signal non lentillé dans la direction  $\hat{\mathbf{n}}'$  déterminée à partir de la valeur de la déflexion dans cette direction en suivant l'équation 3.12. Sur la sphère la direction  $\hat{\mathbf{n}}'$  est estimée en se déplaçant d'une distance  $|\nabla\phi|$  sur une géodésique dans la direction donnée par  $\nabla\phi$ . Il faut pour cela faire l'approximation que le gradient reste constant sur cette distance, ce qui est compatible avec l'hypothèse de Born faite pour le calcul de l'intégrale du champ gravitationnel (voir section 3) et qui correspond à une bonne approximation étant donnée l'amplitude des déflexions. À partir de la décomposition en harmoniques sphériques  $a_\ell^m$  du champ non lentillé il est facile d'estimer sa valeur dans une direction arbitraire  $\hat{\mathbf{n}}'$  en faisant une transformée inverse :

$$T(\hat{\mathbf{n}}') = T(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{m=\ell} a_\ell^m Y_\ell^m(\theta, \varphi) \quad (6.7)$$

où les  $Y_\ell^m$  sont les fonctions de base des harmoniques sphériques. Cette technique permet d'estimer la valeur exacte du signal, cependant, lorsque la direction ne coïncide pas avec le centre d'un pixel de la carte, il n'existe pas d'algorithme rapide pour effectuer une transformation en harmonique sphérique inverse. Le temps de calcul devient alors rapidement trop important même pour des cartes possédant une résolution bien inférieure à celle de PLANCK.

Le code LENSPIX propose alors le choix d'estimer la valeur du signal grâce à une interpolation. Ce n'est plus une méthode exacte mais elle permet toutefois d'atteindre une précision de plus de 0.1% sur les spectres de puissance jusqu'à  $\ell$  égal à 2000, ce qui est largement suffisant pour les besoins de PLANCK. La technique employée est basée sur une projection cylindrique équidistante d'une carte suréchantillonnée, suivie d'une interpolation cubique. C'est cette option que nous avons utilisé pour simuler nos cartes car c'est la seule envisageable étant donné la résolution requise et le nombre important de réalisations dont nous avons besoin.

### 6.3.2 Simulation sur le plan

Pour réaliser des simulations de l'effet de lentille gravitationnelle sur le plan, nous avons utilisé une méthode similaire à la méthode précédente sur la sphère. La première étape de la procédure consiste donc à déterminer le champ de déflexion. Pour cela on génère tout d'abord une carte du potentiel de lentille  $\phi$ . C'est un champ gaussien dont le spectre de puissance est connu et qui peut être calculé par CAMB, par exemple, pour un jeu de paramètres cosmologiques donné. On peut donc employer l'outil de génération de cartes gaussienne décrit dans la section 6.2 pour simuler une réalisation. Le champ de déflexion est alors déduit de l'équation 3.12. La détermination du champ de déflexion nécessite alors d'estimer le gradient du potentiel de lentille (voir équation 3.7). Le calcul est effectué dans l'espace de Fourier, en utilisant la propriété de la transformée de Fourier de la dérivée d'une fonction :

$$\mathbf{d}(\mathbf{n}) = -\nabla\phi(\mathbf{n}) = \begin{pmatrix} d_x \\ d_y \end{pmatrix} = \begin{pmatrix} -\text{TF}^{-1} \left[ ik_x \hat{\phi} \right] \\ -\text{TF}^{-1} \left[ ik_y \hat{\phi} \right] \end{pmatrix} \quad (6.8)$$

où,  $\hat{\phi}$  représente la transformée de Fourier du champ de potentielle de lentille. On obtient finalement une carte à deux dimensions, dont chaque pixel donne la valeur du vecteur de déflexion dans la direction donnée par le centre du pixel, qui permet de simuler l'effet de lentille gravitationnelle sur une carte de température de CMB non lentillées.

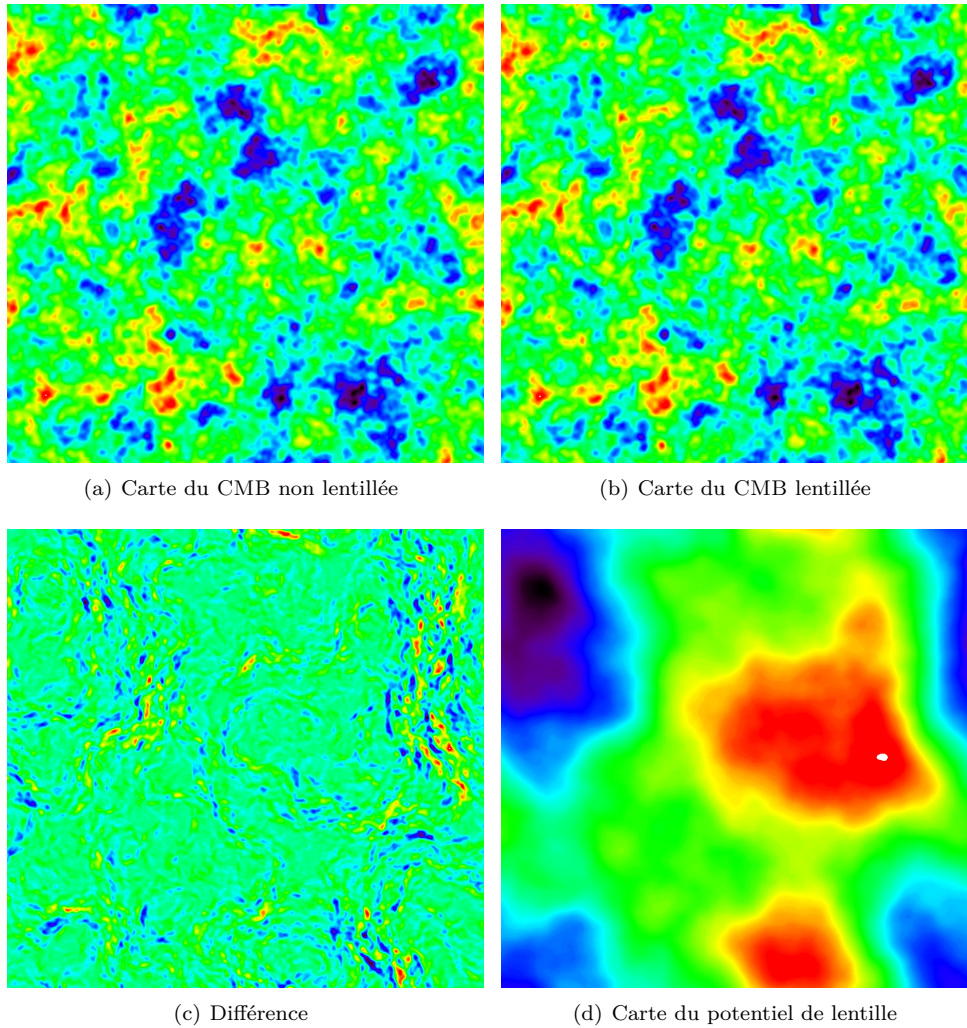


FIGURE 6.4 – **Simulation d'une carte du CMB lentillée dans le plan.** La carte lentillée (figure (b)) est générée à partir des simulations gaussienne des cartes du CMB (figure (a)) et du champ de potentiel de lentille (figure (d)). On remarque sur la figure (c) que l'intensité des différences entre les cartes du CMB lentillées et non lentillées est la plus intense dans les régions où le gradient de la carte du champ de potentiel de lentille est le plus marqué. Chaque carte correspond à une zone de 10 degrés de côté.

La seconde étape se base sur une réalisation d'une carte des anisotropies du CMB idéale. C'est un champ gaussien qui peut être simulé de la même façon que le champ  $\phi$  à partir de son spectre de puissance. Pour obtenir la version lentillée de la carte des anisotropies, il faut alors appliquer l'équation 3.12 en utilisant le champ de déflexion calculé précédemment. Pour simplifier la procédure, on fait appel au principe du retour inverse de la lumière, qui permet d'évaluer le trajet des photons dans le sens opposé à leur propagation. On part donc du centre de chaque pixel de la carte lentillée et on détermine le vecteur de déplacement dans le pixel correspondant de la carte de déflexion, calculée à l'étape précédente. On en déduit finalement la position de laquelle proviennent les photons sur la carte non lentillée (qui correspond à la surface de dernière diffusion).

La dernière étape est la partie la plus délicate de la procédure. En effet, comme le champ de déflexion est un champ aléatoire, l'amplitude et la direction du vecteur de déflexion prennent des valeurs avec une dispersion aléatoire qui dépend de la réalisation du champ considéré. Les photons proviennent donc, en général, d'une position qui ne coïncide pas avec les pixels de la carte non lentillée. Pour estimer l'amplitude d'un pixel de la carte lentillée, il faut donc interpoler le signal non lentillé à la position correcte. Pour éviter d'avoir à interpoler aussi le champ de déflexion, on génère des cartes de déflexion avec la même résolution que celle choisie pour les cartes lentillées, ce

qui permet de connaître directement la valeur de la déflexion pour chaque pixel des cartes lentillées en prenant simplement la valeur de la carte de déflexion dans le pixel correspondant. Ce n'est pas une contrainte très forte et on peut envisager de généraliser la procédure à des cartes de déflexion de résolution différente.

On peut montrer que l'amplitude moyenne des déplacements est de l'ordre de 2 minutes d'arc (voir section 3.4.2). Elle est donc inférieure à la résolution des instruments de PLANCK qui dépend des voies et qui est limitée par la taille du lobe de l'instrument, qui vaut de l'ordre de 5 minutes d'arc dans le meilleur des cas. Les cartes officielles de l'instrument HFI sont réalisées avec un paramètre *nside* de 2048 qui donne un écart moyen entre pixel de 1.7 minutes d'arc. L'effet de lentille gravitationnelle entraîne donc un déplacement du même ordre de grandeur que la distance inter-pixel, ce qui peut poser des problèmes pour l'interpolation. Un moyen simple d'améliorer la précision de l'interpolation consiste donc à générer des cartes lentillées avec une résolution suffisamment élevée pour que la distance inter-pixel soit suffisamment petite comparée à la déflexion moyenne. Par ailleurs, la qualité de l'algorithme d'interpolation joue un rôle important dans la précision des résultats.

Afin de comparer les effets de l'interpolation nous avons implémenté plusieurs méthodes ayant chacune des performances différentes. L'interpolation la plus simple consiste à affecter la valeur du pixel le plus proche de la position recherchée. C'est une méthode rapide mais qui ne donne généralement pas de bons résultats pour des résolutions raisonnables. Nous avons également implémenté un algorithme d'interpolation bilinéaire ainsi que l'algorithme d'interpolation bicubique décrit dans Press (2007) qui donne des résultats très proches de l'interpolation optimale en  $\text{sinc}(x)$  (aussi appelée interpolation de Shannon). Cette dernière est lourde en temps de calcul car elle dépend de la valeur de l'ensemble des points de la carte. Enfin nous avons développé une méthode d'interpolation basée sur une transformée de Fourier. Le principe consiste à calculer la transformée de Fourier de la carte non lentillée et d'évaluer le signal en une position arbitraire en appliquant l'équation de base de la transformée de Fourier. C'est une technique simple mais qui demande beaucoup de temps de calcul spécialement lorsque la résolution de la carte est élevée, car il n'existe pas d'algorithme de transformation inverse rapide en un point arbitraire. Les algorithmes de FFT par exemple, nécessitent une répartition régulière des pixels.

Finalement, on a vu qu'il existe une corrélation entre le champ  $\phi$  et la température du CMB (voir section 3.4.2). Il n'est pas nécessaire d'en tenir compte pour une analyse de l'effet de lentille gravitationnelle. Cependant, il est toutefois possible de prendre en compte cette corrélation, en faisant une simulation contrainte de la carte de température et de potentiel de lentille.

## 6.4 Simulation des effets instrumentaux

Afin de mesurer l'impact des effets instrumentaux sur la qualité des données et de fournir des limites sur les effets systématiques, il est nécessaire de faire des simulations informatiques. Dans PLANCK des programmes de simulation des instruments sont développés depuis de nombreuses années. Ils décrivent de façon précise l'effet de l'ensemble de la chaîne d'observation optique et électronique. Ils permettent également de simuler le balayage du ciel et d'obtenir ainsi un jeu de données comparable aux observations provenant directement du satellites. Ces outils sont toujours en développement. La compréhension et l'estimation de l'impact des effets instrumentaux sur les mesures, est un enjeu majeur et une des principales difficultés de la mission. Cependant, une étude complète des effets systématiques instrumentaux demande beaucoup de temps et va au delà du cadre de cette thèse. Par ailleurs, la complexité de ces simulations ne permet pas de réaliser des tests rapides dont nous avons besoin pour développer un outil de mesure de l'effet de lentille gravitationnelle. Nous avons donc développé un ensemble de simulations indépendantes et rapides, en nous concentrant sur les effets les plus importants et en faisant des approximations qui permettent de simplifier l'étude tout en donnant une bonne idée de l'impact global.

### 6.4.1 Simulation de l'effet du lobe de l'instrument sur les cartes

Un système optique possède une certaine acceptance angulaire qui est une fonction de la position. On définit la direction d'observation comme le maximum de cette fonction. Les détecteurs de PLANCK, placés au foyer du système optique, constitué par les deux miroirs et les cornets, mesure donc une intensité  $\mathcal{I}$  qui est l'intégrale du signal sur le ciel,  $I$ , dans la zone d'acceptance angulaire

Fréquence (en GHz)	100 (P)	143	143 (P)	217	217 (P)	353	353 (P)	545	857
Ellipticité	1.18	1.02	1.04	1.13	1.15	1.07	1.14	1.25	1.03

TABLE 6.1 – **Tableau récapitulatif de l’ellipticité des lobes.** La mesure du lobe des différentes voies a été effectuée à partir des observations de la planète Mars. Elle est tirée de Planck HFI Core Team et al. (2011).

$\mathcal{B}$  :

$$\mathcal{I}(\mathbf{n}) = \int \int I(\mathbf{n}') \mathcal{B}(\mathbf{n} - \mathbf{n}') d\mathbf{n}' \quad (6.9)$$

ce qui déforme l’image obtenue et se traduit par un lissage, ou un flou de l’image (voir figure 6.5(b)). On peut également voir cette fonction comme l’image d’une source ponctuelle par le système optique, ce qui correspond à la convolution de l’image par la réponse impulsionnelle spatiale de l’optique (aussi appelée lobe de l’instrument ou encore PSF pour le terme anglais Point Spread Function).

Chaque mesure correspond donc à la convolution du ciel dans la direction d’observation par le lobe de l’instrument, ce qui mélange l’information dans une zone comparable à la taille du lobe. La simulation d’une observation avec un instrument possédant un lobe arbitraire est une opération très lourde car chaque point de mesure nécessite le calcul d’un produit de convolution en tenant compte de l’orientation de la mesure. Dans le cas d’un lobe circulaire les calculs sont grandement simplifiés car la réponse impulsionnelle ne dépend plus que de la distance angulaire  $\mathcal{B}(\mathbf{n} - \mathbf{n}') = \mathcal{B}(\mathbf{n} \cdot \mathbf{n}')$  et ne dépend plus de l’orientation. Pour simplifier les calculs on peut effectuer l’opération de convolution dans l’espace de Fourier. Elle se résume alors à la simple multiplication des transformées de Fourier du ciel et du lobe orienté dans la direction d’observation. Dans le cas particulier d’un lobe circulaire la transformée de Fourier du lobe est la même quelle que soit l’orientation et elle ne dépend que du module de  $\mathbf{k}$ . Enfin si le lobe est en plus un lobe gaussien l’opération de convolution se résume par la multiplication de la transformée de Fourier du signal par la fonction :

$$\mathcal{B}_{gauss}(k) = e^{-k^2 \sigma_{gauss}^2 / 2} \quad (6.10)$$

où  $\sigma_{gauss}$  est l’écart type de la gaussienne. Les caractéristiques du lobe sont en général plus souvent données en fonction de la largeur à mi-hauteur de la gaussienne  $whm_{gauss}$ , relié à  $\sigma_{gauss}$  par l’expression  $\sigma_{gauss} = \frac{whm_{gauss}}{\sqrt{(8 \log 2)}}$ .

Dans PLANCK le profil des cornets a été conçu afin d’obtenir un lobe principal le plus gaussien et circulaire possible. On peut mesurer le lobe de l’instrument avec une très bonne précision en observant une source suffisamment petite pour être considérée comme ponctuelle et suffisamment intense pour avoir un bon rapport signal sur bruit. En vol, les mesures pour PLANCK ont été effectuées grâce à l’observation des planètes car elles possèdent une forte intensité. On pourrait également empiler un grand nombre d’observations de sources ponctuelles ayant une plus faible intensité afin d’obtenir un bon rapport signal sur bruit. Les mesures au sol puis en vol (Planck HFI Core Team et al. (2011)), ont montré que dans une bonne approximation on pouvait considérer le lobe principal de l’instrument HFI de PLANCK gaussien et les lobes secondaires négligeables. L’ellipticité mesurée est également faible (voir tableau 6.1). Par ailleurs, lorsqu’on combine les observations pour générer une carte, l’effet du lobe résultant est un effet moyen qui tend à symétriser d’autant plus le lobe car on combine des observations dans des directions différentes. Enfin, l’effet des lobes asymétriques sur un estimateur quadratique, a été étudiée dans Hanson et al. (2010b) et a une influence principalement aux bas  $\ell$ , auxquels nous n’avons pas accès par la méthode par morceaux que nous avons développé qui utilise un estimateur quadratique pour mesurer l’effet de lentille gravitationnelle. Nous nous sommes donc limité à des lobes gaussiens circulaires. Pour les simulations, nous réalisons donc la convolution des cartes par le lobe à l’aide de l’équation 6.10. On peut voir un exemple sur la figure 6.5(b).

Enfin, dans certains cas il est nécessaire de déconvoluer les cartes. Dans le cas d’un lobe gaussien circulaire, le calcul consiste simplement à diviser l’amplitude des coefficients de Fourier par le lobe décrit dans l’équation 6.10, au lieu de la multiplier. Cette opération augmente donc exponentiellement la puissance aux grands  $k$ , ce qui produit parfois des effets indésirables, notamment une augmentation du bruit qui est en général proche d’un bruit blanc et qui est donc présent à toutes les fréquences.

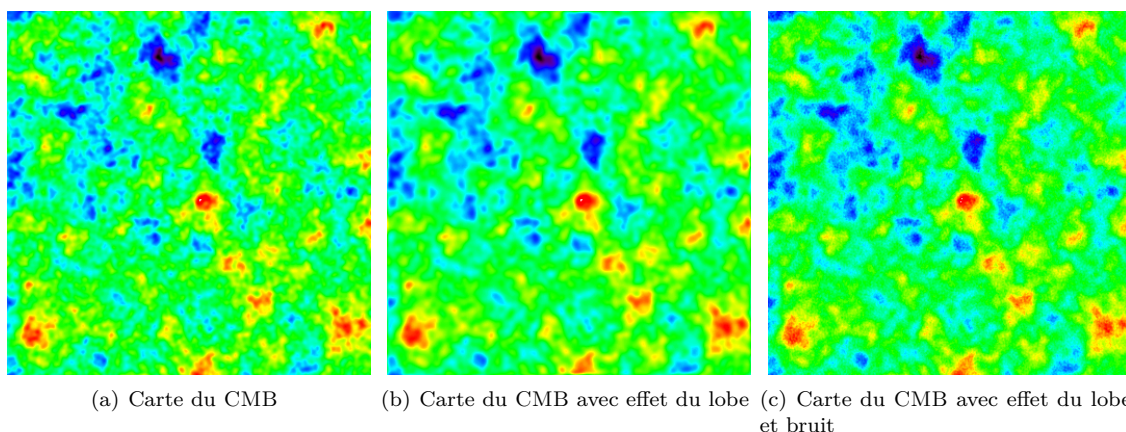


FIGURE 6.5 – **Simulation des effets du lobe de l’optique et du bruit.** L’effet du lobe de l’instrument se traduit par un lissage des petites échelles, visible sur la figure (b) comparé au signal initial (figure (a)). La carte finale (figure (c)) contient en plus du bruit. Il faut appliquer les effets dans cet ordre car le lobe ne lisse que le signal. Le bruit est produit ensuite par la chaîne d’acquisition. La largeur à mi-hauteur du lobe et le niveau de bruit utilisés sont représentatifs de ceux qu’on mesure sur les voies à 143GHz. La carte simulée fait 10 degrés de coté.

### 6.4.2 Simulation du bruit instrumental

La chaîne d’acquisition de PLANCK a été conçue afin de minimiser au maximum le bruit sur les mesures, en particulier, la cryogénie, l’électronique de polarisation et de lecture des bolomètres de HFI ainsi que l’étage d’amplification ont tous été réalisés de façon spécifique pour atteindre les objectifs ambitieux de PLANCK (voir Piat (2000)). Afin de simuler de façon précise le bruit résiduel sur les cartes, il faut appliquer la procédure globale de traitement des données en partant du signal ordonné en temps brut non démodulé. Des efforts collectifs afin de réaliser un jeu de simulations complètes et réalistes sont en cours depuis de nombreux mois, c’est un processus lourd qui demande de faire tourner un grand nombre de programmes différents, qui nécessite de fortes ressources en calcul et qui n’est toujours pas terminé à l’heure actuelle. C’est une étape essentielle à la compréhension des effets fins liés au bruit mais qui va au delà du cadre de cette thèse. Nous nous attacherons donc à simuler un bruit au niveau des cartes sans passer par le pipeline d’analyse complet.

Pour comprendre les propriétés du bruit, il est nécessaire de connaître les principales étapes d’analyse qui les influencent. En effet, le bruit résultant sur les cartes, est issu d’un processus complexe. Dans un premier temps, le bruit est généré par l’ensemble de la chaîne d’acquisition analogique, puis le signal est numérisé. À partir de cette étape le signal ne subit plus que des traitements numériques. Il est ensuite démodulé et filtré au niveau des TOI puis moyenné dans le pixel lors de la génération des cartes. Il faut ajouter en plus l’influence des rayons cosmiques dont le taux d’impact est relativement élevé et qui produisent un bruit additionnel, notamment les résidus issus de la soustraction des queues. Enfin la compression ajoute également un niveau de bruit de quantification qui même si il est faible participe au bilan total (voir à ce sujet le chapitre 5).

Les mesures actuelles montrent que dans une bonne approximation, le bruit n’est pas corrélé au signal. Pour simuler une observation en présence de bruit on peut donc simplement générer une carte de signal et lui ajouter un bruit simulé de façon indépendante. Les mesures montrent cependant que le spectre de puissance du bruit n’est pas tout à fait blanc. Il est possible d’obtenir une bonne approximation du spectre du bruit pour les différents canaux en utilisant des *Jack-Knives* sur les données. Cette méthode consiste à séparer les données en deux jeux et de faire l’analyse sur les deux séparément. On peut par exemple utiliser seulement la première moitié des cercles dans un cas puis la seconde moitié, ou encore un cercle sur deux, ou la première couverture du ciel puis la seconde. En supposant que le bruit est indépendant du signal, ce qui est une bonne approximation, on mesure alors le spectre de puissance de la différence des deux jeux de données. Le signal étant supposé constant dans une même région il est alors soustrait, seul le bruit reste présent. Cette opération n’est valable que pour un jeu de données particulier et doit être renouvelée si on change

les observations sur lesquelles on s'appuie. En effet plus on intègre de données plus le bruit diminue.

Pour simuler une carte de bruit on se base finalement sur le spectre de puissance estimé par *Jack-Knife* et on réalise un tirage gaussien. Cependant, après projection le bruit n'est pas homogène car certaines zones du ciel sont observées plus de fois que d'autres. Le nombre d'échantillons qui tombent dans chaque pixel HEALPIX est très variable (voir figure 8.1), ce qui conduit à une variation du bruit d'un pixel à l'autre. Pour avoir une distribution réaliste du bruit il faut donc pondérer chaque pixel de la carte par le nombre d'échantillons correspondant. Cette information est disponible avec chaque carte du signal générée dans PLANCK. Chaque pixel  $n_i$  de la carte de bruit est donc modulée de la façon suivante :

$$p_i = p'_i \sqrt{\frac{N}{n_i \sum_j \frac{1}{n_j}}} \quad (6.11)$$

où  $N$  est le nombre total d'échantillons de la carte ( $N = \sum_i n_i$ ),  $n_i$  est le nombre d'échantillons du pixel  $i$  et  $p'_i$  est la valeur du bruit gaussien homogène du pixel  $i$ . De façon simple le bruit diminue comme l'inverse de la racine carrée du nombre d'échantillons par pixel. On pourrait alors simplement pondérer chaque pixel  $i$  par  $1/\sqrt{n_i}$ . L'expression 6.11 ne constitue qu'une généralisation de l'expression simple précédente mais elle permet en plus de conserver la normalisation du spectre. La carte ainsi obtenue peut alors être directement additionnée à la simulation du ciel et possède directement les bonnes propriétés.

Enfin, dans le cas où on veut simuler à la fois le bruit instrumental et l'effet du lobe de l'instrument, il est important d'appliquer chaque effet dans un ordre précis. En effet, le bruit apparaît après le passage à travers le système optique. Il faut donc convoluer les cartes idéales du ciel par le lobe de l'instrument dans un premier temps, puis leur ajouter le bruit dans un second temps. On voit qu'en présence de bruit, la déconvolution d'une carte entraîne une explosion du bruit à haute fréquence car on convolue le signal uniquement mais on déconvolue le signal plus le bruit. Si on considère un bruit blanc on observe donc une explosion exponentielle du bruit après déconvolution. Il n'est pas possible de s'affranchir de cet effet.

## 6.5 Estimation du spectre de puissance du CMB

### 6.5.1 Estimation sur la sphère

Comme on l'a vu dans la section 2, le CMB est un champ gaussien isotrope dont les propriétés statistiques sont entièrement décrites par son spectre de puissance angulaire. L'estimation de ce spectre se base sur les propriétés de la décomposition en harmoniques sphériques d'un champ gaussien. Dans un premier temps, il faut commencer par estimer les coefficients  $a_{lm}$  de la décomposition en harmonique sphérique du signal. On mesure dans la pratique des fluctuations sur une sphère pixelisée. Si le ciel est complet, on détermine les  $a_{lm}$  par le calcul de l'équation (2.8) qui s'effectue de manière exacte par une quadrature de Gauss-Legendre en  $\cos\theta$  et une quadrature uniforme en  $\phi$ . C'est l'avantage d'utiliser la pixelisation HEALPIX dont les centres des pixels sont répartis de façon uniforme sur des cercles d'isolatitudes.

On obtient donc un ensemble de coefficients mesurés  $\{a_{lm}\}$  à partir desquels on peut construire l'estimateur de maximum de vraisemblance pour le spectre de puissance. La fonction de vraisemblance pour des variables gaussiennes décorréélées est donnée par :

$$\mathcal{L}(c_l; \{a_{lm}\}) = \prod_m \frac{1}{\sqrt{2\pi C_l}} e^{-\frac{|a_{lm}|^2}{2C_l}} \quad (6.12)$$

En maximisant par rapport à  $C_l$  on obtient l'estimateur :

$$\hat{C}_l = \frac{1}{2\ell + 1} \sum_m |a_{lm}|^2 \quad (6.13)$$

C'est donc l'estimateur optimal au niveau statistique que l'on peut construire.

La variance de cet estimateur peut être obtenue en constatant que la variable  $\frac{a_{lm}}{\sqrt{C_l^{th}}}$ , où  $C_l^{th}$  est le spectre *théorique* suit une loi normale et que l'estimateur donné par l'équation 6.13 suit donc la

loi  $C_l^{th}/(2\ell + 1)\chi_{2\ell+1}^2$ . Ainsi sa variance vaut :

$$\text{Var}(\hat{C}_l) = \frac{2(C_l^{th})^2}{2\ell + 1} \quad (6.14)$$

Il s'agit de la variance *cosmique* qui est irréductible. On constate qu'elle dépend de l'amplitude du spectre théorique et de l'échelle angulaire. Elle augmente aux bas  $\ell$  (nombre de modes réduits) et pour les valeurs de plus grande amplitude de  $C_l^{th}$ .

Pour estimer le spectre du puissance sur la sphère, nous utilisons la version c++ du code `anafast`, distribué avec les bibliothèques HEALPIX. Il effectue simplement les calculs décrit ci-dessus. Finalement, il faut noter que dans le cas d'un ciel partiellement masqué (où les harmoniques sphériques ne constituent plus une base orthogonale) l'estimation est plus compliquée ce qui a conduit au développement de plusieurs méthodes (décrites par exemple dans Tristram & Ganga (2007)) mais dont nous n'avons pas besoin pour la suite.

## 6.5.2 Estimation sur le plan

Dans le cas où on s'intéresse à de petites zones du ciel, possédant une taille caractéristique inférieure à une vingtaine de degrés, on peut approximer la sphère par un plan et effectuer une analyse bi-dimensionnelle. En particulier on peut calculer les coefficients de Fourier du signal  $\tilde{T}(\mathbf{l})$  :

$$\tilde{T}(\mathbf{l}) = \int T(\mathbf{n})e^{i\mathbf{l}\cdot\mathbf{n}} \quad (6.15)$$

où le vecteur  $\mathbf{n}$  parcourt la grille bi-dimensionnelle. Il faut noter que dans ce cas, le  $\mathbf{l}$  ne désigne plus une valeur entière.

On obtient la relation classique :

$$\langle \tilde{T}(\mathbf{l}_1)\tilde{T}^*(\mathbf{l}_2) \rangle = P(l_1)(2\pi)^2\delta(\mathbf{l}_1 - \mathbf{l}_2) \quad (6.16)$$

Ces coefficients se relient à ceux des harmoniques sphériques par (Okamoto & Hu (2003a), détails des calculs dans Perotto (2006)) :

$$T(\mathbf{l}) = \sqrt{\frac{4\pi}{2\ell + 1}} \sum_m i^{-m} a_{lm} e^{im\phi_l} \quad (6.17)$$

$$a_{lm} = \sqrt{\frac{2\ell + 1}{4\pi}} i^m \int \frac{d\phi_l}{2\pi} T(\mathbf{l}) e^{-im\phi_l} \quad (6.18)$$

où  $\ell$  désigne la norme de  $\mathbf{l}$  et  $\phi_l$  sa phase.

En égalant les fonctions d'autocorrélations, on trouve White et al. (1999) que le spectre de puissance donné par l'équation (6.16) *correspond au spectre obtenu sur la sphère*.

Plus précisément, si on travaille sur une grille carrée de taille  $L$ , les modes de Fourier correspondent à des harmoniques des vecteurs de déplacement  $\Delta\mathbf{k} = \frac{2\pi}{L}$  dans les 2 directions :

$$\mathbf{k} = \Delta\mathbf{k}(n\mathbf{i} + m\mathbf{j}) \quad (6.19)$$

avec ces conventions, le spectre de puissance associé est simplement  $C_{|\mathbf{k}|}$  qui désigne une interpolation du spectre harmonique. Cette relation est valable pour  $\ell \gtrsim 70$ .

De la même manière que sur la sphère, dans le cas d'un champ gaussien dans le plan, chaque mode de Fourier peut être considéré comme une variable aléatoire décorrélée gaussienne, avec une phase uniforme. L'estimateur de maximum de vraisemblance de la puissance en deux dimensions, se déroule de la même manière que sur la sphère. Pour un mode de module donnée  $k = |\mathbf{k}|$ , si on a  $N$  réalisations des coefficients de Fourier de partie réelle  $a_i$  et imaginaire  $b_i$  :

$$\tilde{C}_k = \frac{1}{N} \sum_i (a_i^2 + b_i^2) \quad (6.20)$$

En utilisant l'ergodicité, on peut remplacer la moyenne d'ensemble par la moyenne des coefficients de Fourier de même module sur la grille. On peut aussi autoriser une tolérance sur celle-ci, ce qui



revient à prendre les coefficients  $(a_i, b_i)$  à l'intérieur d'une couronne  $[k - \Delta, k + \Delta]$ , c'est-à-dire à binner le spectre.

Finalement, il faut garder à l'esprit que l'objectif est de travailler sur des images découpées dans une sphère. Dans le cas général les patchs ainsi obtenus ne sont pas périodiques. Il est alors généralement nécessaire pour l'estimation du spectre de puissance, d'apodiser les bord afin de limiter les discontinuités qui biaisent le spectre. De plus, il est conseillé de soustraire la moyenne avant de déterminer les coefficients de Fourier. Par analogie avec la méthode d'estimation d'un spectre de puissance sur un signal à une dimension par périodogramme de Welsh, lorsque les images sont découpées sur la sphère, il est important d'obtenir un recouvrement entre les patchs afin d'améliorer l'estimation du spectre de puissance. Dans le cas idéal, à une dimension, il a été montré qu'un recouvrement qui compense la perte de puissance introduite par l'apodisation permet de minimiser les erreurs (Oppenheim & Schafer (1975)). Dans le cas à deux dimensions, il n'est pas possible de paver la sphère de façon uniforme en assurant cette condition entre chaque patch. Cependant un recouvrement moyen entre tout couple de patch de 25% est un bon compromis et s'approche du cas à une dimension.

## 6.6 Estimation du spectre de déflexion

Au cours de leur parcours de la surface de dernière diffusion vers le détecteur, les photons du CMB sont légèrement déviés de leur trajectoire, par l'effet de lentille gravitationnelle produit par les grandes structures. Dans cette section nous allons montrer comment l'empreinte caractéristique laissée par ces petites déviations, dans les anisotropies de température et de polarisation du CMB, peut être utilisée pour produire un estimateur du spectre de puissance de la matière.

### 6.6.1 Présentation

Le champ aléatoire des fluctuations de la température du CMB, au niveau de la surface de dernière diffusion, qu'on notera  $T$ , est un champ gaussien. Cependant, la position d'arrivée des photons observés au niveau du détecteur, est légèrement modifiée par les déflexions causées par l'effet de lentille gravitationnelle des grandes structure rencontrées par les photons pendant leur trajet. Le champ observé, qu'on notera  $\tilde{T}$ , n'est donc plus tout à fait gaussien. Il est décrit par l'équation 3.12, dans laquelle le déplacement en chaque direction est donné par un vecteur  $\mathbf{d}$ , appelé champ de déflexion. Il est courant d'exprimer ce champ comme le gradient d'un potentiel de lentille  $\phi$  (voir section 3.7) qui correspond au potentiel gravitationnel à trois dimensions, intégré le long du parcours.

La question qui se pose alors est de trouver le meilleur estimateur du champ  $\phi$  (et de son spectre de puissance). Il s'agit d'un problème difficile, qui a été résolu par W. Hu (voir Hu (2001) et Hu & Okamoto (2002)) dans le cadre de l'approximation plane, à l'aide d'un estimateur quadratique. Ce résultat a été généralisé sur la sphère dans Okamoto & Hu (2003a). Plus tard il a été montré par Hirata & Seljak (2003b) que cet estimateur correspond bien à celui du maximum de vraisemblance dans le cas de bruits du niveau de ceux de Planck. L'estimateur de Hu et Okamoto peut être appliqué à des cartes de polarisation également, cependant, étant donnés les niveaux de bruit de l'expérience PLANCK, notamment pour les cartes polarisées, il a été montré qu'on ne gagne pas à prendre en compte la polarisation. On donnera par la suite la dérivation de W.Hu, pour la température uniquement, en approximation plane qui est plus intuitive.

### 6.6.2 Dérivation dans le plan

L'amplitude des déflexions étant très faible, de l'ordre de quelques minutes d'arc (voir la section 3.3.3), on peut commencer par effectuer le développement limité de l'équation 3.12 afin de comprendre le développement de l'estimateur. Au premier ordre en  $\phi$ , on obtient alors l'équation suivante :

$$\begin{aligned} \tilde{T}(\hat{\mathbf{n}}) &= T(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})) \\ &\simeq T(\hat{\mathbf{n}}) + \nabla\phi(\hat{\mathbf{n}}) \cdot \nabla T(\hat{\mathbf{n}}) + \dots \end{aligned} \tag{6.21}$$

Dans le cadre de l'approximation plane, on peut décomposer le champ de température lentillé dans l'espace de Fourier. On note :

$$\tilde{T}(\mathbf{K}) = \int d^2\hat{\mathbf{n}} \tilde{T}(\hat{\mathbf{n}}) e^{-i\mathbf{K}\cdot\hat{\mathbf{n}}} \quad (6.22)$$

la transformée de Fourier du champ lentillé  $\tilde{T}$  et

$$\tilde{T}(\hat{\mathbf{n}}) = \int \frac{d^2\mathbf{K}}{(2\pi)^2} \tilde{T}(\mathbf{K}) e^{i\mathbf{K}\cdot\hat{\mathbf{n}}} \quad (6.23)$$

sa transformée inverse. En utilisant l'équation 6.21, on peut exprimer le champ lentillé dans l'espace de Fourier en fonction du champ non lentillé  $T$ , sous la forme :

$$\begin{aligned} \tilde{T}(\mathbf{K}) &= \int \tilde{T}(\hat{\mathbf{n}}) e^{-i\mathbf{K}\cdot\hat{\mathbf{n}}} d^2\hat{\mathbf{n}} \\ &= \int T(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})) e^{-i\mathbf{K}\cdot\hat{\mathbf{n}}} d^2\hat{\mathbf{n}} \\ &\approx \int [T(\hat{\mathbf{n}}) + \nabla\phi(\hat{\mathbf{n}}) \cdot \nabla T(\hat{\mathbf{n}})] e^{-i\mathbf{K}\cdot\hat{\mathbf{n}}} d^2\hat{\mathbf{n}} \\ &\approx \int T(\hat{\mathbf{n}}) e^{-i\mathbf{K}\cdot\hat{\mathbf{n}}} d^2\hat{\mathbf{n}} + \int [\nabla\phi(\hat{\mathbf{n}}) \cdot \nabla T(\hat{\mathbf{n}})] e^{-i\mathbf{K}\cdot\hat{\mathbf{n}}} d^2\hat{\mathbf{n}} \end{aligned} \quad (6.24)$$

qu'on peut réécrire :

$$\tilde{T}(\mathbf{K}) \approx T(\mathbf{K}) + \int [\nabla\phi(\hat{\mathbf{n}}) \cdot \nabla T(\hat{\mathbf{n}})] e^{-i\mathbf{K}\cdot\hat{\mathbf{n}}} d^2\hat{\mathbf{n}} \quad (6.25)$$

en développant et en identifiant  $T(\mathbf{K})$ , à la forme équivalente de l'équation 6.22. En remplaçant alors les champs  $T(\hat{\mathbf{n}})$  et  $\phi(\hat{\mathbf{n}})$  par leur décomposition en terme de coefficients de Fourier sous forme équivalente à l'équation 6.23 et en effectuant le gradient dans l'espace de Fourier, on obtient alors :

$$\begin{aligned} \tilde{T}(\mathbf{K}) &\approx T(\mathbf{K}) + \int \left[ \int -i\mathbf{k}_1 \phi(\mathbf{k}_1) e^{i\mathbf{k}_1\cdot\hat{\mathbf{n}}} \frac{d^2\mathbf{k}_1}{(2\pi)^2} \cdot \int -i\mathbf{k}_2 T(\mathbf{k}_2) e^{i\mathbf{k}_2\cdot\hat{\mathbf{n}}} \frac{d^2\mathbf{k}_2}{(2\pi)^2} \right] e^{-i\mathbf{K}\cdot\hat{\mathbf{n}}} d^2\hat{\mathbf{n}} \\ \tilde{T}(\mathbf{K}) &\approx T(\mathbf{K}) - \iint \mathbf{k}_1 \cdot \mathbf{k}_2 \phi(\mathbf{k}_1) T(\mathbf{k}_2) \frac{d^2\mathbf{k}_1}{(2\pi)^2} \left[ \int e^{-i(\mathbf{K}-\mathbf{k}_1-\mathbf{k}_2)\cdot\hat{\mathbf{n}}} d^2\hat{\mathbf{n}} \right] \frac{d^2\mathbf{k}_2}{(2\pi)^2} \end{aligned} \quad (6.26)$$

Le membre entre crochet s'identifie à la fonction delta de Dirac  $\delta(-\mathbf{k} + \mathbf{k}_1 + \mathbf{k}_2)$ . Finalement en intégrant sur  $\mathbf{k}_1$  on obtient :

$$\tilde{T}(\mathbf{K}) = T(\mathbf{K}) - \int \frac{d^2\mathbf{k}_2}{(2\pi)^2} T(\mathbf{k}_2) \phi(\mathbf{K} - \mathbf{k}_2) [(\mathbf{K} - \mathbf{k}_2) \cdot \mathbf{k}_2] + \dots \quad (6.27)$$

où les termes  $\mathbf{K}$ ,  $\mathbf{k}_1$  et  $\mathbf{k}_2$  sont des vecteurs dans l'espace de Fourier et sont liés par la relation  $\mathbf{K} = \mathbf{k}_1 + \mathbf{k}_2$ . On peut noter que le produit présent dans l'équation 6.21 se transforme en convolution dans l'espace de Fourier.

En utilisant le développement de l'équation 6.27, on peut alors estimer la moyenne sur un ensemble de réalisations de CMB, d'un produit de deux modes du champ de température lentillé  $\tilde{T}(\mathbf{k}_1)$  et  $\tilde{T}(\mathbf{k}_2)$ , pour une réalisation du champ  $\phi$  fixé. On trouve alors :

$$\langle \tilde{T}(\mathbf{k}_1) \tilde{T}(\mathbf{k}_2) \rangle_{\text{CMB}} = \delta(\mathbf{k}_1 + \mathbf{k}_2) C_{k_1}^{TT} + f_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) \phi(\mathbf{K}) \quad (6.28)$$

avec

$$f_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) = \mathbf{k}_1 \cdot \mathbf{K} C_{k_1}^{TT} + \mathbf{k}_2 \cdot \mathbf{K} C_{k_2}^{TT} \quad (6.29)$$

où les  $C_k^{TT}$  désignent les éléments du spectre de puissance primordial du CMB, c'est-à-dire le spectre de puissance au niveau de la surface de dernière diffusion. Il faut noter que la moyenne s'effectue sur des réalisations différentes du CMB, à  $\phi$  fixé. Dans le cas où le champ  $\phi$  est nul, on n'a pas d'effet de lentille gravitationnelle, le membre de droite de l'équation est nul, ce qui est compatible avec la définition du spectre de puissance de température non lentillé, qui s'écrit dans ce cas :

$$\langle \tilde{T}(\mathbf{k}_1) \tilde{T}(\mathbf{k}_2) \rangle = \langle T(\mathbf{k}_1) T(\mathbf{k}_2) \rangle = \delta(\mathbf{k}_1 + \mathbf{k}_2) C_{k_1}^{TT} \quad (6.30)$$

Pour deux modes  $\mathbf{k}_1$  et  $\mathbf{k}_2$  différents, le premier terme du membre de droite de l'équation 6.28 s'annule et on obtient :

$$\langle \tilde{T}(\mathbf{k}_1)\tilde{T}(\mathbf{k}_2) \rangle_{\text{CMB}} = f_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K})\phi(\mathbf{K}) \quad (6.31)$$

Cependant, la valeur moyenne de cette équation sur un ensemble de réalisations du champ  $\phi$  devient nulle. Cette équation ne donne donc pas directement un estimateur du champ de potentiel de lentille. Elle suggère cependant qu'une combinaison quadratique de différents modes d'une carte de température du CMB peut servir d'estimateur du champ  $\phi$ . En s'inspirant de cette forme, W. Hu a proposé un estimateur quadratique du potentiel (dans le domaine de Fourier) de la forme :

$$\hat{\phi}(\mathbf{K}) = A_K \int \frac{d^2\mathbf{k}_1}{(2\pi)^2} \tilde{T}(\mathbf{k}_1)\tilde{T}(\mathbf{k}_2)F_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) \quad (6.32)$$

où, dans ce cas,  $\tilde{T}$  représente la valeur mesurée sur le ciel lentillé c'est-à-dire le champ lentillé en présence du bruit et du lobe instrumental. Par ailleurs, les vecteurs  $\mathbf{K}$ ,  $\mathbf{k}_1$  et  $\mathbf{k}_2$  sont toujours reliés par la relation  $\mathbf{K} = \mathbf{k}_1 + \mathbf{k}_2$ . Enfin  $A_K$  et  $F$  sont respectivement une normalisation et un filtre qui doivent être optimisés.

La normalisation est obtenue en imposant un estimateur non biaisé pour un champ  $\phi(\mathbf{K})$  fixé. Ainsi si on moyenne sur des réalisations de CMB on veut que l'estimateur vérifie :

$$\langle \hat{\phi}(\mathbf{K}) \rangle_{\text{CMB}} = \phi(\mathbf{K}) \quad (6.33)$$

en injectant l'équation 6.32 dans l'équation précédente et en utilisant la relation 6.31 on obtient :

$$\begin{aligned} \langle \hat{\phi}(\mathbf{K}) \rangle_{\text{CMB}} &= \langle A_K \int \frac{d^2\mathbf{k}_1}{(2\pi)^2} \tilde{T}(\mathbf{k}_1)\tilde{T}(\mathbf{k}_2)F_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) \rangle_{\text{CMB}} \\ &= A_K \int \frac{d^2\mathbf{k}_1}{(2\pi)^2} \langle \tilde{T}(\mathbf{k}_1)\tilde{T}(\mathbf{k}_2) \rangle_{\text{CMB}} F_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) \\ &= A_K \int \frac{d^2\mathbf{k}_1}{(2\pi)^2} f_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K})\phi(\mathbf{K})F_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) \\ &= \phi(\mathbf{K}) \end{aligned} \quad (6.34)$$

ce qui donne :

$$A_K = \left[ \int \frac{d^2\mathbf{k}_1}{(2\pi)^2} f_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K})F_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) \right]^{-1} \quad (6.35)$$

La variance de cet estimateur (en séparant les moyennes sur des réalisations de CMB, de celles sur le potentiel de lentille) s'écrit :

$$\langle \langle \hat{\phi}^*(\mathbf{K})\hat{\phi}(\mathbf{K}') \rangle_{\text{CMB}} - \langle \hat{\phi}^*(\mathbf{K}) \rangle_{\text{CMB}}\langle \hat{\phi}(\mathbf{K}') \rangle_{\text{CMB}} \rangle_{\phi} = \delta(\mathbf{K} - \mathbf{K}') [N_K^{(0)} + \dots] \quad (6.36)$$

Le second terme du membre de gauche de l'équation précédente, fait intervenir le spectre du potentiel de lentille puisque  $\langle \hat{\phi}(\mathbf{K}) \rangle_{\text{CMB}} = \phi(\mathbf{K})$  et que le champ  $\phi$  est gaussien  $\langle \phi(\mathbf{K})\phi^*(\mathbf{K}') \rangle_{\phi} = \delta(\mathbf{K} - \mathbf{K}')C_K^{\phi}$

Le terme  $N_K^{(0)}$  correspond au bruit gaussien de l'estimateur, c'est-à-dire à la puissance qui apparaît lorsque les cartes ne sont pas lentillées. Il existe des termes d'ordre supérieur qui sont couplés au champ gravitationnel : ils ont été calculés dans Kesden et al. (2003) et Hanson et al. (2011) (voir section 6.6.3). Ils sont néanmoins d'amplitude plus faible. Le bruit peut se mettre sous la forme (Hu & Okamoto (2002)) :

$$N_K^{(0)} = [A(K)]^2 \int \frac{d^2\mathbf{k}_1}{(2\pi)^2} 2C_{k_1}^{\tilde{T}_{obs}}\tilde{T}_{obs}C_{k_2}^{\tilde{T}_{obs}}\tilde{T}_{obs}F_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) \quad (6.37)$$

où  $\mathbf{K} = \mathbf{k}_1 + \mathbf{k}_2$  et  $C_k^{\tilde{T}_{obs}}\tilde{T}_{obs}$  désigne le spectre de puissance du CMB observé c'est-à-dire le spectre lentillé tenant compte en plus des effets du bruit et du lobe instrumentaux. On peut l'exprimer sous la forme :

$$C_k^{\tilde{T}_{obs}}\tilde{T}_{obs} = C_k^{\tilde{T}\tilde{T}} + C_k^{nn} \quad (6.38)$$

où  $C_k^{mn}$  est le spectre du bruit instrumental, déconvolué du lobe de l'instrument. Il est donné par :

$$C_k^{nn} = f_{sky} \frac{4\pi}{t_{pix} N_{pix}} \left( \frac{s}{T_{CMB}} \right) e^{k^2 \sigma^2} \quad (6.39)$$

où,  $f_{sky}$  est la fraction du ciel couverte,  $N_{pix}$  le nombre de pixels total de la carte,  $t_{pix}$  le temps d'intégration par pixel,  $s$  la sensibilité du détecteur et  $\sigma$  l'écart type du lobe instrumental.

On obtient alors le filtre optimal en minimisant la variance de l'estimateur,  $N_K^{(0)}$ , par rapport à  $F$ , ce qui donne :

$$F_{\mathbf{k}_1 \mathbf{k}_2}(\mathbf{K}) = \frac{f_{\mathbf{k}_1 \mathbf{k}_2}(\mathbf{K})}{C_{k_1}^{\tilde{T}_{obs} \tilde{T}_{obs}} C_{k_2}^{\tilde{T}_{obs} \tilde{T}_{obs}}} \quad (6.40)$$

où l'on retrouve la fonction  $f$  donnée par l'équation 6.29. On en déduit que :

$$N_K^{(0)} = A_K \quad (6.41)$$

Finalement, l'estimateur quadratique optimal s'écrit donc :

$$\hat{\phi}(\mathbf{K}) = A_K \int \frac{d^2 \mathbf{k}_1}{(2\pi)^2} T(\mathbf{k}_1) T(\mathbf{k}_2) f_{\mathbf{k}_1 \mathbf{k}_2}(\mathbf{K}) / (2 C_{k_1}^{\tilde{T}_{obs} \tilde{T}_{obs}} C_{k_2}^{\tilde{T}_{obs} \tilde{T}_{obs}}) \quad (6.42)$$

Il faut prendre garde au fait que les  $C_k^{TT}$  de l'équation 6.29 sont ceux du spectre primordial du CMB au niveau de la surface de dernière diffusion, tandis que les  $C_k^{\tilde{T}_{obs} \tilde{T}_{obs}}$  du filtre (équation 6.40) correspondent à ceux du spectre mesuré sur les données (avec bruit et après déconvolution du lobe). L'estimateur effectue donc une forme de filtrage de Wiener sur les données Hu (2001).

Cette estimateur permet la reconstruction de la carte de  $\hat{\Phi}$  à partir de laquelle on peut estimer le spectre de puissance angulaire du potentiel de lentille. Il est donnée par :

$$\begin{aligned} \hat{C}_K^{\phi\phi} &= \langle |\hat{\phi}(\mathbf{K})|^2 \rangle + N_K \\ &= C_K^{\phi\phi} + N_K \end{aligned} \quad (6.43)$$

où  $N_K = N_K^{(0)} + N_K^{(1)} + N_K^{(2)} + \dots$  et le terme  $N_K^{(0)}$  est donné par l'équation 6.41 et dépend donc du jeu de donnée considéré par l'intermédiaire du filtre  $F$ . Le champ estimé est un champ scalaire de variance  $C_K^{\phi\phi} + N_K$ , la précision de l'estimation prend donc la forme classique (Hu & Okamoto (2002)) :

$$\Delta \hat{C}_K^{\phi\phi} \simeq \frac{C_K^{\phi\phi} + N_K}{\sqrt{K \Delta K} f_{sky}} \quad (6.44)$$

où  $\Delta K$  est la taille des intervalles sur lesquels sont moyennées les valeurs. On voit que la variance dépend directement de la fraction du ciel couverte. Par ailleurs, le terme  $N_K$  dépend des valeurs de  $C_k^{\tilde{T}_{obs} \tilde{T}_{obs}}$  qui dépend directement de la taille du lobe et du temps d'intégration comme on peut le voir dans l'équation 6.39. Ainsi, on constate que la variance de l'estimateur diminue lorsque le temps d'intégration dans les pixels augmente. Enfin, une expérience avec un lobe instrumental plus faible est également capable de réaliser une mesure avec une plus faible variance.

### 6.6.3 Bruits de l'estimateur

La dérivation de l'estimateur du champ de déflexion est basé sur un développement limité du gradient du potentiel de lentille. Dans un premier temps, Hu et Okamoto ont développé leurs travaux en travaillant sur une expansion au premier ordre en  $\phi$ . Il a été montré par la suite par Kesden et al. (2003) qu'en poussant le développement au second ordre on trouvait un biais additionnel  $N_K^{(1)}$  non négligeable. Son amplitude relative domine même celle du spectre de la déflexion à haut  $\ell$ . Ce terme de biais dépend du spectre de puissance de la déflexion.

Récemment, il a été montré qu'un biais résiduel était présent dans l'estimation du spectre de la déflexion sous la forme d'un manque de puissance pouvant atteindre 15%. Ce biais résiduel est principalement significatif aux  $\ell$  inférieurs à 250. Il a été montré dans Hanson et al. (2011) que ce biais était bien décrit par le terme d'ordre quatre en  $\phi$ . Par ailleurs, il a aussi été montré dans

ces mêmes travaux, que les termes d'ordre supérieur ne produisaient pas de biais supplémentaire significatif.

Finalement, en suivant les notations de Hanson et al. (2011), l'expression des biais  $N_K^{(1)}$  et  $N_K^{(2)}$ , dans le cadre de l'approximation plane, est donnée par :

$$N_K^{(1)} = -4A_K^2 \int \frac{d^2\mathbf{k}_1}{(2\pi)^2} \frac{d^2\mathbf{k}_3}{(2\pi)^2} F_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) F_{\mathbf{k}_3\mathbf{k}_4}(\mathbf{K}) [\mathbf{k}_1 \cdot (\mathbf{k}_1 - \mathbf{k}_3) \mathbf{k}_2 \cdot (\mathbf{k}_1 - \mathbf{k}_3) C_{|l_1-l_3|}^{\phi\phi} C_{l_1}^{TT} C_{l_2}^{TT} + \mathbf{k}_1 \cdot (\mathbf{k}_1 - \mathbf{k}_4) (-\mathbf{k}_3) \cdot (\mathbf{k}_1 - \mathbf{k}_4) C_{|l_1-l_4|}^{\phi\phi} C_{l_1}^{TT} C_{l_3}^{TT}] \quad (6.45)$$

et

$$N_K^{(2)} = 4C_K^{\phi\phi} A_K^2 \left( \int \frac{d^2\mathbf{k}_3}{(2\pi)^2} F_{\mathbf{k}_3\mathbf{k}_4}(\mathbf{K}) f_{\mathbf{k}_3\mathbf{k}_4}(\mathbf{K}) \int \frac{d^2\mathbf{k}_1}{(2\pi)^2} F_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K}) \left( \int \frac{d^2\mathbf{k}}{(2\pi)^2} \mathbf{k} \cdot \mathbf{K} [\mathbf{k} \cdot (\mathbf{k} - \mathbf{k}_1)]^2 C_k^{TT} C_{|l-l_1|}^{\phi\phi} - R \mathbf{k}_1 \cdot \mathbf{K} k_1^2 C_{k_1}^{TT} \right) \right) \quad (6.46)$$

avec  $f_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K})$  et  $F_{\mathbf{k}_1\mathbf{k}_2}(\mathbf{K})$  définis par les équations 6.29 et 6.40 respectivement et  $R = \frac{1}{4\pi} \int K^3 C_K^{\phi\phi} dK$ .

#### 6.6.4 Généralisation à la sphère

Le principe de l'estimation du spectre de puissance de la déflexion sur la sphère est le même que celui expliqué précédemment. L'estimation du potentiel de lentille est basé sur une combinaison quadratique de la carte de température associée à un filtrage dans l'espace des harmoniques sphériques. Le développement est cependant plus calculatoire. On travaille dans ce cas avec des harmoniques sphériques spinnées de spin 2. Les gradients sont calculés dans l'espace des harmoniques sphériques mais les dérivées spatiales sont remplacées par des dérivées covariantes sur la sphère. La dérivation complète peut être trouvée dans Okamoto & Hu (2003a). Les biais  $N_L^{(1)}$  et  $N_L^{(2)}$  donnés dans le cas du plan, sont similaires sur la sphère. Leur calcul peut être trouvé dans Hanson et al. (2011).

## Chapitre 7

# Développement d'une méthode d'analyse par patch pour Planck

### 7.1 Outils de découpe d'une sphère en patch

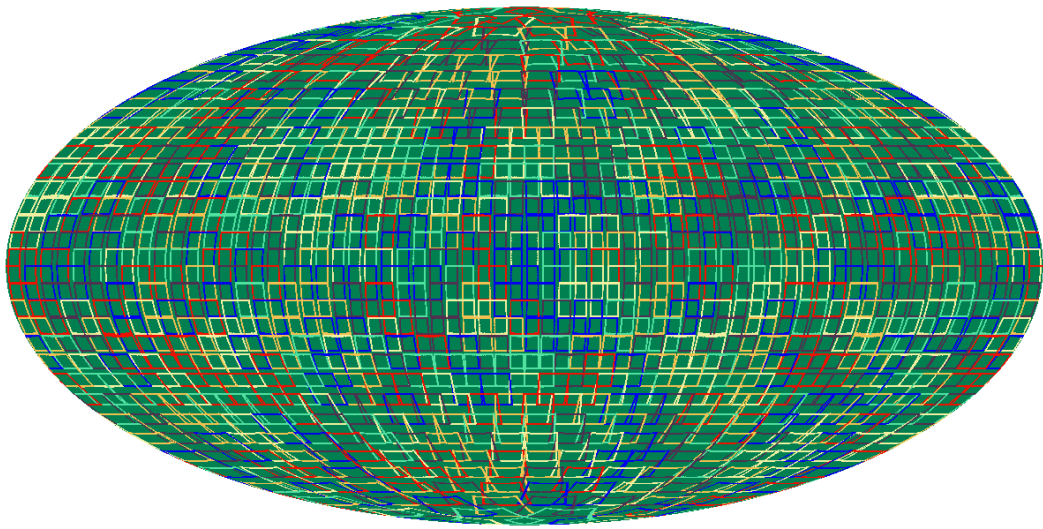
#### 7.1.1 Problématique

Aux longueurs d'onde de PLANCK, le signal est composé de plusieurs composantes : l'émission du CMB, l'émission de la poussière galactique, l'émission synchrotron, du rayonnement de freinage, des sources compactes... Ils est possible d'effectuer une séparation des composantes, mais il est plus difficile dans ce cas de quantifier les erreurs introduites et les cartes obtenues par les différentes méthodes contiennent toujours des zones contaminées. Pour étudier le CMB, il est donc généralement nécessaire de masquer les portions du ciel les plus contaminées par les autres émissions. Les émissions galactiques sont en général plus intenses dans le plan galactique mais des zones à haute latitude possèdent également une intensité importante. Il est souvent difficile de tenir compte de l'effet d'un masque de taille et de forme arbitraire dans les estimations sur toute la sphère. La pixelisation de la sphère étant un problème complexe, de nombreuses pixelisations existent. Suivant la répartition des pixels dans l'espace, les algorithmes peuvent être complexes à écrire. La pixelisation HEALPIX a été pensée pour pouvoir faire des calculs rapides sur des cercles de latitude constante. Cependant la répartition des pixels sur la sphère reste complexe et dans de nombreux cas les algorithmes demandent des calculs complexes qui nécessitent donc un temps d'exécution long et de grandes ressources en calcul. Les données de PLANCK sont projetées sur des sphères HEALPIX avec un paramètre *nside* de 2048. Cela correspond à plus de 50 millions de pixels ( $12nside^2$ ). Pour cette raison il est intéressant de pouvoir sélectionner des portions de ciel les moins contaminées afin de faire une étude par morceaux, sur un plan, sur des zones réduites du ciel.

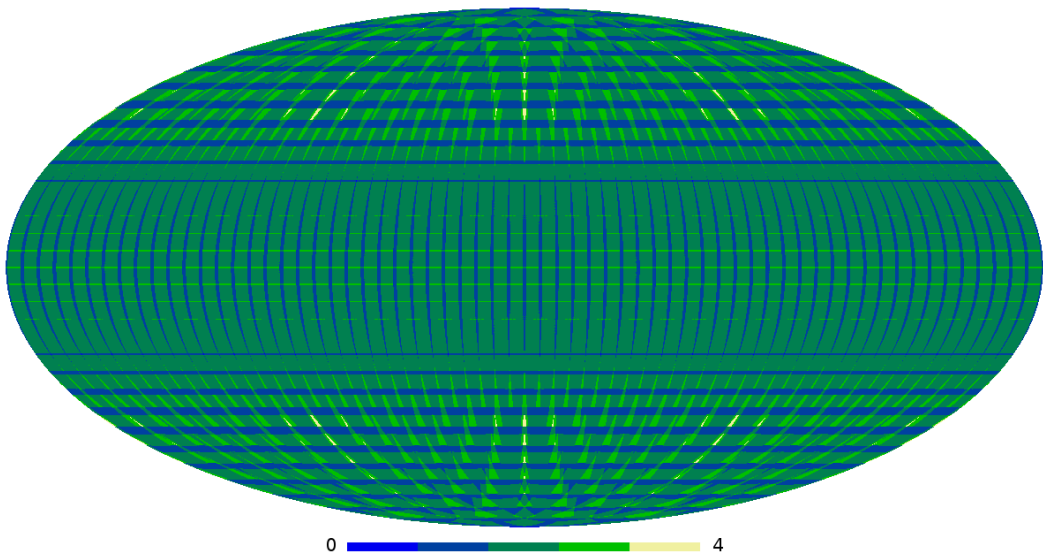
La projection d'une portion de la sphère permet de générer des images pixelisées de façon régulière dans un plan. Beaucoup d'algorithmes sont optimisés pour utiliser la transformée de Fourier discrète rapide (FFT). Le travail sur des patches permet donc de se limiter aux zones du ciel les moins contaminées, de travailler sur des portions plus petite à la fois et d'utiliser des algorithmes optimisés. Ce travail peut être également utile pour comparer des observations différentes qui ne couvrent pas nécessairement la totalité du ciel. Enfin, couplé à un inpainting local, l'analyse par patch permet de s'affranchir de l'inpainting du plan galactique.

#### 7.1.2 Pavage de la sphère

Le pavage d'une surface consiste à trouver un ensemble de tuiles permettant de recouvrir la totalité de la surface sans recouvrement. Le cas particulier du pavage de la surface d'une sphère est un problème complexe et ancien. Une technique simple connue depuis l'antiquité consiste à projeter sur la surface de la sphère, les arêtes d'un solide convexe entièrement contenu dans la sphère afin de délimiter les différentes tuiles. Dans notre cas, la problématique est légèrement différente puisque nous cherchons à répartir un ensemble de patches de forme rectangulaire qui représentent une portion de la sphère projetée. Le seul solide convexe constitué uniquement de rectangles est



(a) Pavage de la sphère



(b) Recouvrement

FIGURE 7.1 – **Pavage de la sphère et recouvrement.** La figure (a) montre un exemple de pavage d'une sphère par un ensemble de patches carrés de 10 degrés de coté, centrés sur les centres des pixels d'une carte HEALPIX avec un paramètre  $n_{side}$  de 8. Les contours de chaque patch sont représentés par les traits de couleur qui ont été choisis aléatoirement afin de pouvoir bien distinguer chaque patch. Cette configuration permet de couvrir 99.96% des pixels. La figure (b) donne le recouvrement correspondant obtenu en comptant le nombre de patches chevauchant chaque pixel. Pour deux patches pris deux à deux on a un recouvrement d'environ 25%, ce qui donne un recouvrement global d'environ 100% en comptant les quatre patches voisins. On voit que la couverture est moins homogène aux pôles mais elle reste toutefois relativement constante sur le ciel. Le recouvrement de chaque pixel varie de 0 à 4. On peut remarquer que la valeur 2 (en vert sombre) domine largement, ce qui est conforme aux attentes. Les pixels non observés (en bleu vif) ne sont pas visibles sur la figure. Ils sont situés dans les régions polaires.

un parallélépipède rectangle qui ne possède que six faces. Il est possible de subdiviser ces faces mais dans ce cas les distortions introduites par la projection varient d'un patch à l'autre. Si on veut découper plus de six patches, il n'existe donc pas de solution optimale permettant de remplir à la fois les conditions de couverture totale et de non chevauchement des patches. Suivant le but recherché il faut trouver un compromis entre la surface perdue et le recouvrement des patches, ce qui ne correspond donc plus exactement à un pavage au sens mathématique du terme.

Nous ne cherchons cependant pas nécessairement à recouvrir la totalité de la sphère, au contraire la méthode que nous avons développée essaye entre autre de répondre au problème du masquage des données et doit donc pouvoir être utilisée sur une fraction incomplète du ciel. Par ailleurs, pour l'estimation du spectre de puissance du CMB, nous avons vu dans la section 6.5.2 qu'il est judicieux d'avoir un recouvrement proche de 25% entre deux patch pris deux à deux, par analogie à l'estimation d'un spectre de puissance à une dimension par la méthode du périodogramme de Welsh. Afin d'obtenir une résolution en fréquence identique dans les deux directions des patches, nous avons choisi de nous limiter à des patches carrés pour notre étude, cependant l'ensemble des codes sont adaptés à des tailles arbitraires. Les paramètres libres sont donc la position du centre des patches et leur angle de rotation. Le problème consiste alors à optimiser la répartition des centres des patches et leur orientation afin d'obtenir une couverture du ciel la plus homogène possible. On cherche par ailleurs à obtenir un chevauchement radial d'environ 50%, à partir du centre des patches, pour des patches pris deux à deux, afin de compenser approximativement la perte de puissance introduite par l'apodisation. Si on suppose que chaque patch est entouré de quatre voisins, la condition précédente conduit à une couverture commune d'environ 25% entre des patches voisins pris deux à deux.

Pour estimer l'effet de lentille gravitationnelle, nous avons montré dans la section 7.4.2 que des patches de 10 degrés de coté permettaient de limiter certains effets systématiques. Nous avons donc conservé cette taille pour l'ensemble de l'étude. Un simple calcul montre alors qu'il suffit de découper environ 800 patches sur la sphère pour obtenir un recouvrement proche de 25%. Nous avons ainsi choisi une répartition des centres, basée sur le centre des pixels d'une carte HEALPIX de paramètre *nside* de 8. On obtient alors 768 positions réparties de façon uniforme sur l'ensemble de la sphère. Cette approche à l'avantage d'être relativement simple à mettre en place et nécessite très peu de temps de calcul. Par ailleurs il existe des routines HEALPIX permettant de calculer la position des centres des pixels pour un paramètre *nside* arbitraire. La procédure peut donc être adaptée à des tailles de patches plus ou moins grande ou pour obtenir un recouvrement différent entre les patches. La figure 7.1 montre la répartition de patches de 10 degrés de coté obtenue avec la méthode précédente. On peut également constater que le recouvrement est globalement homogène même si il est inévitable d'avoir de légères différences. Les régions polaires sont moins uniformes que les régions équatoriales.

### 7.1.3 Projection

La projection de données d'une sphère sur un plan est un problème bien connu, étudié depuis de nombreuses années principalement dans le domaine de la cartographie. Plusieurs dizaines de projections ont été inventées. Elles possèdent des propriétés différentes et répondent à des problèmes spécifiques. Nous avons utilisé une projection gnomonique. Elle fait partie de la classe des projections dites azimutales qui regroupe l'ensemble des opérations de projection de la surface d'une sphère sur un plan depuis un certain point de perspective. La projection gnomonique correspond à la projection depuis le centre de la sphère sur un plan tangent (voir figure 7.2). Ainsi un point  $P$  de la sphère situé en  $(\theta, \phi)$  (où  $\theta$  est la colatitute qui varie entre 0 et  $\pi$  avec l'origine au pôle Nord et  $\phi$  est la longitude qui varie entre 0 et  $2\pi$ ) est projeté en un point  $P'$  du plan tangent de coordonnées  $(x, y)$  donnée par :

$$y = \frac{\cos \theta \sin \theta_0 - \sin \theta \cos \theta_0 \cos(\phi - \phi_0)}{\cos \theta \cos \theta_0 + \sin \theta \sin \theta_0 \cos(\phi - \phi_0)} \quad (7.1)$$

$$x = \frac{\sin \theta \sin(\phi - \phi_0)}{\cos \theta \cos \theta_0 + \sin \theta \sin \theta_0 \cos(\phi - \phi_0)} \quad (7.2)$$

où  $\theta_0$  et  $\phi_0$  représente les coordonnées du point d'intersection du plan tangent et de la sphère.

Il est parfois utile de réaliser l'opération réciproque qui consiste à reprojeter le point du plan sur



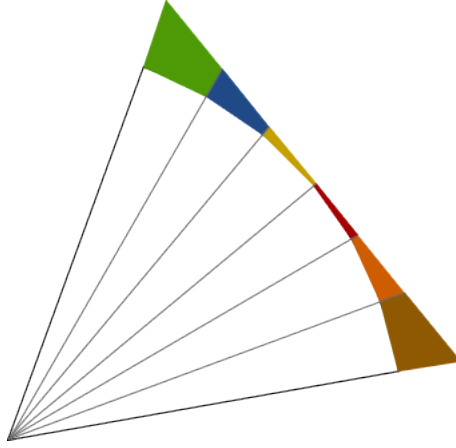


FIGURE 7.2 – **Projection gnomonique.** La projection gnomonique est la projection depuis le centre de la sphère sur un plan tangent. Comme toute projection, elle introduit des distortions. Elles sont d'autant plus importantes qu'on s'éloigne du point tangent. On peut voir que les pixels aux extrémités (brun et vert) sont plus allongés que les pixels proches du centre (rouge et jaune)

la sphère. Les équations sont alors données par :

$$\theta = \tan^{-1} \left( \frac{\sqrt{XX^2 + YY^2}}{ZZ} \right) \quad (7.3)$$

$$\phi = \tan^{-1} \left( \frac{YY}{XX} \right) \quad (7.4)$$

avec

$$\begin{aligned} XX &= \sin \theta_0 \cos \phi_0 - y \cos \theta_0 \cos \phi_0 - x \sin \phi_0 \\ YY &= \sin \theta_0 \sin \phi_0 - y \cos \theta_0 \sin \phi_0 + x \cos \phi_0 \\ ZZ &= \cos \theta_0 + y \sin \theta_0 \end{aligned}$$

La principale propriété de la projection gnomonique est de conserver la plus courte distance entre deux points. Elle transforme ainsi tous les grands cercles en lignes droites. Il faut noter qu'il n'est possible de projeter qu'une fraction d'un hémisphère à la fois ce qui est toutefois suffisant pour des patches de 10 degrés de coté. Comme toute projection, elle introduit des distortions. On peut montrer que les distortion sont d'autant plus importantes qu'on s'éloigne du point tangent. La distance radiale varie avec la distance  $d$  au point tangent comme :

$$r_r(d) = \frac{1}{\cos(d)} \quad (7.5)$$

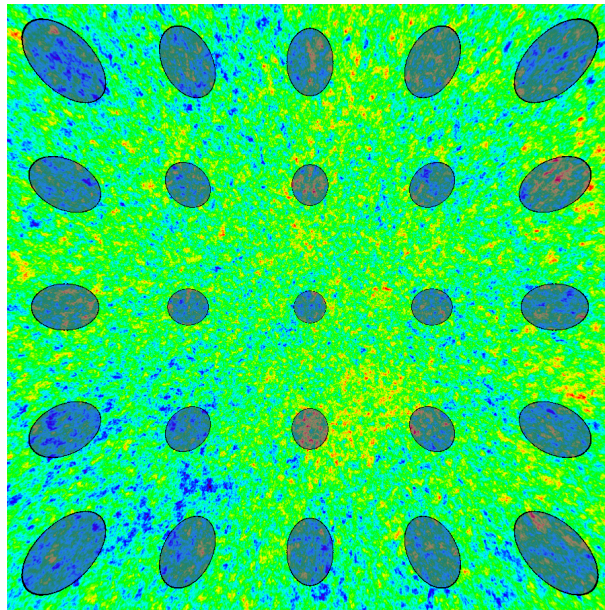
et la distance transverse comme :

$$r_t(d) = \frac{1}{\cos^2(d)}. \quad (7.6)$$

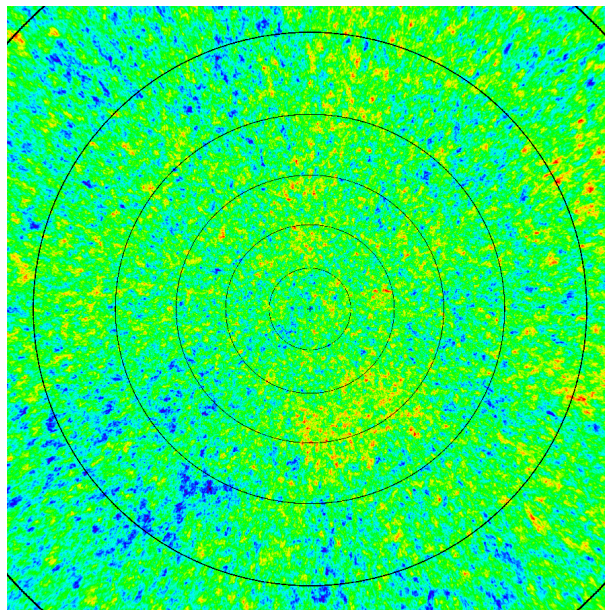
Afin de donner une idée des déformations introduites par une projection, il est courant d'utiliser les indicatrices de Tissot. Ce sont des cercles de même taille, dessinés sur la sphère. Leur image est en général des ellipses qui permettent de mettre en évidence les distortions et de donner une idée des propriétés de la projection (conservation locale des angles ou des surfaces). On peut voir sur la figure 7.3 la déformation introduite par la projection gnomonique sur des figures de test à la surface de la sphère. La projection gnomonique ne conserve ni les angles (les cercle deviennent des ellipses) ni les surfaces (la surface des indicatrices grandit). Il faut noter toutefois qu'aucune projection ne peut vérifier l'ensemble de ces propriétés à la fois.

#### 7.1.4 Estimation du signal sur une grille irrégulière

Les pixels de la sphère, projetés sur le plan, tombent sur une grille irrégulière (voir figure 7.5). L'ensemble des codes d'analyse dans le plan travaille sur des images à deux dimensions avec une

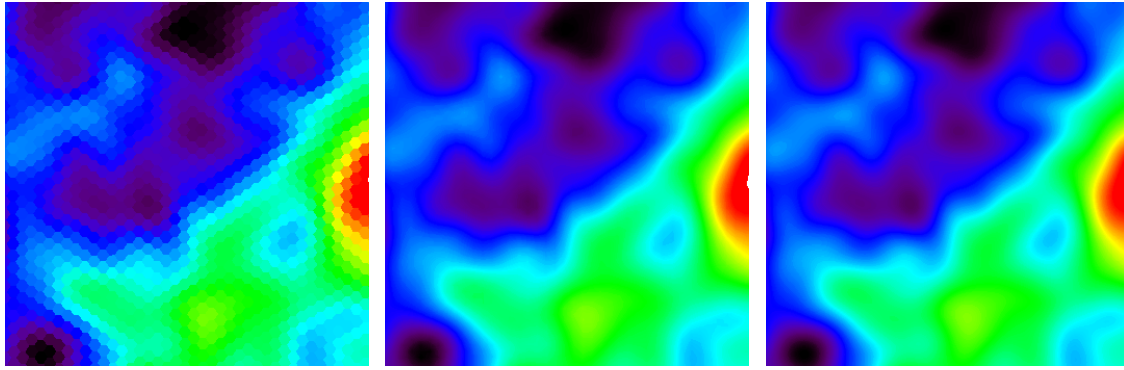


(a) Indicatrices de Tissot



(b) Indicatrices de Tissot

FIGURE 7.3 – **Figures de distortion.** Sur la figure (a) on peut voir les indicatrices de Tissot de la projection gnomonique d'un patch carré, de 105 degrés de coté. Ces figures permettent de donner une idée des distortions introduites par la projection. Elles correspondent à l'image de cercles identiques de la sphère. On observe à la fois une élongation des cercles qui signifie que la projection gnomonique n'est pas conforme (elle ne conserve pas les angles localement) et une augmentation de la surface des indicatrices qui signifie que la projection gnomonique n'est pas non plus une projection équivalente (elle ne conserve pas les surfaces). On peut voir sur la figure (b) la projection de cercles concentriques centrés sur le point tangent, dont les rayons varient par pas de 10 degrés. On observe une augmentation de la distance entre les cercles images, lorsqu'on s'éloigne du centre, à cause de la distortion. Ces figures sont volontairement tracées sur un patch de grande taille afin d'amplifier les effets. Sur un patch de 10 degrés de coté, les effets sont beaucoup plus faibles. Il ne sont même plus visibles à l'oeil.



(a) Interpolation par plus proche voisin (b) Interpolation bilinéaire (c) Interpolation par Fourier

FIGURE 7.4 – **Interpolation pour la projection.** Les pixels d'une carte HEALPIX projetés sur un plan par projection gnomonique ne tombent pas dans une grille régulière. Pour obtenir une image plane, à deux dimensions et régulière il faut interpoler le signal sur la sphère, à la position correspondant aux centres des pixels sur le plan. Trois méthodes d'interpolation (plus proche voisin, bilinéaire et par Fourier) sont comparées, sur un morceau carré de 1 degré de coté extrait d'un patch de 10 degrés de coté à très haute résolution. On voit bien que la méthode par interpolation par plus proche voisin (figure (a)) entraîne des variations discontinues lorsque la valeur la plus proche passe d'un pixel à l'autre sur la sphère. Les deux autres méthodes (figures (b) et (c)) donnent des résultats beaucoup plus proches. Il est difficile de distinguer les deux par une simple comparaison des cartes. Les résultats sur l'étude des spectres de puissance de température ou de déflexion sont toutefois meilleurs avec la dernière méthode y compris en présence des effets d'un bruit et d'un lobe instrumental.

pixelisation régulière carrée. Il est donc nécessaire à partir des valeurs irrégulièrement réparties, d'estimer le signal sur une grille régulière. C'est un problème bien connu d'interpolation. Nous avons implémenté trois méthodes permettant d'y répondre.

La méthode la plus simple consiste à associer à chaque pixel du patch la valeur du pixel de la sphère le plus proche. On appelle ce type d'interpolation, l'interpolation par plus proche voisin. Ses principaux avantages sont sa rapidité et le fait qu'elle n'utilise que des valeurs existantes dans la carte de départ. La répartition des pixels HEALPIX sur des cercles d'équilatitudes puis de façon régulière en longitude, facilite grandement la recherche des pixels les plus proches d'une direction arbitraire sur la sphère. Ainsi pour chaque pixel du patch nous effectuons une projection inverse afin de déterminer la position d'origine sur la sphère qui se projette dans le centre du pixel sur le plan. À partir de cette position, nous utilisons une routine HEALPIX rapide qui retourne les quatre pixels de la sphère les plus proches. Parmi cette liste de pixels nous choisissons finalement le plus proche. Pour cela, il faut évaluer la distance entre le point correspondant au centre du pixel du patch et les centres des quatre pixels de la sphère les plus proches déterminés précédemment. L'estimation des distances doit se faire le long des géodésiques qui sont données par les grands cercles pour une sphère. Afin de faciliter les calculs, nous utilisons une méthode permettant de réaliser le calcul des distances sur le plan. Pour cela nous projetons les pixels de la liste, avec une projection gnomonique centrée sur la position en laquelle nous voulons interpoler le signal. Cette projection transforme les grands cercles en ligne droite et les distortions évoluent radialement avec la distance au centre de projection. Ainsi il suffit de mesurer la distance cartésienne la plus courte entre l'origine du patch et les pixels de la sphère projetés, pour trouver le pixel le plus proche. Comme on peut le voir sur la figure 7.4(a), la méthode d'interpolation par plus proche voisin produit des variations brutales du signal interpolé lorsque le plus proche voisin passe d'un pixel à l'autre de la sphère. Nos études ont montré qu'elle n'était pas adaptée à l'étude de l'effet de lentille gravitationnelle, tout au moins pour des cartes initiales à la résolution de PLANCK (*nside* égal à 2048).

La seconde méthode permet d'obtenir un signal interpolé beaucoup moins discontinu (voir figure 7.4(b)). Elle utilise une interpolation bilinéaire. Cette opération est réalisée avec les routines HEALPIX et fait une approximation plane. C'est-à-dire qu'elle évalue les distances entre pixels de façon cartésienne, ce qui introduit des erreurs qui sont largement négligeables, étant donné qu'à la

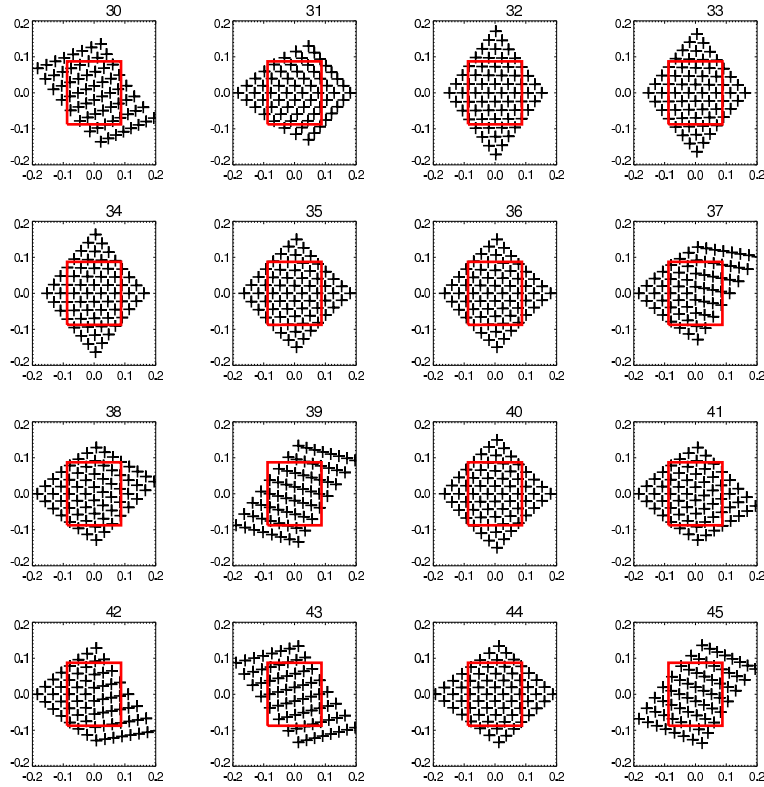


FIGURE 7.5 – **Répartitions des pixels projetés.** Les pixels de la sphère projetés sur le plan ne tombent pas sur une grille régulière. On a tracé ici la répartition des centres des pixels d'une carte HEALPIX de paramètre  $n_{\text{side}}$  de 64 (croix noires) sur un ensemble de 16 patchs carrés de 10 degrés de coté (en rouge). On observe que la répartition n'est pas régulière et dépend de la position du patch considéré.

résolution de PLANCK, la distance moyenne inter-pixels est de seulement 1.7 minutes d'arc. Elle nécessite la détermination des quatre pixels les plus proches de la position en laquelle on veut interpoler le signal. Cette liste de pixels est obtenue par la même routine que précédemment.

Finalement avec l'aide de Guy Lemeur qui est ingénieur au LAL, nous avons développé un algorithme permettant de se passer totalement d'interpolation. Pour cela, nous réalisons pour chaque patch, une estimation des coefficients de Fourier sur une grille régulière de taille définie, à partir de la répartition globale de l'ensemble des points irrégulièrement espacés issus de la projection des pixels de la sphère pour chaque patch (voir l'annexe B). C'est une méthode générale qui peut s'appliquer pour toute répartition arbitraire de l'ensemble des points de départ et peut donc être employée sur des cartes n'utilisant pas la pixelisation HEALPIX. Elle donne les meilleurs résultats pour l'étude de l'effet de lentille gravitationnelle. Sur la figure 7.4(c) on la compare, dans l'espace réel, sur une zone de 1 degré à haute résolution, aux deux autres méthodes qui utilisent elles une interpolation. Pour cela, on calcule le signal dans l'espace réel par transformée de Fourier discrète inverse à partir des coefficients estimés par la méthode d'approximation par série de Fourier.

### 7.1.5 Apodisation

L'estimation des coefficients de Fourier bi-dimensionnels nous permet de reconstruire un spectre de puissance en 2D. On part d'une carte de tout le ciel (donc sphérique) qu'on recouvre de patchs carrés de  $10^\circ$  de coté. On reconstruit les spectres de chacun des patchs en prenant simplement le module carré des coefficients de Fourier reconstruits. Le spectre final de la carte du ciel est obtenu à partir de la moyenne des spectres à deux dimensions de l'ensemble des patchs. On obtient ainsi le spectre bi-dimensionnel présenté sur la figure 7.6.

La figure précédente montre l'aspect isotrope du CMB (dépendance en  $\ell = |\mathbf{k}|$ ) et les pics acoustiques. Cependant, on peut également constater un excès d'énergie autour des axes  $k_x = 0$  et

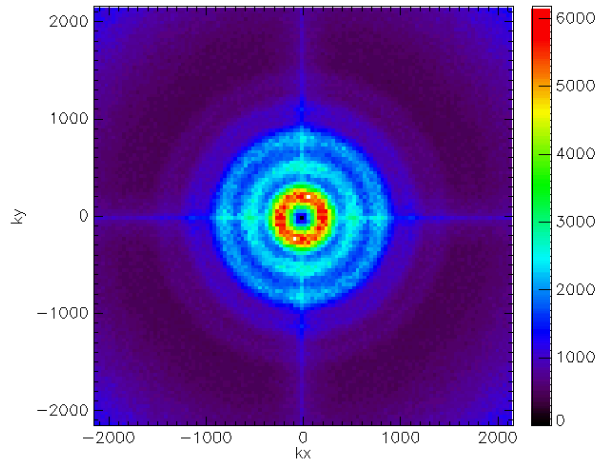


FIGURE 7.6 – **Effet de la taille finie des patches sur le spectre de puissance bi-dimensionnel.** On représente le spectre de puissance bi-dimensionnel reconstruit à partir d'une carte du ciel complet du CMB par la méthode par patches. Les axes  $x$  et  $y$  sont en unité de mode de Fourier  $\mathbf{k}$ . Le niveau de couleur représente  $\frac{|\mathbf{k}|^2}{2\pi} C_{\mathbf{k}}$ . Il est exprimé en  $\mu K^2$ .

$k_y = 0$  qui forme un signal en forme de croix.

Ce problème provient du fait que notre carte est de taille finie ( $10^\circ$  dans ce cas). En effet, par rapport au cas idéal (carte périodique ou infinie), la taille finie de la carte se traduit par l'introduction d'une multiplication dans l'espace réel par une fenêtre plate (aussi appelée « tophat » en anglais pour sa ressemblance à un chapeau). Cette multiplication dans l'espace réel, se traduit par une convolution des coefficients de Fourier par la transformée de Fourier de la fonction fenêtre. Cette fonction oscille très loin comme on peut le voir sur la figure 7.14. Or les coefficients de Fourier estimés, qui sont ceux du CMB plus du bruit, tombent approximativement en  $a_k \propto \frac{1}{k}$ . Ainsi la fuite de puissance due aux lobes secondaires de la fenêtre « tophat » est particulièrement importante pour les faibles valeurs de  $k_x$  et  $k_y$ . Il est donc nécessaire d'apodiser les patches avant d'estimer les coefficients. On cherchera donc une fonction fenêtre permettant de limiter l'amplitude des lobes secondaires (voir la section 7.4.4).

## 7.2 Outils d'inpainting local de la sphère

### 7.2.1 Problématique

Les zones du ciel dans lesquelles le signal n'a pas été observé ou n'a pas pu être estimé, ou celles pour lesquelles la contamination par des composantes astrophysiques autres que la composante étudiée est trop importante, doivent être masquées. Il n'est pas toujours possible de traiter simplement l'effet d'un masque dans l'analyse des données. Il existe deux méthodes principales. La première consiste à développer une méthode d'analyse capable de n'utiliser que les pixels valides, la seconde méthode implique de remplir l'ensemble des pixels non valides et faire l'analyse comme si on disposait d'une carte entière. Dans le second cas, il est nécessaire d'estimer l'erreur introduite par le masque, afin d'apporter les corrections nécessaires aux mesures. Suivant la méthode choisie pour remplir les pixels masqués et la géométrie du masque, il est parfois possible d'estimer l'erreur de façon analytique, dans le cas contraire, il est toujours possible d'utiliser des simulations Monte-Carlo.

Il est souvent difficile d'employer la première méthode, parce qu'il n'existe pas toujours d'algorithmes permettant de travailler sur un sous ensemble des pixels d'un part mais également parce que même dans les cas où des méthodes existent, elles sont souvent inapplicables aux cas de PLANCK, pour des raisons pratiques, à cause de la trop grande résolution des cartes qui nécessite généralement des ressources en terme de puissance de calcul et/ou de mémoire non accessible à l'heure actuelle.

Pour l'estimation de l'effet de lentille gravitationnelle, il existe un estimateur de maximum de vraisemblance développé par Hirata & Seljak (2003a) qu'on peut adapter au travail sur un sous

ensemble des pixels, ce qui permet de s'affranchir des effets d'un masque. Cependant, comme dans beaucoup de cas, il est difficilement envisageable de l'appliquer aux données de PLANCK car cela nécessite notamment d'inverser une matrice de 50 millions d'éléments de coté. Des tentatives sont toutefois en cours par des méthodes de gradients conjugués. Il existe une méthode d'estimation du spectre de la déflexion basée sur un estimateur quadratique. Cette méthode n'est pas applicable sur un sous ensemble incomplet de données. Dans la procédure d'estimation de l'effet de lentille gravitationnelle, le traitement du masque fait donc parti des effets majeurs à prendre en compte. De notre côté, nous avons alors choisi d'utiliser l'estimateur quadratique de Okamoto et Hu (voir section 6.6) couplé à une méthode permettant de remplir les pixels à masquer, avec une correction par Monte-Carlo.

Différentes méthodes ont été développées pour remplir un masque. L'opération la plus simple consiste certainement à remplacer les pixels masqués par une valeur nulle. C'est la technique utilisée par exemple par MASTER (Hivon et al. (2002)), pour estimer le spectre de puissance des anisotropies du CMB. Afin de minimiser les erreurs introduites par le masque, il est naturel de chercher à remplir les données manquantes par un signal le plus proche possible de l'original.

Ce n'est que récemment, suite aux travaux pionniers de Masnou & Morel (1998), que des méthodes efficaces, basées sur des développements innovants en traitement d'image, ont été développées. Ces techniques d'analyse d'image, visent à restaurer des portions manquantes d'un signal à partir des données connues. Elles permettent de répondre au problème précédant en imposant des contraintes fortes pour s'assurer d'influencer le moins possible les mesures. Le principe de base, consiste, par diverses méthodes, à reconstruire un signal dans le masque en forçant ce signal reconstruit à posséder des propriétés statistiques les plus proches possibles de celles du signal observé en dehors du masque, possédant également des propriétés de continuité convenables et ne modifiant pas les valeurs hors du masque. Ces méthodes portent le nom générique d'*inpainting* (qui signifie littéralement en anglais peindre à l'intérieur), qui est issu du nom utilisé par les conservateurs des musées qui restauraient des toiles, en peignant les parties abimées. Ce qualificatif pourrait être employé pour désigner l'ensemble des méthodes permettant de remplir un masque. Cependant pour des raisons historiques, seuls les algorithmes essayant de restaurer le signal à l'intérieur de régions masquées sont dénommés *inpainting*. Elles ont été développées dans un premier temps pour le traitement d'images planes, à deux dimensions, puis ont été étendues au cas de l'*inpainting* d'une sphère (voir 8.1.3).

Toutefois, lorsque le masque couvre une trop grande proportion du ciel, ce qui peut être le cas si on utilise un masque galactique restrictif, par exemple, ou encore si la couverture du ciel n'est pas complète, la convergence vers une solution optimale n'est plus toujours assurée. On peut alors parfois observer un biais pour l'estimation de l'effet de lentille gravitationnelle. De plus, tous les algorithmes d'*inpainting* de la sphère sont des méthodes globales prenant en compte l'ensemble des informations sur tout le ciel non masqué. Un des avantages de la méthode d'analyse des données par patches est d'étudier les propriétés du signal localement sur des portions restreintes du ciel afin de pouvoir mettre en évidence d'éventuelles variations avec la position. Dans cette optique, il est intéressant de pouvoir réaliser un *inpainting* en n'utilisant que les valeurs locales du patch.

Enfin, il a été montré<sup>1</sup> que si on ajoute un masque de sources compactes à un large masque galactique, l'*inpainting* utilisé sur la sphère (voir section 8.1.3) introduit un biais plus important sur la reconstruction de l'effet de lentille gravitationnelle. La morphologie d'un masque de sources compactes qui sont des sources ponctuelles ou quasi-ponctuelles réparties dans tout le ciel, est très différente de celle d'un masque des émissions galactiques qui sont des émissions diffuses couvrant de larges régions du ciel. La proportion de ciel couverte par un masque galactique est également, en général, bien supérieure à celle d'un masque de sources compactes. Ainsi on comprend que la solution au problème d'*inpainting* est dominée dans un tel cas par les contraintes imposées par les pixels situés dans le plan galactique et n'est pas nécessairement la meilleure solution pour des pixels du masque de sources compactes.

Pour s'affranchir de ces problèmes, nous avons développé une méthode d'*inpainting* local par patches. Elle permet d'*inpainter* une carte de tout le ciel, morceau par morceau, en n'utilisant que les informations locales. Elle est bien adaptée à l'*inpainting* des sources compactes qui sont présentes sur tout le ciel mais qui ne couvrent que des petites zones du ciel à la fois. Cette méthode possède toutefois ses limites et n'est pas adaptée à l'*inpainting* de larges zones comme le plan galactique. L'implémentation que nous avons réalisée tient compte de ces contraintes et laisse intactes les zones

1. Voir note interne <http://wiki.planck.fr/index.php/Lensing/TestMrsAlmInpainting2>

trop fortement masquées. Elle est par ailleurs très bien adaptée à une analyse par patch car dans ce cas, il est déjà possible de sélectionner des zones particulières du ciel uniquement et de ne pas prendre en compte les zones pour lesquelles la proportion de pixels masqués est trop importante. On peut diviser la procédure en trois étapes principales :

- découpe de la carte à inpainter en patches
- inpainting des patches
- inpainting de la sphère par reprojektion des patches sur la sphère

Dans la suite, nous décrivons les détails de chacune des étapes.

## 7.2.2 Découpe des patches

La première étape consiste à découper la sphère en patches. Les codes d'inpainting dans le plan travaillent sur des images à deux dimensions avec une pixelisation régulière. Nous utilisons donc l'outil développé dans la section 7.1 afin d'obtenir des patches, dont les pixels tombent sur une grille régulière, à partir des pixels de la sphère. Nous utilisons la méthode d'interpolation bilinéaire pour accélérer la procédure. La découpe doit se faire sans apodisation ni soustraction de la moyenne du patch, car les zones masquées devront être inpaintées puis reprojectées sur la sphère et la modulation de l'amplitude avec la position introduite par l'apodisation ou la soustraction d'un offset, produiraient alors de fortes discontinuités lors de la reprojektion.

Il est important de respecter certaines contraintes particulières pour le choix de la taille des patches, de leur nombre de pixels et de leurs centres. Dans le cas général, il faut essayer de choisir des patches suffisamment larges comparés à la taille des zones à inpainter, pour que l'algorithme d'inpainting puisse converger, tout en conservant une taille raisonnable pour que l'approximation plane soit correcte. Dans la pratique cela correspond à se limiter à des patches de taille inférieure à une vingtaine de degrés de coté. Nous avons également introduit un paramètre qui fixe un seuil, permettant d'éliminer les patches dont la proportion de pixels masqués est supérieure au seuil. Une fois la taille des patches choisie, ce paramètre permet donc de façon simple de s'affranchir des régions où l'inpainting risque de ne pas donner des solutions fiables.

Enfin, à partir des images inpaintées les zones reconstruites sont reprojectées sur la sphère. Comme les pixels HEALPIX ne coïncident pas avec ceux d'un patch, il faut interpoler la grille régulière au point correspondant au centre du pixel à inpainter sur la sphère. Pour cette raison, il faut impérativement choisir une grille suffisamment fine afin de ne pas introduire d'effets indésirables lors de la reprojektion. L'idéal étant d'obtenir une densité de pixels sur les patches comparable à celle de la zone de la sphère découpée. Dans le cas où la sphère va être utilisée dans la suite pour une analyse par patch, il est assez naturel d'utiliser les mêmes centres et la même taille pour l'inpainting que pour l'analyse qui suit mais ce n'est pas impératif et ce choix doit avant tout être motivé par les contraintes précédentes.

## 7.2.3 Inpainting local

Les patches découpés sont ensuite inpaintés. Pour cela, nous utilisons le code FASTLens<sup>2</sup> développé par Pires et al. (2009). Le problème de l'inpainting consiste à reconstruire des régions manquantes d'une image en exploitant au maximum les informations présentes dans les données valides. La méthode utilisée par FASTLens repose sur l'hypothèse qu'il est possible de décomposer le signal sur un dictionnaire dans lequel la représentation du signal possède une grande parcimonie et celle du masque une parcimonie beaucoup plus faible. Cette hypothèse est généralement favorisée par le fait qu'un signal physique et un masque possèdent des structures morphologiques très différentes et ne sont donc pas aussi facilement représentés sur un dictionnaire donné.

Un dictionnaire,  $\Phi$ , peut être vu comme une collection de fonctions, par exemple des sinusoides si on décompose le signal dans l'espace de Fourier, ou des ondelettes, des curvelettes des ridgelettes... ou encore l'union de plusieurs collections de fonctions. À la différence d'une base, un dictionnaire peut être surcomplet, c'est-à-dire que certains éléments du dictionnaire peuvent être décrits par une combinaison d'autres éléments. Une décomposition sur un dictionnaire n'est donc pas nécessairement unique.

2. Voir le site web <http://irfu.cea.fr/Ast/fastlens.software.php>

La parcimonie est liée au nombre de coefficients nécessaires à la représentation d'un signal sur un dictionnaire. Plus ce nombre est faible, plus le signal possède une grande parcimonie. Si on note  $\alpha_X^\Phi$ , les coefficients de la décomposition d'un signal  $X$  sur le dictionnaire  $\Phi$ , une mesure de la parcimonie d'un signal  $X$  peut donc être donnée par :

$$\|\alpha_X^\Phi\|_0 \quad (7.7)$$

où,  $\|z\|_0$  représente la pseudo norme  $l_0$ , c'est-à-dire le nombre d'éléments non nuls du vecteur  $z$ . Cette définition est difficilement applicable au cas de signaux réels, en présence de bruit notamment, car tous les coefficients de la décomposition peuvent être non nuls. Cependant, si la représentation possède une parcimonie suffisamment élevée, il a été montré dans Donoho & Huo (2001), que la pseudo-norme  $l_0$  pouvait être remplacée par la norme convexe  $l_1$ , c'est-à-dire la somme de la valeur absolue des coefficients de la décomposition (soit pour un vecteur  $z$ ,  $\|z\|_1 = \sum_k |z_k|$ ).

Si on considère un masque  $M$  associé au signal  $X$ , le problème de l' inpainting se résume à estimer  $X$  connaissant  $M$  et le signal masqué  $Y = MX$ . En suivant l'hypothèse énoncée précédemment qui suppose que le signal masqué possède une parcimonie plus grande que le signal d'origine, la solution au problème d' inpainting peut alors se poser sous la forme du problème de minimisation suivant :

$$\min_X \|\alpha_X^\Phi\|_1 \quad (7.8)$$

soumis à la contrainte

$$\|Y - MX\|^2 \leq \sigma \quad (7.9)$$

où la norme  $\|\cdot\|$  est la norme  $l_2$  usuelle ( $\|z\| = \sum_i z_i^2$ ) et  $\sigma$  représente la variance du bruit qui doit être donnée en paramètre. Si on utilise la norme  $l_1$  comme mesure de la parcimonie, le problème précédant peut être résolu par une méthode de seuillage itératif appelé MCA (voir Elad et al. (2005) ou également Pires et al. (2009) pour l'implémentation). Chaque itération  $n + 1$  peut s'écrire en fonction de l'itération précédente, sous la forme :

$$X^{n+1} = \Delta_{\Phi, \lambda_n}(X^n + M(Y - X^n)) \quad (7.10)$$

où l'opération non linéaire  $\Delta_{\Phi, \lambda_n}(Z)$  se décompose en trois étapes :

- décomposition de  $Z$  sur le dictionnaire  $\Phi$  pour obtenir les coefficients  $\alpha_Z^\Phi$
- effectuer un seuillage des coefficients de la façon suivante  $\tilde{\alpha}_Z^\Phi = \sin(\alpha_Z^\Phi) \max(0, |\alpha_Z^\Phi|) - \lambda$ . Le paramètre  $\lambda$  diminue à chaque itération, jusqu'à devenir nul. Les détails des différentes stratégies envisageables pour diminuer le seuil peuvent être trouvés dans Pires et al. (2009).
- reconstruction du signal  $\tilde{Z}$  à partir des coefficients après le seuillage

Le nombre d'itération est choisi en paramètre. La valeur initiale de  $X$  est choisie à  $Y$ . Dans le cas où on a du bruit dans les données, le seuillage peut-être arrêté en fonction du niveau de bruit. Cependant en général on cherche également à reconstruire le bruit dans le masque afin de conserver des propriétés statistiques proches de celles en dehors du masque. Le seuil est donc choisi à 0, ce qui équivaut à imposer une stricte égalité pour la contrainte 7.9 (voir Pires et al. (2009) pour plus de détails).

L'efficacité de l'algorithme dépend du choix de la représentation du signal qui doit posséder la plus grande parcimonie. Tous les dictionnaires ne sont donc pas équivalents. La difficulté de la méthode réside donc en partie dans le choix d'un dictionnaire répondant aux meilleures contraintes de parcimonie. La différence de parcimonie entre le signal masqué et non masqué est assurée dans la plus part des cas car le signal et le masque possèdent généralement des morphologies distinctes. En revanche, la parcimonie du signal non masqué peut dépendre elle fortement du dictionnaire et il est nécessaire de réaliser des tests avant de faire un choix. Comme l'estimation du signal reconstruit repose sur un processus itératif qui nécessite à chaque itération de décomposer le signal de l'itération précédente sur le dictionnaire puis d'effectuer la reconstruction dans l'espace direct du signal estimé à l'itération courante, en général, il est pratique de se limiter à des dictionnaires construits à partir de bases de fonctions pour lesquelles il existe des implémentations des décompositions directes et inverses rapides afin d'accélérer le processus itératif.

Il a été montré par Pires et al. (2009) que pour l'étude de l'effet de lentille gravitationnelle une décomposition sur une base de cosinus (aussi connue sous le nom de DCT pour le terme anglais Discrete Cosine Transform voir par exemple González & Woods (2002)) permet d'obtenir une représentation



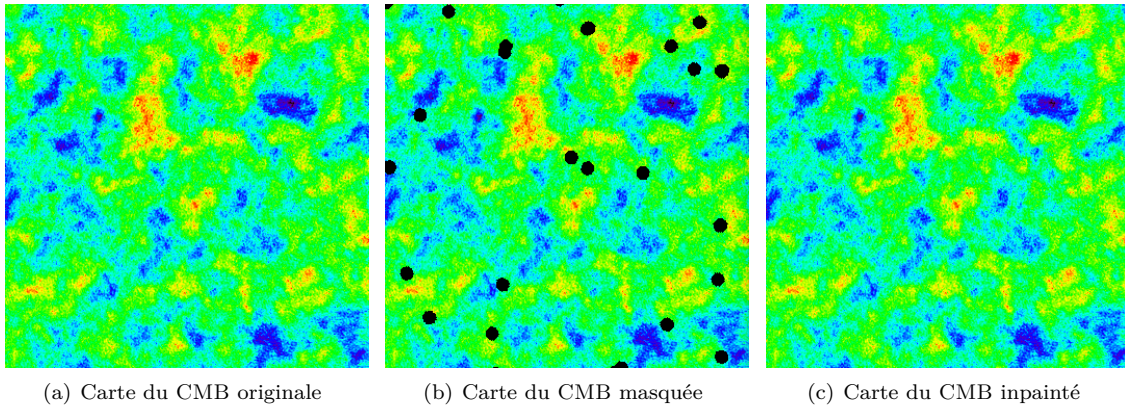


FIGURE 7.7 – **Inpainting par patch d’une carte de CMB.** La figure (a) montre une carte du CMB de 10 degrés de côté, possédant un bruit et un lobe comparables à ceux des canaux à 217GHz de l’instrument HFI. Les pixels masqués (en noir sur la figure (b)) sont ensuite inpaintés avec le programme FASTLens. Le masque est issu d’un patch de 10 degré de côté, découpé dans le masque de sources ponctuelles à 217GHz. Il n’est pas possible de distinguer la carte inpaintée (figure (c)) de l’originale à l’oeil. L’inpainting introduit cependant un faible effet systématique sur le spectre de deflexion qu’il est possible de corriger par Monte-Carlo.

avec une grande parcimonie. C’est donc l’approche que nous avons utilisée. On peut voir sur la figure 7.7 un exemple de carte du CMB lentillée et inpaintée ainsi que le masque correspondant. Sans connaître le masque il n’est pas possible de retrouver les zones masquées. Cependant, l’algorithme d’inpainting n’est pas capable de recréer un effet de lentille gravitationnelle dans les zones masquées. Le spectre estimé à partir de cartes inpaintées possède donc un biais lié à un manque global de puissance. Le manque de puissance est directement proportionnel à la surface masquée et peut donc être corrigé facilement.

## 7.2.4 Reprojection

L’étape finale de l’inpainting local par patch consiste à reconstruire le signal dans les régions masquées de la sphère à partir des patches inpaintés. Cette procédure peut être résumée de la façon suivante :

- parcourir l’ensemble des pixels de la sphère et déterminer s’ils doivent être inpaintés en testant la valeur des pixels correspondants du masque
- si un pixel doit être inpainté, chercher les patches dans lequel il tombe
- sélectionner le patch le plus adapté
- déterminer la position projetée du centre du pixel à inpainter, sur le patch sélectionné
- interpoler le signal sur le plan à cette position et affecter la valeur obtenue au pixel de la sphère à inpainter

Il faut noter que dans certains cas, suivant la forme et la taille du masque, il n’existe aucun patch remplissant les conditions nécessaires ce qui empêche d’inpainter certains pixels. Par ailleurs les pixels valides sont inchangés, seuls les pixels masqués sont modifiés.

L’étape la plus délicate est la sélection du meilleur patch à utiliser pour estimer le signal à inpainté d’un pixel donné. Pour cela, dans un premier temps nous recherchons l’ensemble des patches recouvrant le pixel considéré parmi les patches possédant une fraction de pixels masquée inférieure au seuil défini à l’étape précédente. Comme les patches peuvent se recouvrir, il y en a généralement plusieurs. Si aucun patch ne vérifie cette condition, il n’est pas possible d’inpainter le pixel, dans le cas contraire, il faut choisir le patch le plus adapté à la reconstruction du signal dans le pixel considéré. On pourrait être tenté de déterminer la valeur du pixel comme la moyenne des valeurs estimées avec chaque patch. Cependant, comme chaque patch couvre une zone du ciel différente, la solution vers laquelle converge l’inpainting est différente dans chaque cas. En effet la solution du problème dans chaque cas ne représente pas une solution intrinsèque. Elle peut être vue comme une réalisation particulière satisfaisant les conditions imposées, qui changent pour chaque patch. En

d'autres mots, la valeur reconstruite n'a aucune raison d'être proche de la valeur réelle, elle constitue simplement la solution du problème imposé avec des conditions spécifiques. Ainsi, nous choisissons parmi les patches retenus, celui dans lequel le pixel à inpainter tombe le plus près du centre. Cela assure que les distortions provoquées par la projection sont les plus faibles possible.

On voit que cette méthode d'inpainting a des limites. Lorsqu'il n'existe pas de patch recouvrant un pixel et possédant une fraction masquée suffisamment faible, il n'est pas possible d'inpainter le pixel. L'inpainting local par patch n'est donc pas adapté à la restauration de larges zones contiguës. Cependant, ses performances sont très bonnes pour inpainter les masques de sources ponctuelles. Couplé à une étude par patch, il permet de réaliser une analyse en s'assurant que la statistique locale n'est pas influencée par celle de régions du ciel éloignées.

## 7.3 Validation

### 7.3.1 Algorithme d'approximation par série de Fourier

Afin de s'assurer de la validité du code d'estimation de fonction par série de Fourier, que nous utilisons pour évaluer le signal dans un patch à partir des pixels de la sphère projetés, nous avons comparé ses résultats dans des cas simples, à ceux d'une transformée de Fourier discrète ordinaire sur une grille régulière. Pour cela nous partons d'une image gaussienne à deux dimensions avec un spectre en  $k^{-2}$  (proche d'un spectre de CMB). Le signal est généré sur une grille régulière, de sorte qu'il est possible d'estimer sa transformée de Fourier discrète par FFT. Nous utilisons alors l'outil général d'estimation de signal par ajustement de séries de Fourier que nous avons développé (voir appendice B). Nous l'utilisons habituellement sur une grille irrégulière mais il est adapté à toutes répartitions des pixels. Nous l'utilisons donc sur l'image régulière précédente, pour effectuer la comparaison à une transformée de Fourier discrète.

L'estimation des coefficients de Fourier est effectuée sur la même grille que la celle de la transformée de Fourier discrète. Sur la figure 7.8 on peut observer la comparaison des parties réelles et imaginaires des coefficients estimés par les deux approches. À partir des coefficients de Fourier estimés, il est possible de faire une transformation inverse afin d'obtenir la représentation du signal dans l'espace réel, ce qui permet de comparer également le signal reconstruit à l'original (voir figure 7.9). On observe dans les deux cas que les performances sont remarquablement bonnes. En effet, l'amplitude des différences est de l'ordre de  $10^{-15}$ , ce qui correspond à la précision des erreurs numériques pour des calculs sur des nombres à virgule flottante de double précision.

Dans le test précédent, nous partons d'un signal idéal sur une grille régulière. Nous avons estimé autant de coefficients de Fourier indépendants que de nombre dans l'espace réel. Dans un second temps il est intéressant de tester l'algorithme dans des conditions comparables à celles que nous allons utiliser et de chercher ses limites. En particulier, nous avons essayé de voir l'influence du choix du nombre de coefficients de Fourier à estimer. Pour cela nous partons d'une carte de CMB HEALPIX, de paramètre *nside* égal à 2048, avec un bruit et un lobe instrumental comparables à ceux des données à 217GHz. On découpe ensuite un ensemble de patches carrés de 10 degrés de coté, en faisant varier la taille de la grille de coefficients, de 112 par 112 à 350 par 350. La limite inférieure de l'intervalle de variation est dictée par la taille minimale nécessaire afin d'obtenir des fréquences jusqu'à  $\ell_{max} = 2000$  qui sont celles qui nous intéressent pour notre étude. La limite supérieure correspond approximativement au cas d'une transformée de Fourier discrète sur une grille régulière. En effet, étant donnée la taille des patches, on dispose d'environ  $350^2$  pixels de la sphère HEALPIX pour estimer les coefficients. Par analogie au cas régulier, on se limite donc à une grille de taille maximale identique au nombre de pixels dans l'espace réel.

Tout d'abord, on observe que l'algorithme commence à ne plus être capable de converger vers une solution, dans un temps fini, pour des grilles de taille supérieure à environ 260 par 260. C'est un problème technique qui peut être résolu en utilisant un préconditionnement afin de s'approcher de la solution et de faciliter la convergence (on pourrait utiliser par exemple la valeur de la transformée de Fourier discrète calculée sur le patch découpé avec une interpolation). Nous n'avons donc obtenu que des résultats pour ces tailles de grille.

Finalement, on cherche à estimer la dispersion des modules et des phases des coefficients de Fourier estimés sur des grilles de tailles différentes. Pour cela, nous conservons uniquement la zone commune à chaque estimation qui correspond donc à la zone centrée sur la fréquence nulle de taille

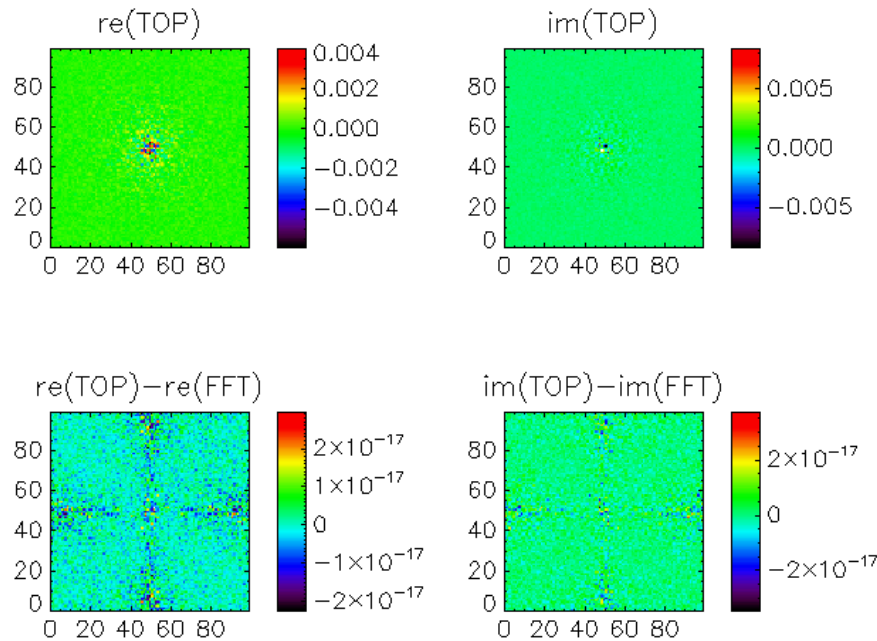


FIGURE 7.8 – **Comparaison dans l'espace de Fourier des estimations par FFT et par ajustement de séries de Fourier.** On compare la partie imaginaire et réelle des coefficients de Fourier obtenus avec une transformée de Fourier discrète, calculée par FFT, à ceux obtenus par ajustement de séries de Fourier, en partant d'une même image possédant des pixels réguliers. L'accord est excellent, l'amplitude de la différence est de  $10^{-15}$  et correspond au niveau des erreurs numériques attendues.

112 par 112. On estime finalement, le module et la phase de chaque coefficient et on détermine la dispersion entre les différentes estimations. Les résultats sont présentés sur la figure 7.10. On peut observer que l'erreur absolue sur la phase et le module est assez uniforme. L'erreur relative varie elle en fonction du module de la fréquence mais reste cependant toujours inférieure au pourcent. On en déduit que la taille de la grille sur laquelle on estime les coefficients de Fourier n'influence pas la qualité de l'estimation.

### 7.3.2 Méthode d'estimation du signal projeté sur une grille régulière

Nous avons développé trois méthodes afin d'estimer le signal, issu des pixels de la sphère projetés sur un patch, qui sont répartis sur une grille irrégulière (voir section 7.1.4). Il est nécessaire d'utiliser la meilleure des trois méthodes pour réaliser l'analyse de l'effet de lentille gravitationnelle. Pour cela nous réalisons une étude complète sur une carte du ciel de CMB, simulée à partir du spectre WMAP, avec un bruit et un lobe instrumental correspondant à ceux de la carte à 217GHz du détecteur HFI. Nous découpons, avec les trois méthodes différentes, un ensemble de 768 patchs carrés de 10 degrés de coté, centrés sur les centres des pixels HEALPIX de paramètre  $n_{side}$  égal à 8 et apodisés avec une fenêtre de Hann. Sur la figure 7.11 on peut voir un exemple de patch obtenu avec les trois méthodes différentes.

Sur la figure 7.12, nous comparons l'estimation du spectre de puissance des anisotropies de température, effectuée à l'aide d'une analyse par patch avec les trois méthodes de découpe que nous avons développées. La mesure est effectuée sur l'ensemble des patchs de la sphère afin d'avoir une bonne statistique. On peut noter que les méthodes de découpe utilisant une interpolation bilinéaire ou par plus proche voisin introduisent un lissage des données, ce qui produit un biais dans l'estimation du spectre de puissance de température. Le biais n'est pas présent sur le spectre estimé à partir des cartes découpées avec la méthode d'ajustement de séries de Fourier. Ce test montre l'avantage d'utiliser une méthode optimisée comme celle que nous avons développée.

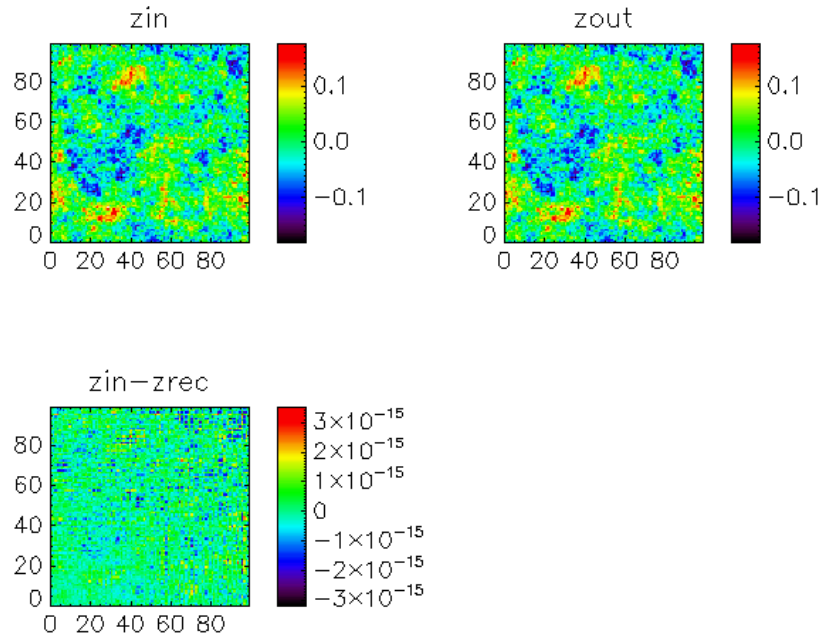


FIGURE 7.9 – **Comparaison dans l’espace réel des estimations par FFT et par ajustement de séries de Fourier.** On compare l’image reconstruite par transformée inverse, à partir des coefficients de Fourier estimés par la méthode par ajustement de séries de Fourier à l’image initiale. La différence obtenue est négligeable et correspond au niveau des erreurs numériques attendues pour des calculs en virgule flottante de double précision.

Les trois méthodes ne sont pas équivalentes. Celles utilisant une interpolation, permettent une découpe rapide. En effet, il faut moins d’une seconde pour découper un patch carré de 10 de degrés de côté sur une carte HEALPIX de paramètre  $n_{side}$  égal à 2048. Cependant elles ne sont pas adaptées aux études que nous voulons réaliser. À la vue de ces résultats, nous avons décidé d’utiliser la méthode de découpe par approximation de fonction par ajustement de série de Fourier pour l’ensemble de notre étude. Elle est moins efficace en terme de temps de calcul (il faut une trentaine de secondes pour découper des cartes de 10 degrés de côté dans une carte HEALPIX de  $n_{side}$  égal à 2048), mais elle conduit à une estimation plus juste du signal sous-jacent.

## 7.4 Optimisations

Dans les sections précédentes nous avons présenté un ensemble de procédures visant à estimer l’effet de lentille gravitationnelle. Elles dépendent d’un ensemble de paramètres qui sont laissés libres dans les discussions précédentes. Il est toutefois nécessaire d’adapter leur valeur à chaque problème particulier et un choix judicieux influence grandement la qualité des mesures.

### 7.4.1 Borne de l’intégrale pour l’estimateur du potentiel de lentille

On a vu dans la section 6.6 que l’estimateur de la déflexion était basé sur l’intégrale dans l’espace de Fourier d’un produit quadratique de modes de la carte des anisotropies de température du CMB, associé à un filtrage en fréquence. Cette intégrale doit théoriquement être estimée jusqu’à l’infini. Cependant, la normalisation de l’estimateur, qui est également une intégrale de la même forme que le terme principal, compense les effets de la coupure. Il est ainsi possible de fixer une borne supérieure d’intégration de valeur finie. Dans la pratique, la borne supérieure est donc un paramètre de l’estimateur.

Nous travaillons sur des cartes déconvoluées, sans soustraction du bruit. La puissance du signal

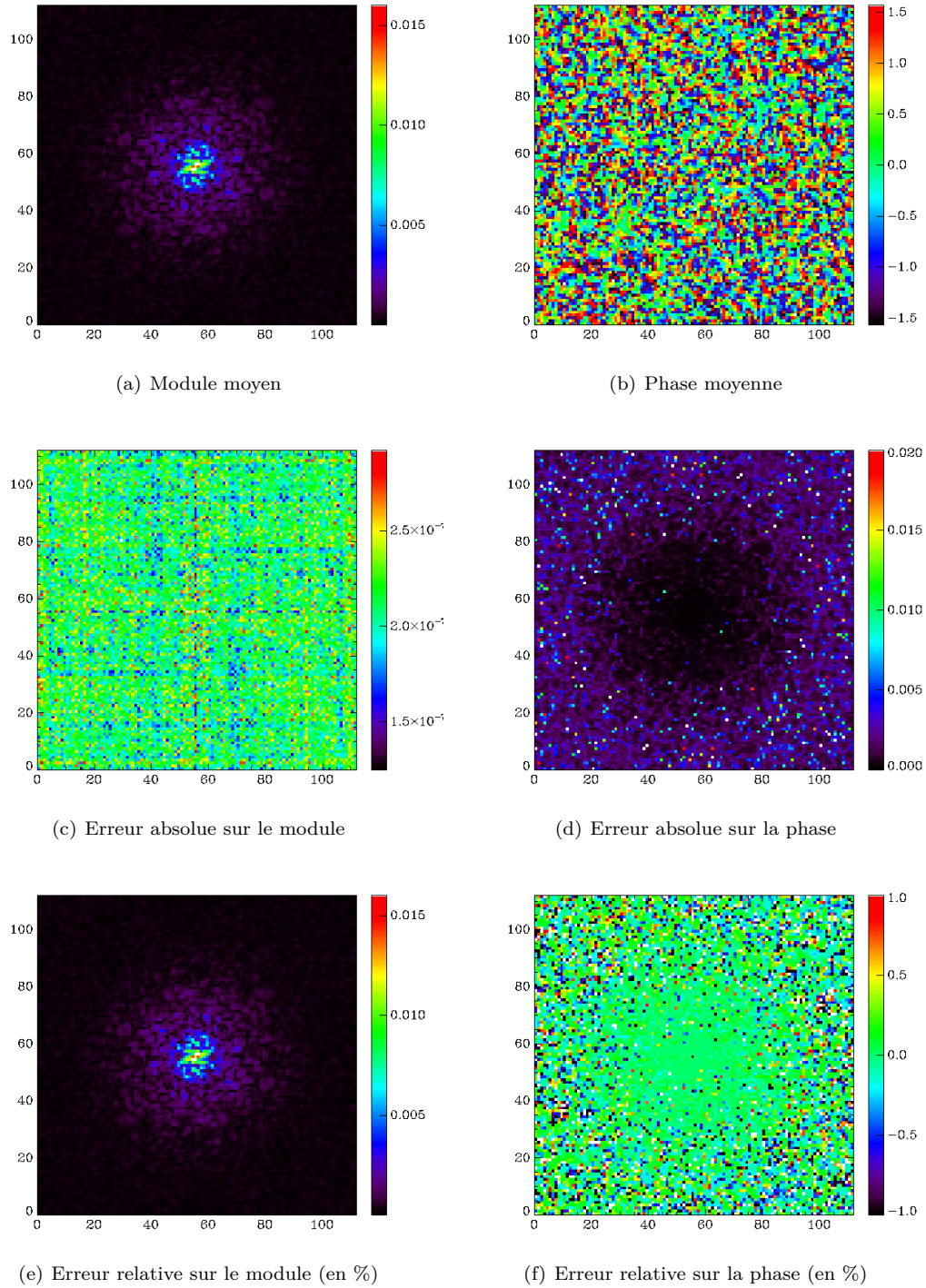
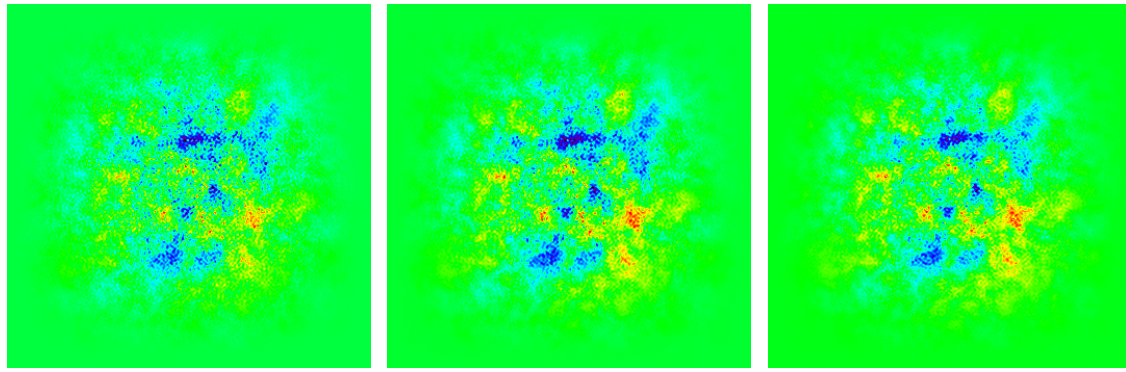
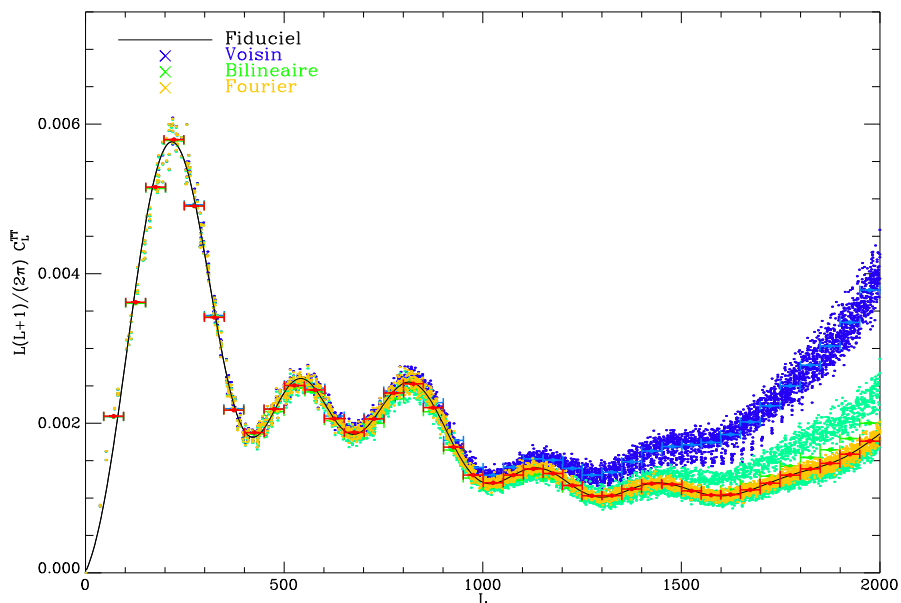


FIGURE 7.10 – **Dépendance de la taille de la grille, de la méthode d'estimation par série de Fourier.** En partant d'une carte du ciel du CMB, avec un lobe et un bruit instrumental comparables à ceux des données de HFI à 217GHz, un ensemble de patches ont été découpés avec la méthode d'approximation par séries de Fourier, afin de tester l'influence de la taille de la grille de Fourier sur la détermination des coefficients. L'estimation a été menée sur des tailles de grilles différentes en partant du même jeu de données. La dispersion sur le module et la phase est alors estimée et comparée à la valeur moyenne des estimations pour l'ensemble des coefficients en commun sur les différentes grilles. L'erreur absolue sur la phase et le module est uniforme et l'erreur relative ne dépasse pas 1%.



(a) Découpe par interpolation par plus proche voisin (b) Découpe par interpolation bilinéaire (c) Découpe par estimation de fonction par ajustement de coefficients de Fourier

**FIGURE 7.11 – Comparaison des méthodes de découpe des patches.** Lorsqu'on projette une portion du ciel sur le plan, les centres des pixels HEALPIX sont irrégulièrement répartis. Afin d'obtenir un patch sur une grille régulière nous avons développé trois méthodes dans la section 7.1.4. Deux d'entre elles utilisent des interpolations (par plus proche voisin ou bilinéaire), la dernière se base sur un ajustement global du signal de l'ensemble des pixels projetés, sur des séries de Fourier. La même région du ciel a été découpée avec les trois méthodes. Avec la méthode par séries de Fourier, on obtient une représentation dans l'espace de Fourier de la fonction et on a utilisé une transformée de Fourier inverse pour obtenir la représentation dans l'espace réel et pouvoir la comparer aux autres. La structure générale est dominée par l'apodisation par la fonction fenêtre de Hann. Les trois méthodes donnent des résultats similaires si on compare simplement les carte dans l'espace réel à l'oeil. Il faut comparer les spectres de puissance des anisotropies de température et de la déflexion pour détecter des différences notables.



**FIGURE 7.12 – Comparaison des spectres de puissance de température pour les différentes méthodes de découpe.** On observe que les méthodes par apodisation introduisent un lissage des données, ce qui se traduit par une perte de puissance notable au delà de  $\ell$  égal à 1000. La méthode la plus robuste est celle par ajustement de séries de Fourier.

total augmente très fortement à haut  $\ell$ , à cause de la déconvolution du bruit, comme expliqué dans la section 6.4.2. Au-delà d'une certaine échelle, la mesure est alors totalement dominée par le bruit. Il est donc raisonnable de couper l'intégrale à partir de cette valeur. Par ailleurs, le nombre de terme à évaluer pour le calcul de l'intégrale croit linéairement avec la valeur maximale de la borne. Le

temps de calcul dépend directement de la valeur de la borne supérieure.

Nous voulons évaluer le spectre de puissance de la déflexion jusqu'à  $\ell$  égal à 2000 car au delà, il est dominé par le bruit. Afin de limiter le temps de calcul, dans la suite, sauf mention du contraire, nous utilisons une borne supérieure pour l'intégrale, de 2000.

### 7.4.2 Taille des patchs

La taille des patchs est un paramètre important qu'il est nécessaire d'optimiser pour la recherche de l'effet de lentille gravitationnelle. En effet, notre étude est basée sur une estimation dans le plan. L'approximation plane est bien vérifiée uniquement jusqu'à une échelle d'une vingtaine de degrés, au-delà de laquelle, la courbure de la sphère devient trop importante pour être négligée. Les distortions introduites par la projection de la sphère sur un plan augmentent également avec la distance au point tangent de la projection (voir section 7.1.3), ce qui nécessite de se limiter à des tailles raisonnables et fonctions du problème.

Par ailleurs, l'effet de lentille gravitationnelle corrèle le signal sur des échelles de plusieurs degrés. Ils est donc indispensable d'utiliser des patchs suffisamment grands pour pouvoir bien prendre en compte ces effets. De plus, l'estimation du potentiel de lentille se fait dans l'espace de Fourier et la résolution de l'échantillonnage de cet espace dépend également directement de la taille des patchs. Un patch plus grand permet également d'utiliser un plus grand nombre de pixels pour contraindre la méthode d'interpolation par ajustement des coefficients de Fourier (voir section 7.2.2). Ce dernier point est toutefois à nuancer. En effet, plus le nombre de pixels devient grand, plus le temps de calcul pour que la méthode converge devient grand. Il faut donc trouver un juste milieu, afin d'obtenir un temps d'exécution compatible avec le grand nombre de simulations que nous devons réaliser.

Comme on l'a vu dans la section 7.1.2, le pavage de la sphère par des patchs dépend lui aussi, fortement de leur taille et constitue une contrainte supplémentaire à prendre en compte dans le choix de la taille des patchs. Pour une répartition des centres fixée, plus la taille augmente plus le recouvrement est important. Pour estimer le spectre de puissance des anisotropies de température, on a vu dans la section 6.5.2 qu'il est intéressant d'avoir un recouvrement d'environ 25% pour estimer le spectre de puissance par patch, par analogie avec l'estimation d'un spectre de puissance à une dimension par périodogramme de Welsh. La figure 7.1(b) montre le recouvrement obtenu avec des patchs carrés de 10 degrés de côté, répartis sur les centres des pixels d'une carte HEALPIX de paramètre *nside* égal à 8. Cette configuration permet d'obtenir un recouvrement très proche de 25%. Si on augmente la taille des patchs il faut donc changer la répartition des centres et diminuer leur densité.

Enfin, lorsque la taille du patch augmente, l'effet de l'apodisation diminue. En effet, la multiplication par une fenêtre d'apodisation dans l'espace réel correspond à une convolution dans l'espace de Fourier. La convolution va corrélérer les modes sur une distance correspondant à la taille caractéristique de la fenêtre d'apodisation dans l'espace de Fourier. Cette échelle est inversement proportionnelle à la taille de la fenêtre d'apodisation dans l'espace réel. Ainsi, l'augmentation de la taille de la fenêtre d'apodisation diminue l'influence de l'apodisation. On peut voir sur la figure 7.13 que l'effet de l'apodisation sur le spectre de puissance de la déflexion se manifeste principalement à bas  $\ell$  et dépend fortement de la taille des patchs. À partir d'une taille de 10 degrés, on voit que les effets sont faibles au-delà de  $\ell$  égal à 100.

Pour conclure, il est nécessaire de trouver un compromis entre ces différentes contraintes. Nous avons donc cherché à utiliser une taille de patchs la plus grande possible afin de minimiser les effets de l'apodisation et d'avoir la meilleure contrainte sur les coefficients de Fourier, tout en se limitant à des échelles inférieures à une vingtaine de degrés, à un nombre de pixels raisonnable afin d'obtenir des temps de calcul compatibles avec un grand nombre de simulations et en cherchant à obtenir une stratégie simple de répartition des centres des patchs donnant un recouvrement correct. Au vu des résultats de la figure 7.13 et étant donnée la simplicité de la répartition des patchs sur les centres de pixels HEALPIX, nous avons choisi d'utiliser des patchs carrés de 10°de côté. Comme on peut le voir dans la figure 7.1(b) cette configuration optimise également le chevauchement des patchs.

### 7.4.3 Taille de la grille de coefficients à estimer

On a vu dans la section 7.3.2 que l'estimation du signal, à l'aide de la méthode basée sur l'ajustement des coefficients de Fourier sur une grille régulière donne les meilleurs résultats. Le seul

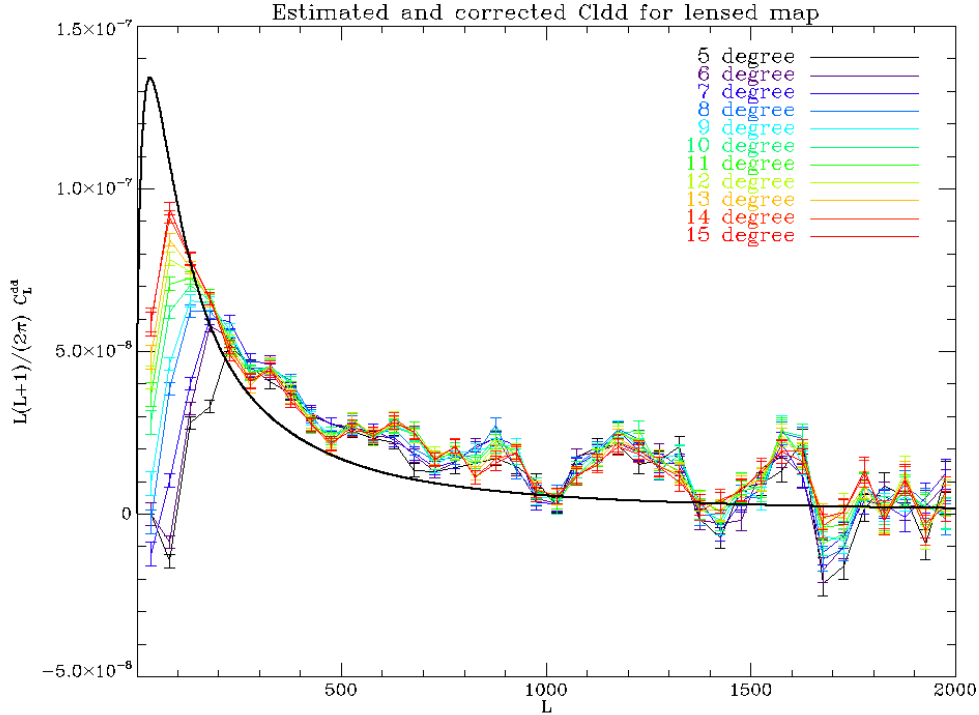


FIGURE 7.13 – **Influence de la taille des patches sur le spectre de la déflexion.** On observe un effet systématique important qui dépend fortement de la taille du patch. Pour une taille supérieure à 10 degrés l’effet est visible uniquement en dessous de  $\ell$  égal à 100. Les oscillations résiduelles visible à plus haut  $\ell$  sont les biais intrinsèques dûs à l’estimateur qui ne sont pas corrigés dans ce cas mais qui sont identiques pour les 11 tailles de patches (voir section 6.6.3).

paramètre libre de cette méthode est la taille de la grille des coefficients à estimer.

Il est nécessaire d’optimiser ce paramètre afin d’obtenir les meilleures performances. La densité moyenne de pixels de la sphère tombant dans des patches de taille donnée, est très constante car les pixels HEALPIX ont tous la même surface. Pour des patches de 10 degrés de coté, extraits de cartes HEALPIX de paramètre *nside* égal à 2048 (qui est la valeur pour les cartes de HFI), on mesure une densité très proche de  $350^2$  pixels par patch, ce qui fixe une limite supérieure pour la grille de Fourier. En effet, il n’est pas possible d’estimer plus de coefficients que le nombre de pixels du patch. Dans l’espace de Fourier à deux dimensions, le pas de la grille des vecteurs  $k$ , est uniquement fixé par la taille du patch, indépendamment du nombre de pixels de la carte dans l’espace réel. Pour une taille de  $L_x$  par  $L_y$ , on obtient des pas en fréquence respectifs de  $\Delta f_x = 2\pi/L_x$  et  $\Delta f_y = 2\pi/L_y$  dans les directions  $x$  et  $y$ . Cette propriété fixe une taille minimale de la grille nécessaire pour obtenir la puissance jusqu’à une valeur donnée. La formule générale, pour la taille minimale de la grille dans chaque direction, est alors donnée par :

$$s = 2 \lceil \frac{\ell_{max}}{(2\pi/L)} \rceil \quad (7.11)$$

où l’opérateur  $\lceil \cdot \rceil$  correspond à l’arrondi à l’entier supérieur. Le facteur 2 provient des propriétés de symétrie des fréquences dans l’espace de Fourier (on a autant de fréquences négatives que positives).

Nos tests ont montré que la méthode d’estimation des coefficients de Fourier par ajustement que nous avons développée, est très robuste. Les variations sur les coefficients de Fourier estimés sur des grilles dans l’espace de Fourier de tailles différentes, à partir du même jeu de données dans l’espace réel, sont inférieures au pourcent (voir 7.4.3). On peut donc estimer les coefficients de Fourier sur la grille dans l’espace de Fourier, la plus petite possible, pour une valeur maximale en fréquence donnée, ce qui permet d’optimiser les temps de calcul qui dépendent eux fortement du nombre de coefficients à estimer.

Pour des patches carrés, de 10 degré de coté, on obtient un pas pour la grille des vecteurs dans l’espace de Fourier de  $\Delta f_x = \Delta f_y = 2\pi/L_x = 36$ . Nous avons choisi d’estimer le spectre de puissance



jusqu'à  $\ell$  égal à 2000, car avec les niveaux de bruit de PLANCK, au delà de cette valeur le signal de déflexion est totalement dominé par le bruit. Cette limite donne une taille minimale de 112 par 112 pour la grille des coefficients de Fourier que nous avons arrondi à 120. Il faut garder à l'esprit que dans l'espace réel les cartes sont échantillonnées à la résolution d'une carte HEALPIX de paramètre  $n_{side}$  égal à 2048, ce qui correspond à environ 350 pixels carrés. La grille de 120 fixe juste la taille de la grille dans l'espace de Fourier sur laquelle nous estimons les coefficients. On a vu que la valeur de ces coefficients dépendait très peu de la taille de la grille dans l'espace de Fourier choisie (voir 7.4.3). On a donc utilisé la grille dans l'espace de Fourier de taille minimale étant donné que nous n'utilisons pas l'information au-delà de  $\ell$  égal à 2000. Sauf mention du contraire, nous utilisons donc dans la suite une grille de 120 par 120. Finalement, pour des patches carrés de 10 degrés de coté, découpés dans une carte HEALPIX de paramètre  $n_{side}$  égal à 2048, on constate que le nombre de pixels contenus dans chaque patch, permet de réaliser l'ajustement des coefficients de Fourier pour obtenir un signal sur une grille régulière carré de 120 éléments de coté, en une trentaine de secondes.

#### 7.4.4 Choix de la fenêtre d'apodisation

Une solution pour le problème rencontré dans la section 7.1.5 consiste à multiplier les patches avant l'estimation des coefficients de Fourier, par une fonction fenêtre qui présente des lobes secondaires d'amplitude moins importante dans l'espace de Fourier. Cependant, le choix d'une fenêtre d'apodisation résulte toujours d'un compromis entre la taille du lobe primaire et celle des lobes secondaires. En effet, l'énergie totale de la fenêtre est fixée en vertu du théorème de Parseval. On a choisi d'étudier la famille des fonctions de Kaiser définie par :

$$W_{\alpha}(x) = \frac{I_0(\pi\alpha\sqrt{1 - (\frac{x}{L/2})^2})}{I_0(\pi\alpha)} \quad (7.12)$$

$$W_{\alpha}(x, y) = W_{\alpha}(x)W_{\alpha}(y) \quad (7.13)$$

car elles sont analytiques et possèdent un paramètre  $\alpha$  qui permet de faire varier facilement le rapport entre la puissance du lobe principal et des lobes lointains Oppenheim & Schafer (1975). On a représenté sur la figure 7.14 leur forme (à une dimension) dans l'espace direct et réciproque, pour  $\alpha = 0.5$  et 1, pour un patch carré de  $10^{\circ}$  de coté. On compare l'allure de ces fenêtres à celle d'une fenêtre « tophat ». On voit que le paramètre  $\alpha$  permet d'atténuer l'amplitude des lobes lointains ce qui élargit le lobe principal en contre partie.

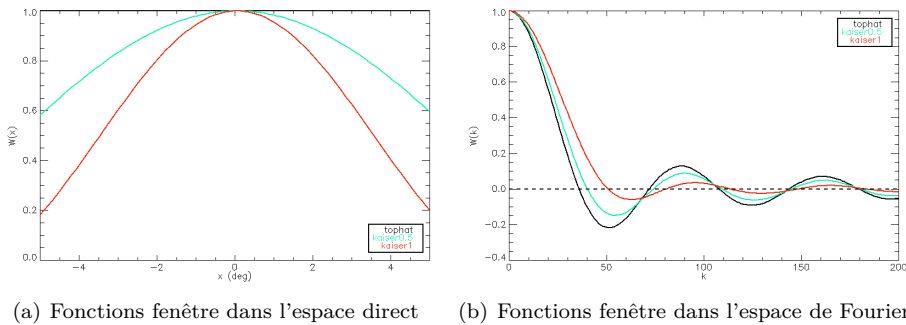


FIGURE 7.14 – Comparaison des fonctions fenêtre de Kaiser pour  $\alpha = 0.5$  et 1 et de la fonction « tophat ». Sur la figure (a) on a représenté les fonctions dans l'espace direct et sur la figure (b) dans l'espace de Fourier, pour un patch de taille  $L = 10^{\circ}$  de coté. Le calcul de l'intégrale de Fourier est fait numériquement dans le cas des fonctions de Kaiser et vaut  $\text{sinc}(kL/2)$  dans le cas « tophat ».

Sur la figure 7.15 on peut voir l'estimation du spectre de puissance obtenu dans les mêmes conditions que précédemment en multipliant cette fois chaque patch par une fenêtre de Kaiser bidimensionnelle. On constate ainsi que  $\alpha = 0.5$  n'est pas suffisant pour faire disparaître la croix des axes mais qu'on obtient un bon résultat pour  $\alpha = 1$ . On notera qu'un paramètre  $\alpha$  plus élevé donne également de bons résultats mais agrandit en contre partie le lobe principal, ce qui n'est pas

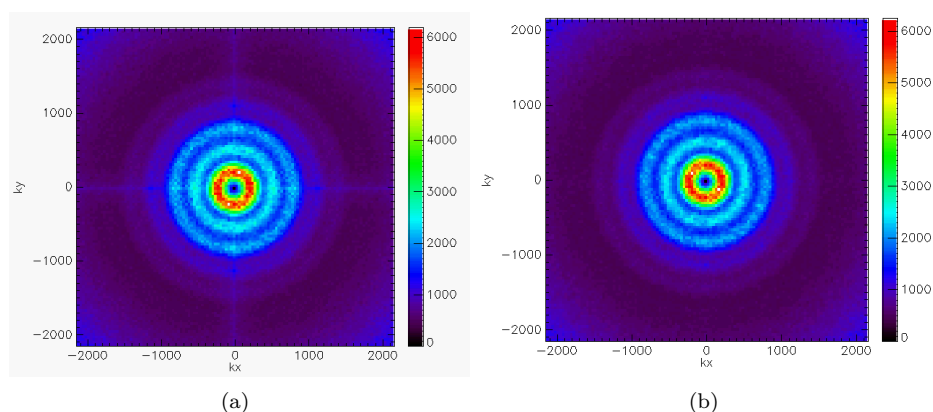


FIGURE 7.15 – **Influence du paramètre  $\alpha$  de la fonction de Kaiser sur le spectre de puissance.** Le spectre de puissance bi-dimensionnel est reconstruit à partir d’une carte CMB par la méthode par patches utilisant une fenêtre d’apodisation de Kaiser, de paramètre  $\alpha = 0.5$  (a) et  $\alpha = 1$  (b). Les axes  $x$  et  $y$  sont en unités de mode de Fourier  $\mathbf{k}$ . Le niveau de couleur représente  $\frac{|\mathbf{k}|^2}{2\pi} C_{\mathbf{k}}$ . Les unités sont des  $\mu K^2$

désirable en particulier pour l’estimateur de l’effet de lentille gravitationnelle qui est basé sur un produit de modes de Fourier et va donc être sensible à l’intégrale de recouvrement entre la puissance des fonctions fenêtre de la Figure 7.14(b). Plus précisément l’effet de l’apodisation sera d’introduire un effet fort dans la région  $\ell \lesssim 100$ , ce qui ne pose pas trop de problème étant donné qu’on travaille dans le cadre de l’approximation plane qui n’est à peu près valide qu’au dessus de cette valeur.

Dans la suite du manuscrit, l’apodisation fera donc toujours référence à la fenêtre de Kaiser définie par l’équation (7.13) avec un paramètre  $\alpha = 1$ .

## 7.5 Traitement des erreurs

### 7.5.1 Méthode Monte-Carlo

La méthode Monte-Carlo, est une technique statistique d’estimation et de correction des erreurs. Elle consiste à simuler un grand nombre de réalisations d’une expérience, reproduisant le mieux possible les données à étudier et les effets instrumentaux associés à la mesure, puis à appliquer, sur chacune des réalisations, la procédure d’analyse pour laquelle on cherche à estimer les erreurs.

En générant un grand nombre de simulations, il est possible d’explorer les distributions des processus aléatoires. En analysant la dispersion entre les différentes réalisations on a alors accès aux erreurs statistiques. Par ailleurs, si on moyenne l’ensemble des réalisations, on peut faire diminuer les erreurs statistiques. En comparant le signal obtenu au signal attendu on a alors accès à une mesure des effets systématiques produits, soit par l’instrument, soit par les méthodes d’analyse.

Les simulations Monte-Carlo demandent en général de lourdes ressources en calcul car elle nécessitent de simuler la totalité de la chaîne d’analyse et de réaliser les opérations suffisamment de fois. Afin de pouvoir distinguer des effets systématiques différents il faut en général générer un jeu de simulations particulier pour chacun des effets. C’est une méthode puissante qui est souvent la seule à pouvoir estimer des erreurs quand aucun modèle théorique n’est disponible. C’est également un moyen efficace de confronter des prédictions.

### 7.5.2 Estimation des erreurs par Monte-Carlo

Dans l’ensemble des analyses que nous avons développées, les barres d’erreurs correspondent aux erreurs statistiques. Elles sont estimées sur des jeux de simulations Monte-Carlo. Pour réaliser les simulations nous nous basons sur les méthodes et les effets décrits dans la section 6. Les simulations ont été générées au centre de calcul de l’in2p3, le « CCIN2P3 ». Elles représentent plusieurs dizaines de milliers d’heures de calculs et occupent plusieurs Téraoctets de données.

### 7.5.3 Correction des erreurs systématiques par Monte-Carlo

Il est possible d'estimer des corrections pour les effets systématiques à l'aide de simulation Monte-Carlo. La précision des corrections dépend de la précision de prise en compte des différents effets dans les simulations. Pour la mesure du spectre de la déflexion, les effets systématiques dominants sont : l'apodisation, le bruit inhomogène, le masquage des données et l'inpainting.

L'avantage de cette technique, est que le signal à mesurer est connu par avance. On peut alors comparer les mesures effectuées sur les simulations, aux valeurs utilisées et en déduire une correction. Le fait de générer un grand nombre de simulations permet de s'affranchir des variations statistiques de la mesure et de déterminer la correction correspondant aux effets systématiques uniquement. C'est cette correction qui est finalement utilisée pour corriger les données réelles. Pour cela il faut s'assurer d'employer la même procédure sur les données réelles que sur les simulations.

On peut envisager deux méthodes différentes pour estimer une correction. La première consiste à partir d'un jeu de simulations sans signal (c'est-à-dire, dans le cas qui nous intéresse des simulations sans effet de lentille gravitationnelle), d'appliquer l'analyse complète et de mesurer l'éventuel signal résiduel. Si la procédure de prise de données et d'analyse n'a aucune influence sur la mesure, cette analyse doit conduire à un signal nul. Ce n'est en général pas le cas et c'est le but de cette correction. On appellera cette hypothèse l'hypothèse H0. Elle consiste donc à partir d'un signal nul puis d'estimer le signal produit par la procédure afin de le soustraire des données. La seconde méthode commence par supposer que le signal est connu et à simuler des réalisations incluant le signal. En comparant le signal mesuré au signal en entrée, on en déduit alors la correction. Cette seconde hypothèse est appelée l'hypothèse H1.

La correction déduite en faisant l'hypothèse H0 suppose que l'effet produit par la chaîne d'analyse en l'absence de signal est le même que celui en présence d'un signal, ce qui correspond à faire l'hypothèse que les effets systématiques ne dépendent pas du signal. Or, ce n'est pas nécessairement le cas et il peut parfois être nécessaire de se baser sur l'hypothèse H1. Cependant, cette dernière suppose de bien connaître le signal à mesurer par avance. Il faut par ailleurs, dans ce cas, tenir compte de l'erreur introduite par l'erreur attendue sur la modélisation du signal, ce qui ajoute une étape de plus dans l'analyse. Finalement, la correction par hypothèse H0 donnant de bons résultats sur les simulations (voir section 8.7.3) et permettant de ne pas faire d'hypothèse sur le signal recherché, c'est cette méthode que nous avons employée pour l'analyse des données de PLANCK.

## Chapitre 8

# Estimation de l'effet de lentille gravitationnelle dans les données de HFI

Dans le chapitre 7, nous avons développé un ensemble de programmes permettant de mesurer le spectre de déflexion de la matière par une méthode originale. Nous avons validé ses performances sur des simulations idéales en tentant de prendre en compte les effets systématiques dominants pour PLANCK. Nous avons montré que dans les conditions de fonctionnement nominales nous étions capables de mesurer le spectre de puissance de la déflexion. Dans ce chapitre, nous allons appliquer la technique développée précédemment pour analyser les données réelles de PLANCK. Nous allons montrer comment nous avons réalisé la première mesure du spectre de déflexion, permettant de confirmer sans ambiguïté l'existence de l'effet de lentille gravitationnelle.

Nous avons réalisé l'analyse sur trois cartes distinctes. Dans un premier temps nous avons utilisé les cartes des observations combinées à 143GHz et à 217GHz qui correspondent aux fréquences où le signal du CMB est dominant sur une large fraction du ciel, comparé à l'émission de la poussière galactique notamment. La carte à 100GHz est plus difficilement utilisable car le lobe de l'instrument à cette fréquence est beaucoup plus grand et plus elliptique. Nous avons également effectué l'analyse de l'effet de lentille gravitationnelle sur une carte issue de la séparation des composantes avec l'algorithme GMCA, afin d'essayer de s'affranchir d'éventuelles contaminations résiduelles.

Dans la première section, nous donnons un aperçu des principales étapes de la procédure de traitement des données permettant de produire les cartes de PLANCK, puis nous décrivons le jeu de données sur lequel se base notre analyse. Les sections suivantes décrivent ensuite les masques des données et la procédure de sélection des patchs. Enfin dans les trois dernières sections nous détaillons la méthode appliquée aux données pour mesurer le spectre de déflexion : les traitements préliminaires, l'estimation du spectre de la déflexion et la correction des effets systématiques par Monte-Carlo.

## 8.1 Préparation des données

Après la réception des données sur Terre, le niveau L1 est chargé de traiter les paquets télémétriques produits par le satellite, de décompresser les données et de les mettre en forme dans un format normalisé, exploitable par les modules de traitement des données du niveau L2 (voir section 4.5). Le point de départ de l'analyse au niveau L2 est donc une suite de mesures brutes, correspondant au signal à la sortie de l'électronique de lecture, après l'étape de compression et décompression décrite dans la section 5.

### 8.1.1 Traitement des données ordonnées en temps

Les données brutes produites par le satellite, pour chaque voie, sont une suite de mesures ordonnées en temps, modulées et exprimées en unité ADU (Unité de base de conversion analogique

numérique). On appelle parfois ces données les TOIs pour le terme anglais Time Ordered Information et leur traitement le *TOI-processing*. Le traitement des données ordonnées en temps a pour but, à partir des données brutes, de fournir un signal au contenu et aux dimensions physiques cohérents et nettoyé le mieux possible des effets systématiques, avant la production des cartes du ciel. Pour cela il faut effectuer une démodulation des données et une transformation du signal en ADU, en puissance déposée sur le bolomètre, en prenant en compte l'ensemble de la chaîne de mesure. Il faut par ailleurs supprimer les effets instrumentaux comme le bruit à haute fréquence ou les variations du signal produites par les variations de température du plan focal et traiter la contamination induite par le dépôt d'énergie des rayons cosmiques.

Depuis le lancement du satellite, le *TOI processing* a connu de nombreuses améliorations visant à réduire au maximum les effets instrumentaux et l'impact des rayons cosmiques sur la qualité du signal, notamment grâce à l'amélioration de la compréhension du comportement des instruments et des populations de rayons cosmiques, à partir des données présent en vol. Auparavant, l'ensemble du traitement était basé sur les données présent lors des campagnes de calibration des instruments au sol.

Les différentes améliorations sont identifiées par un numéro de version. Les données DX4, sur lesquelles est basée notre analyse (voir section 8.2), sont produites à partir de la version 4.1 du *TOI processing*, qui est la version stable la plus performante à l'heure de rédaction de ces lignes. Par rapport aux versions précédentes, elle possède notamment un meilleur modèle de la chaîne électronique, une meilleure détection et un marquage de l'impact des rayons cosmiques de haute énergie ainsi qu'une modélisation plus fine des différentes populations permettant la soustraction de leur queue. Enfin, cette version améliore également la soustraction des harmoniques, produites par les mouvements mécaniques périodiques du piston du compresseur de l'étage cryogénique à 4K.

### 8.1.2 Production des cartes

La production des cartes est effectuée à partir des données ordonnées en temps, après l'étape de *TOI processing*. C'est l'équipe du LAL qui a en charge cette production pour l'instrument HFI de PLANCK. Le principe de base consiste à projeter l'ensemble des échantillons sur une carte, à partir des données de pointage fournies par le senseur stellaire de PLANCK. Pour cela il faut d'abord pixeliser la sphère, puis, la valeur de chaque pixel de la carte est alors calculée grâce à une méthode de type NGP (Nearest grid point) qui affecte la valeur estimée à partir de l'ensemble des échantillons tombant dans un pixel au centre du pixel. Les cartes de l'instrument HFI sont toutes produites avec la pixelisation HEALPIX, avec un paramètre *nside* égal à 2048. Le code utilisé est décrit dans Tristram et al. (2011).

Dans la pratique, la production des cartes est une étape complexe. Elle nécessite par exemple une connaissance très fine du plan focal, car le senseur stellaire ne donne que la direction globale du satellite. Il faut en déduire ensuite de façon précise la direction exacte pointée par chaque détecteur sur le ciel. En plus de la température, il faut également estimer la polarisation du signal. Cela nécessite de bien combiner l'ensemble des mesures, de tenir compte des fuites de polarisation des bolomètres polarisés... Par ailleurs, il faut s'affranchir des dérives lentes produites par le bruit à basse fréquence. En effet, lorsqu'on combine des échantillons dans un même pixel, provenant d'observations ayant été réalisées dans un intervalle de temps suffisamment long pour que les variations du bruit à basse fréquence soient importantes (c'est-à-dire sur une échelle de temps supérieure aux échelles de variations caractéristiques du bruit à basse fréquence), on peut considérer que les échantillons mesurés dans un pixel correspondent à la mesure du même signal plus un décalage (ou *offset*) qui est lié au bruit et qui change sur des temps longs. L'hypothèse faite pour la production officielle des cartes dans PLANCK est qu'estimer un *offset* par *ring* est suffisant, étant donné le temps caractéristique de variation du bruit à basse fréquence (voir Tristram et al. (2011)).

Enfin, afin d'obtenir des cartes dans une unité appropriée, il faut estimer un facteur de calibration pour chaque voie, afin de faire la conversion des Watts déposés sur les bolomètres, dans l'unité voulue. Les cartes de HFI sont produites en température CMB. C'est-à-dire que la puissance reçue par le bolomètre, intégrée dans la bande passante de l'instrument, est convertie en température de corps noir équivalente. Le facteur de calibration inclue l'efficacité de la chaîne optique qui propage une fraction des photons initiaux seulement vers les détecteurs et l'intégration de l'émission sur la bande passante associée à chaque voie. La calibration est effectuée conjointement à l'estimation des *offsets*. À basse fréquence, pour le jeu de données que nous avons utilisé, les coefficients de calibration sont

estimés grâce à un ajustement du signal mesuré, à un patron produit à partir de la somme du dipôle orbital et du dipôle cosmologique mesuré par WMAP. Pour les voies à 545GHz et 857GHz, le signal est dominé par les émissions galactiques et il n'est plus possible d'utiliser la méthode précédente. La calibration est alors basée sur la comparaison à un patron de l'émission galactique, extrapolé à partir des observations de l'instrument FIRAS. La précision de cette méthode est de l'ordre du pourcent. À terme, il est prévu de réaliser, pour les voies à basse fréquence, une calibration basée sur la variation du dipôle orbital à six mois d'intervalle, induite par la variation de vitesse du satellite par rapport au référentiel du CMB, liée au mouvement du satellite autour du Soleil. La précision attendue, de cette méthode de calibration, est de quelques pour mille.

Étant donnée la stratégie de pointage du satellite, certaines zones du ciel sont observées plus de fois que d'autres. La couverture est donc inhomogène comme on peut le voir sur la figure 8.1. De façon générale, plus le nombre d'échantillons tombant dans un pixel est élevé, plus le rapport signal sur bruit va augmenter. Pour améliorer la qualité du signal par pixel, il est possible de produire des cartes en combinant plusieurs voies d'observations à une même fréquence. Enfin, pour étudier les effets systématiques, des cartes peuvent être produites avec un sous ensemble des échantillons, ce qui permet de comparer l'influence des effets qui varient avec le temps. Ainsi, en plus des cartes par fréquence, chaque production officielle est accompagnée de cartes produites avec la première moitié des cercles, puis la deuxième moitié, ou avec les cercles pairs, puis les cercles impairs, ou encore, avec la première couverture du ciel puis la seconde. Il est courant de se référer à ces jeux de données, coupés en deux et leur analyse sous le terme de cartes *Jack-Knives* ou analyse *Jack-Knife* respectivement. On les utilise en général pour tester les effets systématiques.

### 8.1.3 Inpainting

Au cours de la production des cartes, certains pixels ne sont pas reconstruits, soit parce qu'ils n'ont pas été observés soit parce qu'une des étapes de la chaîne d'analyse a rencontré un problème. En plus de cela, certaines régions du ciel sont contaminées par des émissions d'autres composantes que le CMB. Les pixels correspondants doivent donc être masqués. Les masques que nous avons utilisés sont décrits dans la section 8.3. Dans tous les cas, seules les zones masquées tombant dans des patches retenus après les trois sélections décrites dans la section 8.4, ont été inpaintées.

Nous avons utilisé l'inpainting disponible dans les codes FASTLens<sup>1</sup> et MRS<sup>2</sup> respectivement pour inpainter les patches et la sphère. Ils reposent tous les deux sur les mêmes hypothèses de base et supposent qu'il est possible de décomposer le signal sur un dictionnaire dans lequel la représentation possède une grande parcimonie (voir section 7.2). Le code MRS généralise la méthode dans le plan, comme celle employée par FASTLens (voir section 7.2), à des données sur la sphère. En outre, ces hypothèses sont aussi à la base de la méthode de séparation des composantes, décrite dans la section 8.1.4.

Pour les cartes combinées à 143GHz et à 217GHz nous utilisons la méthode d'inpainting locale par patch développée dans la section 7.2. Pour la carte issue de la séparation des composantes avec GMCA, la méthode est plus complexe et nous utilisons deux étapes d'inpainting distinctes. Dans un premier temps, on utilise le code MRS pour inpainter les sources ponctuelles et les zones d'émissions galactiques les plus intenses sur les six cartes en fréquence afin de pouvoir appliquer l'algorithme de séparation des composantes décrit dans la section 8.1.4 et déterminer la matrice de mélange. C'est un inpainting global de la sphère. Dans un second temps on utilise l'algorithme d'inpainting par patch sur la carte finale issue de la séparation des composantes, obtenue à partir de la matrice de mélange comme une simple combinaison linéaire des cartes en fréquence.

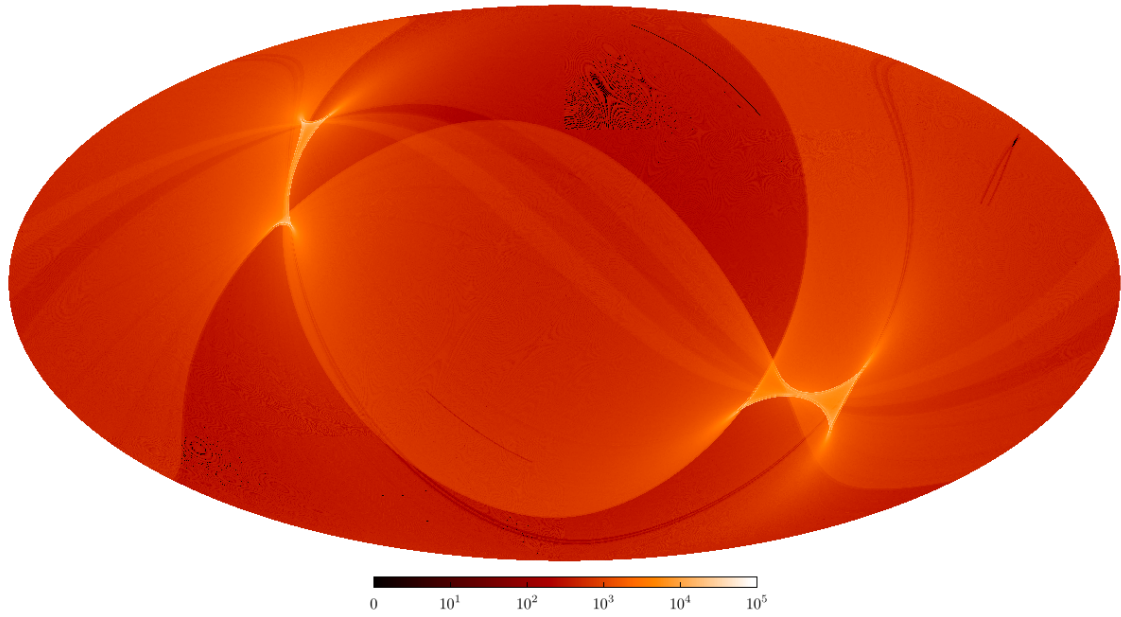
### 8.1.4 Séparation des composantes

En plus de l'émission du CMB, on trouve de nombreuses autres composantes astrophysiques qui émettent aux longueurs d'onde de PLANCK. On peut distinguer des émissions diffuses d'origine galactique, comme l'émission de la poussière, du rayonnement synchrotron et de freinage et des émissions ponctuelles extragalactiques comme les galaxies radios ou infrarouges et les amas SZ, ou galactique comme les nuages moléculaires froids. Le CMB étant émis avant, dans le temps, on

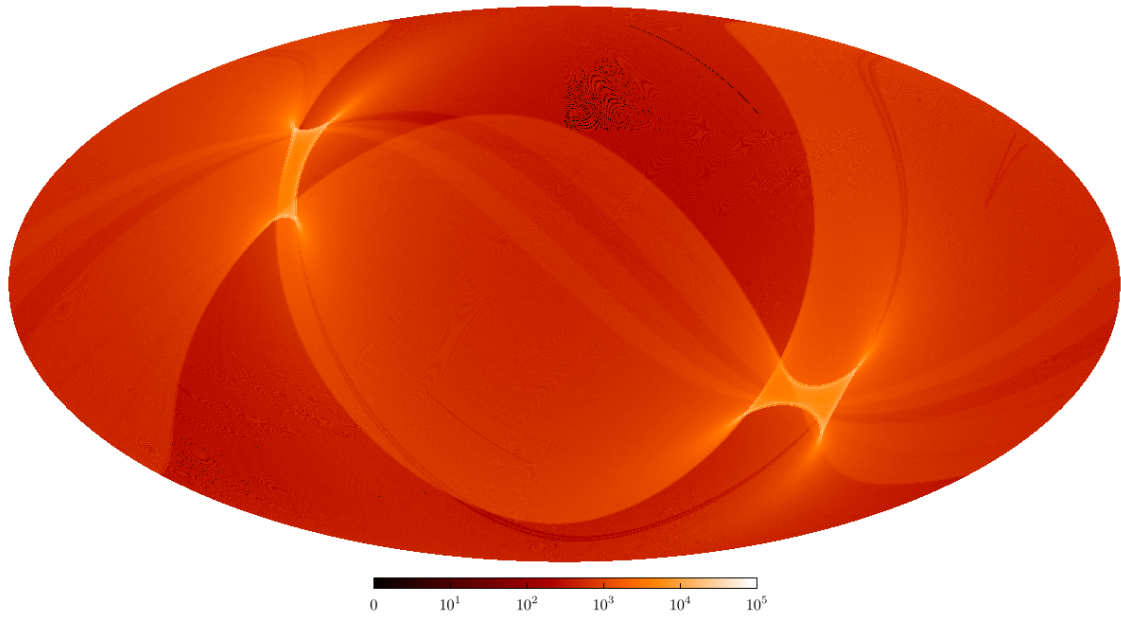
---

1. Voir le site web <http://irfu.cea.fr/Ast/fastlens.software.php>

2. Voir le site web [http://irfu.cea.fr/en/Phoce/Vie\\_des\\_labos/ast\\_visu.php?id\\_ast=895](http://irfu.cea.fr/en/Phoce/Vie_des_labos/ast_visu.php?id_ast=895)



(a) Nombre d'échantillons par pixel pour la carte combinée à 143GHz



(b) Nombre d'échantillons par pixel pour la carte combinée à 217GHz

FIGURE 8.1 – **Nombre d'échantillons par pixel pour les cartes à 143GHz et 217GHz du jeu de données DX4.** Chaque production de cartes est accompagnée des cartes contenant le nombre d'échantillons dans chaque pixel (parfois appelées cartes de *hit*). Elles permettent notamment d'estimer les variations de bruit d'un pixel à l'autre qui sont directement dépendantes du nombre de mesure. Étant donnée la fréquence d'échantillonnage des bolomètres de HFI, de  $F_{samp} = 180.38Hz$ , chaque échantillon correspond à un temps d'intégration d'environ 6 ms. L'échelle de couleur est une échelle logarithmique. On remarque bien que la stratégie de pointage de PLANCK produit des zones très denses autour des pôles écliptiques.

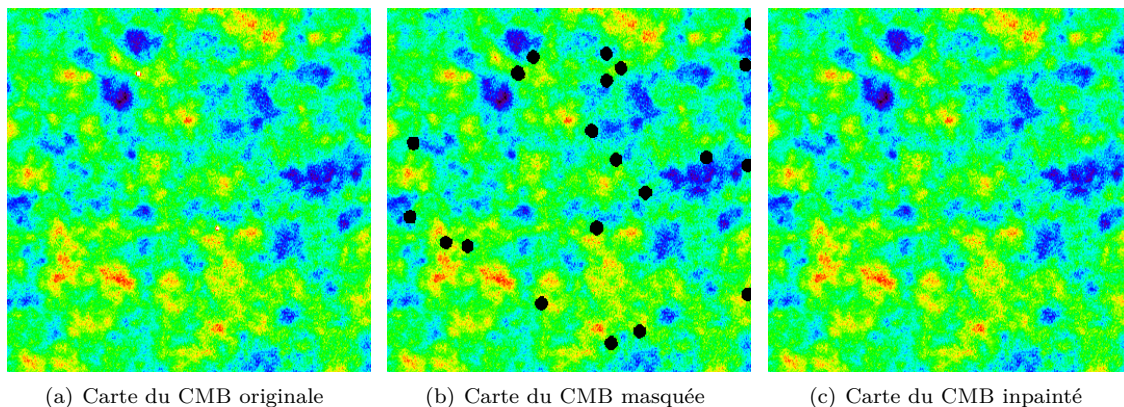


FIGURE 8.2 – **Exemple d’inpainting de la carte à 217GHz.** La figure (a) montre un patch de 10 degrés de coté découpé dans la carte des observations combinées à 217GHz, sur une position centrée à  $-35.7^\circ$  de latitude et  $225^\circ$  de longitude. Les pixels masqués (en noir sur la figure (b)) sont ensuite inpaintés avec l’inpainting local par patch décrit dans la section 7.2. Les zones inpaintées de la carte (c) sont très similaires aux zones à l’extérieur du masque. Les sources ponctuelles, dont certaines sont fortement visibles sur la carte originale, ont totalement disparues.

regroupe souvent l’ensemble des émissions précédentes sous le terme d’avant-plans. C’est la somme de toutes les émissions qui est capturée par les détecteurs. Lorsqu’on veut étudier l’émission du CMB par exemple, les autres composantes constituent alors une contamination qui modifie les propriétés statistiques, ce qui peut entraîner notamment une dégradation de l’estimation de l’effet de lentille gravitationnelle.

Dans la plupart des analyses et en particulier pour l’analyse de l’effet de lentille gravitationnelle, il est intéressant de travailler sur une carte la moins contaminée possible par les autres émissions. Pour cela on peut essayer de séparer la contribution des différentes émissions afin d’obtenir une carte par composante. C’est évidemment un exercice très difficile, qui occupe un groupe de travail entier dans le consortium PLANCK, le WG2. Afin de séparer les différentes émissions, les algorithmes tirent partie de la différence de structure morphologique et de répartition spatiale et spectrale des émissions. Pour atteindre les meilleures performances, la séparation de composante est réalisée à partir d’une analyse multi-fréquence. Les neuf bandes de fréquence de PLANCK ont été prévues principalement pour cette analyse.

Le rayonnement mesuré sur le ciel par chaque détecteur peut être modélisé comme une combinaison linéaire de l’émission des différentes composantes. On peut alors écrire le signal observé à une fréquence  $\nu$ , dans une direction  $\mathbf{n}$  sous la forme :

$$T_\nu^{obs}(\mathbf{n}) = A_{\nu i} s^i(\mathbf{n}) + n_\nu(\mathbf{n}) \quad (8.1)$$

où,  $s^i(\mathbf{n})$  est l’émission associée à chacune des  $n_c$  composantes  $\{s^i(\mathbf{n})\}_{i=1,\dots,n_c}$ , la matrice  $A$  est la matrice de mélange des différentes émissions en fonction de la fréquence et  $n_\nu(\mathbf{n})$  est le bruit instrumental à la fréquence  $\nu$ . Le but de la séparation des composantes est alors d’estimer l’ensemble des  $s_i$  à partir des observations  $T_\nu^{obs}$  ce qui correspond à estimer la matrice  $A_{\nu i}$ .

Huit méthodes concurrentes ont été développées dans le cadre du WG2, pour essayer de répondre au mieux aux besoins de PLANCK. En règle générale, elles se basent donc sur un certain nombre de cartes à différentes fréquences en entrée et en déduisent les cartes de l’émission des différentes composantes ainsi qu’une carte de résidus et de bruit instrumental, en faisant parfois certaines hypothèses supplémentaires sur les propriétés des différentes émissions. Les méthodes diffèrent principalement par les contraintes extérieures qu’elles utilisent et par les modèles physiques d’émission qu’elles supposent. Une revue des différents algorithmes de séparation de composantes et une comparaison de leurs résultats, issus du travail du WG2, peuvent être trouvées dans Leach et al. (2008).

Pour notre étude, nous avons utilisé la méthode GMCA (Generalized Morphological Component Analysis) développée par Bobin et al. (2008), qui réalise une décomposition basée sur une analyse morphologique à l’aveugle. C’est la seule méthode de séparation de composantes dont l’influence



sur l'effet de lentille gravitationnelle a été étudiée à l'heure actuelle (voir à ce sujet Perotto et al. (2010)). Son principal avantage est l'analyse à l'aveugle qui ne nécessite pas de réglages particuliers. Le fonctionnement de l'algorithme repose sur des travaux récents en traitement du signal (voir les références dans Bobin et al. (2008)).

Pour réaliser la séparation des composantes, on part du principe que les différentes émissions, possèdent chacune, de par leur origine physique distincte, une morphologie et une répartition spatiale propres et distinctes de celles des autres. L'algorithme GMCA s'appuie alors sur l'hypothèse principale qu'on peut trouver un dictionnaire  $\mathcal{D}$  de fonctions (voir section 7.2.3) sur lequel la décomposition de chaque composante possède une grande parcimonie et est distincte de celle des autres. Si on appelle  $\alpha_{ik}$  les coefficients de la décomposition sur les  $k$  composantes de  $\mathcal{D}$  de la source  $s^i$ , il a été montré dans Bobin et al. (2008) qu'un bon estimateur de la parcimonie est la somme de la valeur absolue des  $\alpha_{ik}$ . La détermination de la matrice de mélange et des différentes composantes du signal se fait alors par une approche bayésienne de maximisation de la parcimonie de la décomposition, qui peut s'écrire sous la forme du problème de minimisation suivant :

$$\min_{\{s^i\}, \{A_{\nu i}\}} \sum_{ik} |\alpha_{ik}| \quad (8.2)$$

soumis à la contrainte :

$$\|T_{\nu}^{obs}(\mathbf{n}) - A_{\nu i} s^i(\mathbf{n})\| < \epsilon \quad (8.3)$$

où  $\epsilon$  est un paramètre fixant l'erreur de reconstruction autorisée et  $\|\cdot\|$  représente la norme  $\ell_2$ . (voir Bobin et al. (2008)).

Finalement chaque composante du signal peut être obtenue à partir des observations  $\{T_{\nu}^{obs}(\mathbf{n})\}$  en appliquant le pseudo-inverse de la matrice de mélange :

$$s^i(\mathbf{n}) = \sum_{\nu} A_{\nu i}^{+} T_{\nu}^{obs}(\mathbf{n}) \quad (8.4)$$

où  $A_{\nu i}^{+}$  est l'élément  $i, \nu$  du pseudo-inverse de  $A$  donné par  $A^{+} = (A^T A)^{-1} A^T$ . Chaque composante est donc une combinaison linéaire des observations ce qui permet de propager facilement les erreurs et les propriétés du bruit instrumental.

## 8.2 Jeux de données de départ pour l'analyse

### 8.2.1 Données DX4

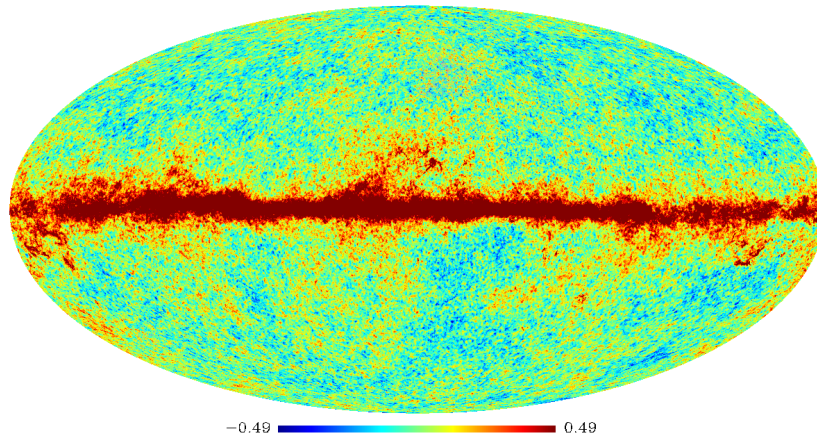
Dans la collaboration PLANCK, le traitement des données des deux instruments HFI et LFI est réalisé de façon indépendante par les deux consortiums, en charge des instruments respectifs. Les deux détecteurs utilisent des technologies différentes, qui nécessitent des traitements spécifiques et qui ne sont pas sensibles aux mêmes effets systématiques. Pour chacun des deux instruments, l'ensemble du traitement des données est effectué au sein d'un *Core Team* qui est un groupe défini de personnes qui pilote l'analyse et qui a un accès privilégié aux mesures jusqu'en 2013, date à laquelle l'ensemble des données sera rendu public.

Les deux *Core Team* procèdent régulièrement à des échanges de données sur un ensemble de mesures fixé à l'avance. Au cours du mois d'août 2010, le quatrième échange a eu lieu. Le jeu de données associé porte le nom de DX4, pour le terme anglais Data eXchange 4. Ils est basé sur les observations du 12 août 2009 à 14h13 au 06 juin 2010 à 14h53, soient les *ring* 240 à 9467, ce qui représente un peu moins de deux observations du ciel complet.

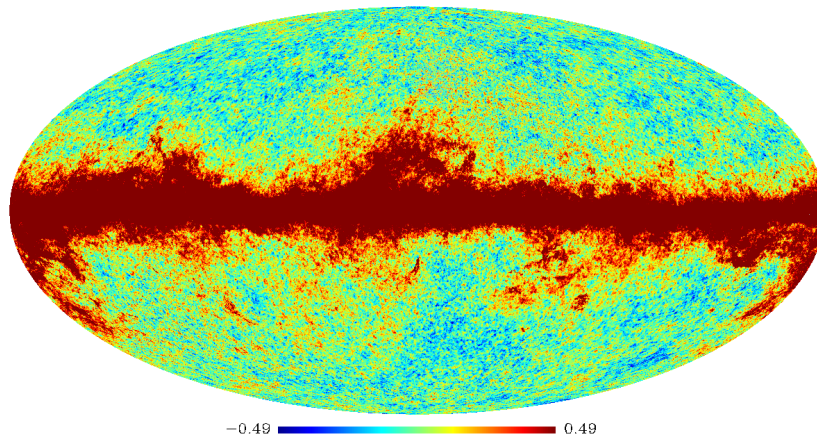
Nous avons basé notre étude sur ce jeu de données car il a été l'objet d'un nombre important de tests et d'études en vue de l'échange, mais également car c'est le jeu de données sur lequel se base les premiers résultats scientifiques publiés par la collaboration en janvier 2011. De ce fait on dispose donc également d'un ensemble de mesures intéressantes pour l'analyse, notamment un large catalogue d'objets compacts.

### 8.2.2 Carte des mesures combinées à 143GHz

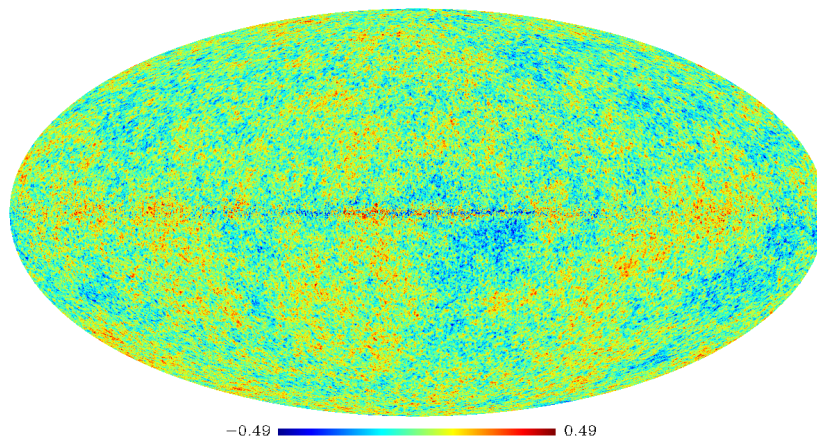
L'instrument HFI possède 12 voies de mesures à 143GHz. Cependant, le bolomètre 143\_8 possède un bruit télégraphique très important (voir section 5.5.5) qui rend ses mesures inutilisables. Dans



(a) Carte du ciel, des observations combinées à 143GHz

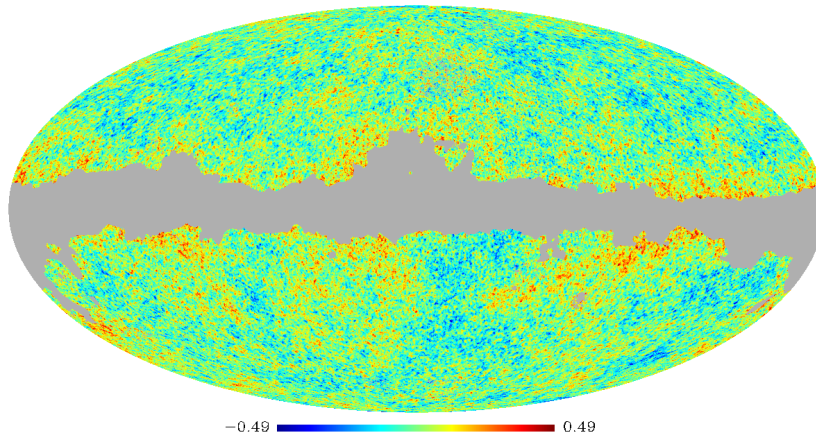


(b) Carte du ciel, des observations combinées à 217GHz

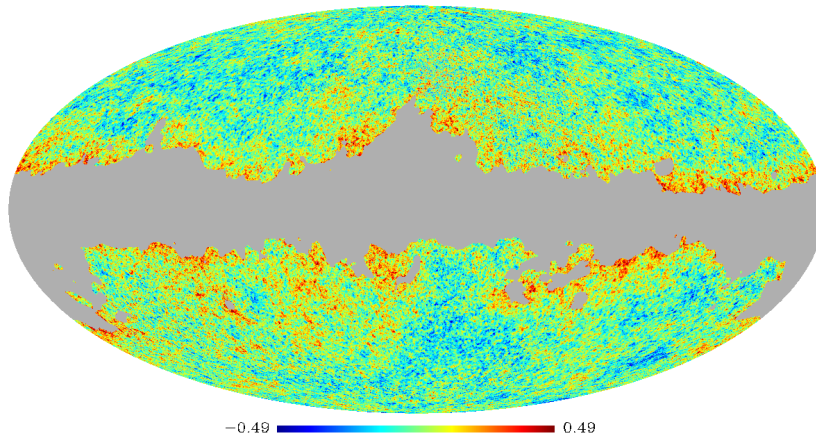


(c) Carte du CMB, après séparation de composantes avec GMCA

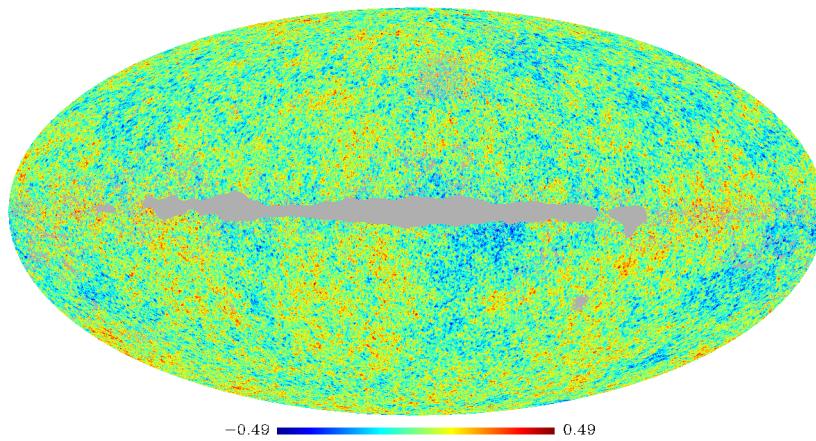
FIGURE 8.3 – **Cartes du jeu de données DX4, utilisées pour l'analyse.** Les cartes ((a)) et ((b)) sont obtenues à partir de la combinaison des observations à 143GHz et 217GHz respectivement. La carte ((c)) est issue de la séparation des composantes avec l'algorithme GMCA. On voit que plus la fréquence augmente plus la contamination de la poussière est importante. La carte issue de la séparation des composantes ne possède qu'un faible résidu galactique, visible dans la partie centrale. L'unité des cartes est le mK.



(a) Carte du ciel, des observations combinées à 143GHz, masquée et inpaintée



(b) Carte du ciel, des observations combinées à 217GHz, masquée et inpaintée



(c) Carte du CMB après séparation de composantes avec GMCA, masquée et inpaintée

FIGURE 8.4 – **Cartes du jeu de données DX4, utilisées pour l'analyse, masquées et inpaintées.** Les cartes sont masquées avec les masques combinés associés, décrits dans la section 8.3.4. Elles sont ensuite inpaintées avec l'inpainting local par patch. Les zones dont les patches possèdent plus de 5% des pixels masqués ne sont pas inpaintées, c'est pourquoi des zones masquées sont toujours visibles sur les trois cartes.

l'état actuel de la compréhension de l'origine de ce bruit, aucune solution satisfaisante n'a été trouvée pour corriger les erreurs introduites et les voies possédant du bruit télégraphique ne sont donc pas prises en compte dans la production des cartes combinées. La carte à 143GHz que nous avons utilisée est issue de la production officielle pour le DX4. Elle a été produite à partir de la combinaison des 11 voies restantes à 143GHz.

À 143GHz, l'émission est dominée par celle du CMB, sur une grande portion du ciel (voir figure 8.3(a)). Cependant en l'absence de séparation des composantes, certaines zones restent très contaminées par l'émission des autres composantes astrophysiques et doivent être masquées. Nous repérons les pixels invalides grâce au masque combiné de sources compactes, de l'émission galactique et de la couverture MCOMB143 (voir section 8.3.4), qui couvre 22.36% du ciel. Les pixels masqués sont ensuite inpaintés avec la méthode d'inpainting local par patch développée dans la section 7.2.

Pour cela, nous utilisons la méthode de projection couplée à l'interpolation bilinéaire qui constitue un bon compromis entre rapidité et qualité de l'inpainting. Nous découpons des patches de 10 degrés de coté et 350 pixels carrés, centrés sur les centres des pixels d'une carte HEALPIX de paramètre *nside* égal à 8. Nous fixons la valeur du paramètre de seuil à 5%. Ainsi seules les régions du ciel tombant dans un patch possédant moins de 5% des données masquées sont inpaintées. Grâce à la souplesse de l'analyse par patch, les zones du ciel non inpaintées ne sont pas prises en compte dans le reste de l'analyse.

### 8.2.3 Carte des mesures combinées à 217GHz

L'instrument HFI possède également 12 voies de mesure à 217GHz. Encore une fois, nous avons utilisé la carte issue des mesures combinées en utilisant l'ensemble des voies à cette fréquence. À 217GHz, l'émission de la poussière est plus importante qu'à 143GHz (voir figure 8.3(b)) et nous utilisons le masque combiné MCOMB217, comparable à celui à 143GHz construit de façon classique à partir du masque de couverture, du masque de sources compactes mais avec un masque galactique plus important qu'à 143GHz. Il couvre 31.52% des pixels. Nous avons réalisé l'inpainting des zones masquées en appliquant une procédure identique à celle utilisée pour la carte à 143GHz.

### 8.2.4 Carte issue de la séparation des composantes avec GMCA

On a vu dans la section précédente que pour augmenter le rapport signal sur bruit, on peut combiner les observations des différentes voies à une même fréquence. Pour obtenir un rapport signal sur bruit encore meilleur, on peut être tenté de vouloir projeter directement, l'ensemble des mesures à toutes les fréquences afin d'obtenir un grand nombre d'échantillons par pixels. Cependant cette opération n'est pas avantageuse car, à haute fréquence, le signal est dominé par les émissions galactiques et de façon plus générale, chaque bande de fréquence possède une certaine contamination. Produire une carte à partir de l'ensemble des voies additionne plus les problèmes qu'elle ne les résout.

Cependant, il est intéressant d'essayer de combiner l'ensemble des informations disponibles de façon particulière pour obtenir la carte de CMB la plus propre possible. C'est le rôle de la séparation des composantes (voir la section 8.1.4). Nous avons donc choisi d'utiliser une carte du CMB issue de la séparation des composantes. Nous avons sélectionné l'algorithme GMCA, notamment car nous avons déjà montré Perotto et al. (2010) que cette méthode de séparation était adaptée à une analyse de l'effet de lentille gravitationnelle.

Nous partons des six cartes en fréquence des observations combinées des détecteurs de HFI. L'algorithme de GMCA n'est pas adapté à la séparation des sources ponctuelles. Avant de réaliser la séparation des composantes nous avons donc effectué un inpainting de chaque carte à partir du masque des objets compacts et du masque de couverture associés à chaque fréquence (voir section 8.3), en utilisant la méthode m2 du code MRS (voir section 8.1.3). Après la séparation avec GMCA, on obtient donc une carte de l'émission du CMB. Elle est obtenue à partir de la combinaison linéaire des cartes combinées originales (sans inpainting) à partir du pseudo inverse de la matrice de mélange.

Dans la zone du plan galactique, la séparation des composantes est difficile car les émissions galactiques dominent à toutes les fréquences. La carte de CMB contient donc dans cette région des résidus galactiques. Par ailleurs comme les cartes utilisées pour la combinaison linéaire ne sont pas inpaintées, la carte de CMB contient des sources ponctuelles et des pixels non observés. La dernière étape consiste donc à inpainter les zones problématiques à partir de l'union des masques des objets compacts tenant compte du poids de chaque fréquence, des masques de couverture des

	Carte 143GHz	Carte 217GHz	Carte GMCA
masque de couverture	0.17%	0.12%	0.35%
masque galactique	20.03%	30.19%	3.87%
masque de sources compactes	4.45%	3.57%	6.14%
masque combiné	22.36%	31.52%	9.27%

TABLE 8.1 – **Tableau récapitulatif du pourcentage de pixels masqués pour les différents masques.** En l'absence de séparation de composante, la fraction des pixels masqués est dominée par ceux des masques galactiques. La taille de la région masquée autour des sources compactes est fonction de la taille du lobe. On observe donc un proportion de pixels masqués par le masque de sources compactes plus faible à 217GHz car le lobe de l'instrument est plus petit. Le nombre de pixels masqués par les masques combinés n'est pas la somme des pixels masqués par les masques individuels car ils peuvent posséder des pixels en commun. Enfin les masques de couverture et de sources compactes de la carte issue de GMCA combinent les masques à toutes les fréquences et possèdent donc un plus grand nombre de pixels masqués. Au contraire le masque galactique pour la carte séparée de composantes est le plus petit car seuls quelques résidus subsistent.

six cartes combinées et du petit masque galactique MGALGMCA (voir section suivante). Pour cela nous utilisons l'inpainting local par patch avec les mêmes paramètres que pour les cartes à 143GHz et 217GHz.

## 8.3 Masquage des données

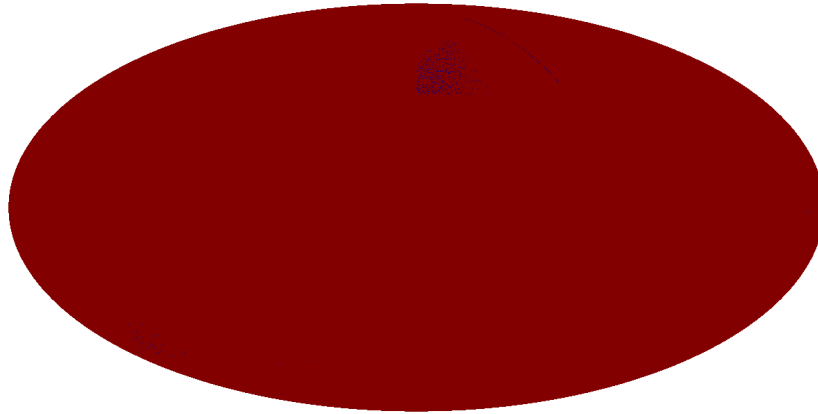
Au cours de la production des cartes, à cause de la stratégie de pointage ou de problèmes instrumentaux par exemple, le signal de certains pixels ne peut pas être reconstruit. Par ailleurs, les pixels tombant dans des zones du ciel, pour lesquelles, la contamination par des composantes autres que le CMB est trop forte, ne peuvent pas être pris en compte dans certaines étapes de l'analyse. Il est alors utile de définir des masques binaires permettant de rejeter ou non les pixels en fonction des besoins. Pour cela, nous avons défini trois types de masques de base, ainsi que différentes combinaisons produites à partir de l'union d'un à trois de ces masques. Chacun des masques combinés est alors utilisé pour une étape particulière de l'analyse.

Pour des raisons de simplicité et afin d'alléger la lecture, nous avons donné des noms spécifiques à chaque masque détaillé dans la suite. L'identifiant n'est pas un identifiant officiel mais permet cependant de nous référer de façon simple et non ambiguë à un masque particulier.

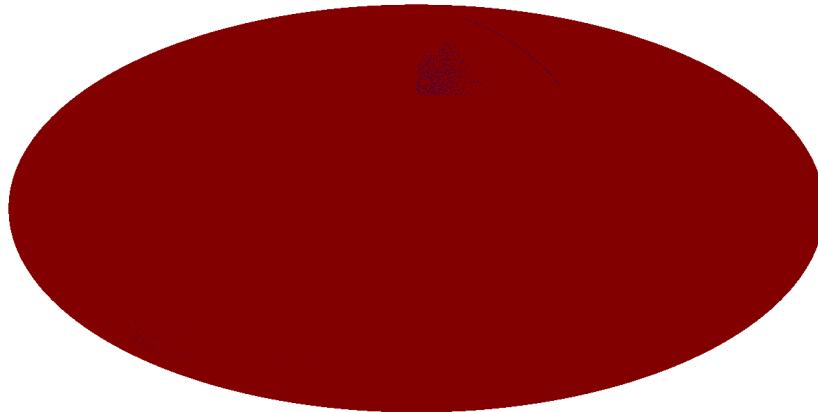
### 8.3.1 Masques de couverture

Les données DX4 intègrent approximativement 10 mois d'observation (du 12 août 2009 au 6 juin 2010). Le 13 août 2010, qui est la date officielle de la fin du second *survey*, on avait 95% de couverture commune entre les deux observations. Les cartes DX4 sont donc issues d'un peu moins de deux couvertures du ciel. Dans tous les cas, il faut environ 7 mois pour couvrir 99% du ciel, mais la couverture reste toutefois toujours très inhomogène. En effet, la stratégie de pointage de PLANCK, basée sur des observations sur des cercles, entraîne une grande redondance des observations aux pôles écliptiques. Par ailleurs, après 10 mois de prise de données, certaines zones du ciel sont déjà observées pour la deuxième fois alors que d'autres n'ont vu qu'un passage.

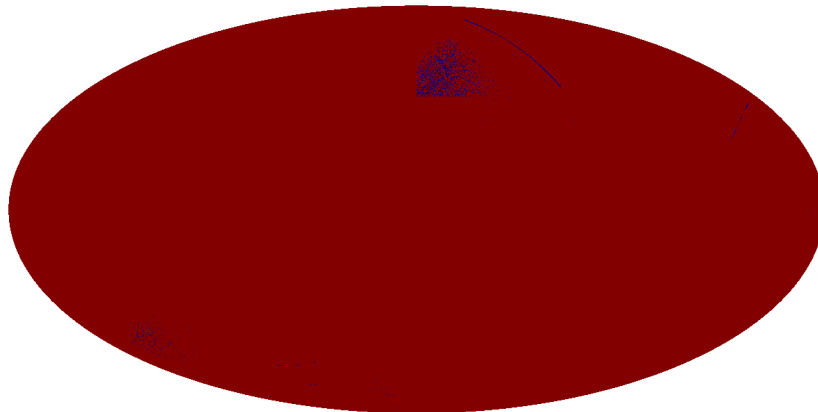
L'étude par patches nous a permis de réaliser dès décembre 2009, une première analyse avec une petite fraction du ciel uniquement. C'est pourquoi nous avons intégré un masque tenant compte de la couverture du ciel, dans la chaîne d'analyse. Évidemment, au fur et à mesure des observations, la proportion du ciel couverte augmente. Cependant, étant données la haute résolution des cartes produites, la stratégie de pointage de PLANCK et la géométrie des pixels, même après 10 mois, on trouve, pour chaque détecteur, un certain nombre de pixels qui n'ont toujours pas été observés, ou pour lesquels la valeur du signal n'a pas pu être reconstruite. Par ailleurs, comme chaque détecteur observe une région légèrement différente du ciel, le fait de combiner les mesures de plusieurs détecteurs permet toutefois de diminuer le nombre de pixels vides. Nous intégrons dans tous les cas, systématiquement, le masque de couverture à l'analyse.



(a) Masque de couverture pour la carte combinée à 143GHz



(b) Masque de couverture pour la carte combinée à 217GHz



(c) Masque de couverture pour la carte issue de la séparation de composante

FIGURE 8.5 – **Masques de couverture.** Plus de 99.8% des pixels sont observés à toutes les fréquences (En rouge sur les figures. Voir également le tableau 8.1). Seules deux zones distinctes contiennent un grand nombre de pixels pour lesquels le signal n'a pas pu être reconstruit (En bleu sur les figures). Ce problème est causé par une combinaison particulièrement défavorable de la stratégie de pointage, de la fréquence d'échantillonnage et de la géométrie des pixels HEALPIX, qui permet l'apparition de pixels dans lesquels aucun échantillon ne tombe. Une longue bande de pixels masqués est également visible au Nord-Est. Ces pixels ont été masqués suite à un problème connu et compris, dans le traitement des données ordonnées en temps après l'observation d'une source intense (la planète Jupiter).

Fréquence (Hertz)	100	143	217	353	545	857
Taille des disques (minutes d'arc)	21	15	10	9.5	9.5	9

 TABLE 8.2 – **Tableau récapitulatif de la taille des cercles masqués en fonction de la fréquence.**

Les masques de sources compactes sont créés à partir d'un ensemble de catalogues. Un cercle de taille variable est masqué autour de la position de chaque objet. La taille du cercle dépend de la taille du lobe de l'instrument et varie donc avec la fréquence.

Nous avons produit un masque de couverture pour chacune des six cartes combinées. Ils sont appelés MCOVFFF, où FFF est la fréquence d'observation des cartes auxquelles correspondent les masques. Chaque masque indique simplement les pixels qui n'ont pas été observés et qui ne doivent pas être intégrés aux mesures. Les masques sont créés à partir de la carte de *hit* associée à chaque carte produite, qui donne le nombre d'échantillons tombant dans chaque pixel. Ils sont alors construits en identifiant les pixels vides.

### 8.3.2 Masques de sources compactes

En plus des émissions diffuses, comme celle de la poussière galactique, on trouve un grand nombre d'objets compacts, d'origine galactique ou extragalactique, qui émettent aux longueurs d'onde de PLANCK. Ce sont principalement, des galaxies radios, des blazars, des galaxies infrarouges lumineuses, des amas SZ, des nuages moléculaires galactiques froids... Ils peuvent produire une importante contamination des données s'ils ne sont pas masqués.

L'effet de ces objets sur le signal, dépend de leur amplitude à la longueur d'onde considérée, comparée à l'amplitude de l'émission du CMB. Les sources suffisamment puissantes, dont l'amplitude dépasse l'amplitude moyenne locale du CMB, sont visibles individuellement et modifient de façon significative la valeur locale du signal. Les sources dont l'amplitude est trop faible et qui sont noyées dans le bruit, ne sont pas visibles individuellement. Leurs résidus contribuent toutefois de façon subtile et non négligeable et peuvent modifier les propriétés statistiques du signal par exemple. Ces sources produisent donc un bruit de fond. Pour réaliser ce masque, nous avons privilégié l'exhaustivité, car un petite zone masquée en plus à un effet négligeable comparé à celui d'une source oubliée.

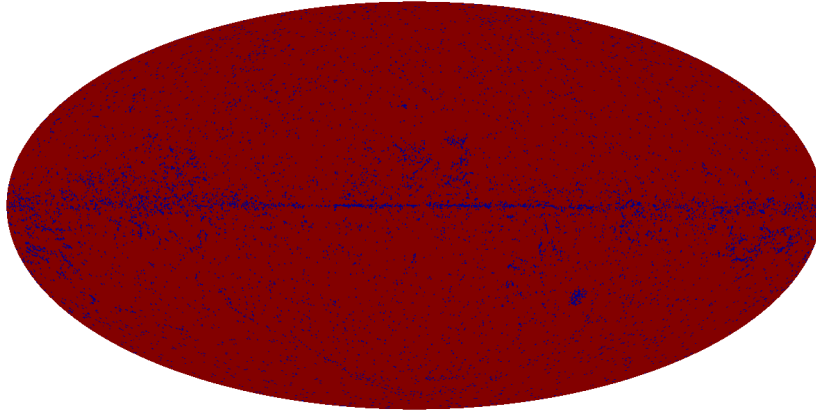
Beaucoup de ces objets sont suffisamment lointains pour avoir une taille angulaire petite devant la taille du lobe de l'instrument et sont traités comme des sources ponctuelles. Dans ce cas, on les appelle parfois, également, objets non résolus car leur image est totalement diluée par la réponse impulsionnelle spatiale de l'instrument. De par la nature ponctuelle ou quasi-ponctuelle de ces objets, le masque de sources compactes est généré à partir d'un ensemble de catalogues qui fournissent la position de chaque objet dans le ciel. On masque alors l'ensemble des pixels qui tombent dans un cercle dont le diamètre est fixé par la taille du lobe à chaque fréquence (voir le tableau 8.2).

Dans la mesure du possible nous avons utilisé les catalogues basés sur les données de PLANCK car c'est la première fois qu'on fait des observations sur tout le ciel aux plus hautes fréquences de l'instrument. Les masques de sources compactes sont ainsi construits à partir du catalogue ERCSC de PLANCK Planck Collaboration et al. (2011), du catalogue de sources de la bande W de WMAP7 et de catalogues extérieurs et compilés par Laurence Perotto<sup>3</sup>. Le catalogue ERCSC inclut un catalogue de corps froids qui sont des objets galactiques détectés à l'aide de l'algorithme Cococodet (Montier et al. (2010)), un catalogue d'amas SZ détectés avec une méthode basée sur Melin et al. (2006) et un catalogue de sources ponctuelles détectées avec les programmes PowellSnakes (Carvalho et al. (2009)) pour les cartes aux fréquences inférieures à 143GHz et SExtractor (Bertin & Arnouts (1996)) pour les cartes aux fréquences supérieures à 217GHz. Le catalogue ERCSC est de loin le plus complet de tous et il n'est basé que sur des observations de PLANCK.

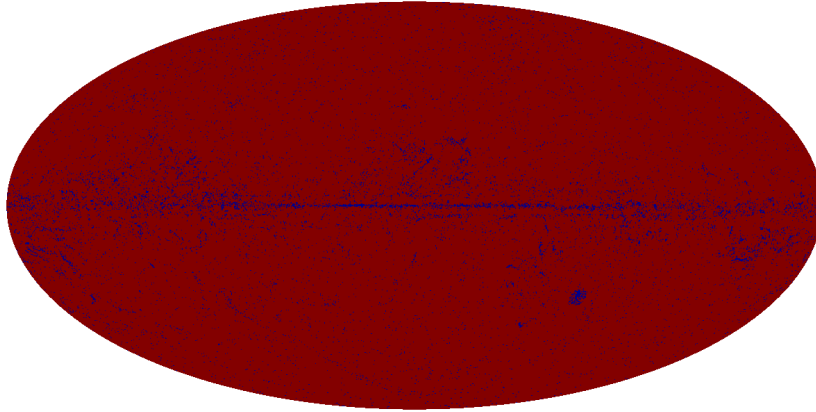
### 8.3.3 Masques galactiques

Les masques galactiques sont produits à partir des données de PLANCK. On utilise pour cela la carte des observations combinées à 857GHz. À cette fréquence l'émission est largement dominée par

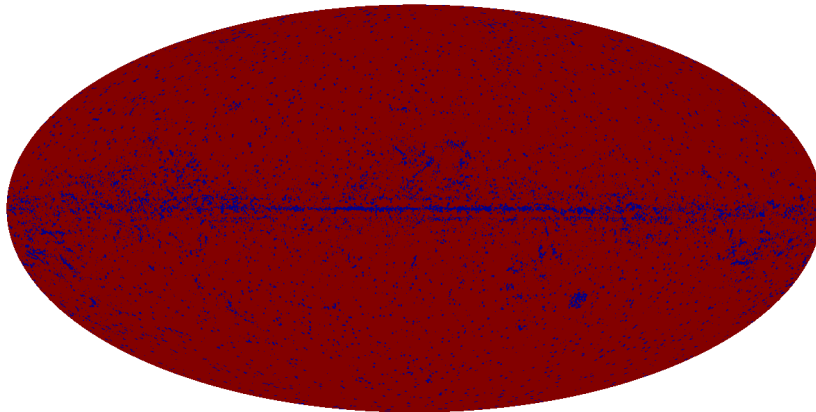
3. voir note interne <http://wiki.planck.fr/index.php/Lensing/DX4masks>



(a) Masque de sources compactes pour la carte combinée à 143GHz



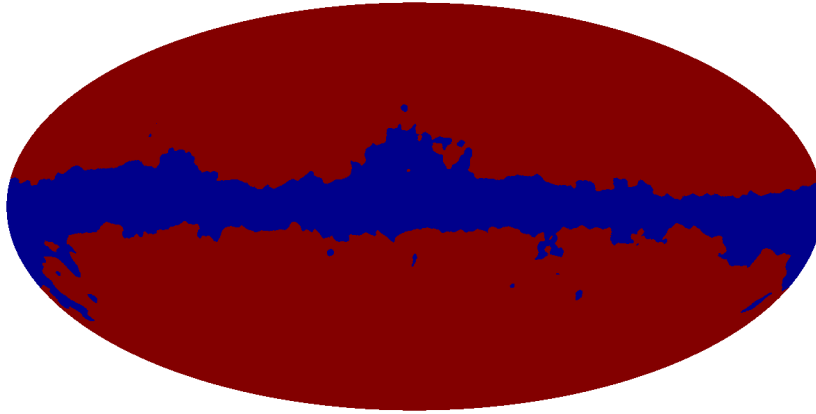
(b) Masque de sources compactes pour la carte combinée à 217GHz



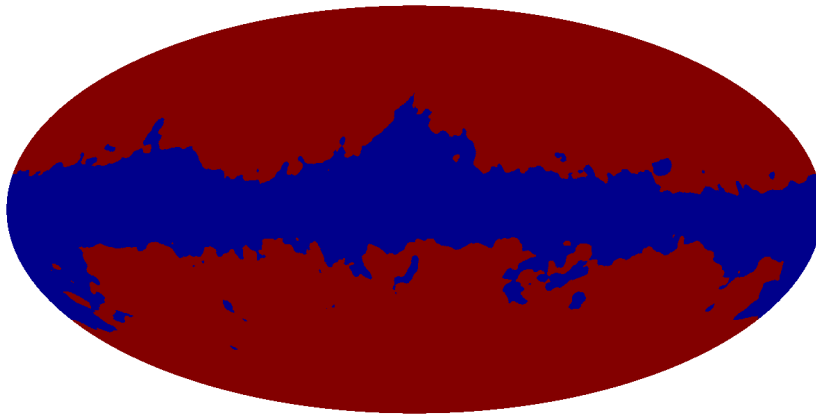
(c) Masque de sources compactes pour la carte issue de la séparation de composantes

FIGURE 8.6 – **Masques de sources compactes.** On observe un grand nombre d'objets compacts sur tout le ciel et à toutes les fréquences. Leur répartition est cependant fortement inhomogène car la proportion des sources galactiques est largement dominante. On voit que les zones les plus denses sont celles du plan galactique. On observe également une forte concentration d'objets dans la partie Sud-Est qui sont des sources extragalactiques proches, situées dans le nuage de Magellan.

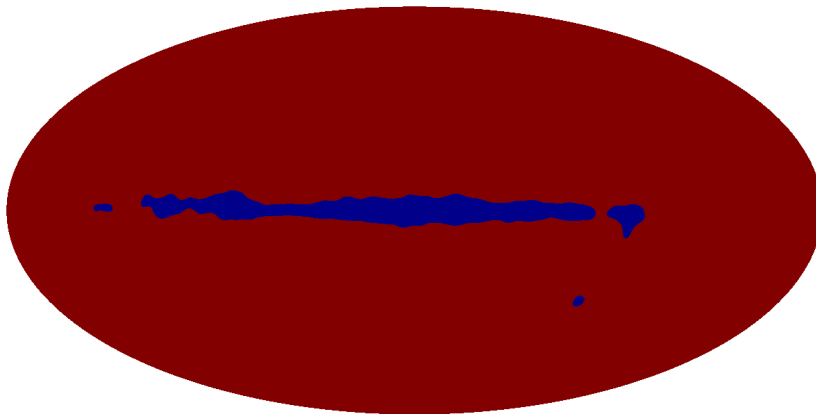




(a) Masque galactique pour la carte combinée à 143GHz



(b) Masque galactique pour la carte combinée à 217GHz



(c) Masque de galactique pour la carte issue de la séparation de composante

FIGURE 8.7 – **Masques galactiques.** La zone contaminée par l'émission de la poussière galactique augmente fortement avec la fréquence. Après séparation des composantes les résidus galactiques sont très faibles et le masque utilisé est plus petit.

l'émission de la poussière galactique, comme on peut le voir sur la figure 2.9. On produit donc les masques galactiques à partir d'un seuil en amplitude de la carte à 857GHz.

Les seuils que nous avons utilisé, conduisent à des masques à 143GHz et 217GHz couvrant respectivement 20.03% et 30.19% des pixels de la carte. Nous les appelons MGAL143 et MGAL217 respectivement. On peut remarquer sur la figure 8.8, par exemple, que certaines zones du ciel en périphérie du masque, semble être toujours contaminées par l'émission de la poussière. On peut noter également, l'absence du petit nuage de Magellan ou de la galaxie d'Andromède (M31) des masques galactiques. L'émission de la galaxie est une émission diffuse et s'étend évidemment bien au delà du masque. Le seuil en amplitude que nous avons utilisé ne permet donc pas de masquer toutes les zones contaminées. Cependant en dehors du masque, on s'attend à ce que l'émission soit dominée par celle du CMB sur la majeure partie de ciel. Par ailleurs, même si cette propriété n'est pas vérifiée dans certaines régions du ciel, les conséquences sur l'analyse restent négligeables. En effet, dans tous les cas, pour s'assurer que les patchs que nous utilisons sont bien dominés par le CMB, nous réalisons une série de sélections basées sur les propriétés statistiques du CMB (gaussianité et isotropie, voir section 8.4) qui ne sont pas vérifiées par l'émission de la poussière et qui permettent d'éliminer les zones contaminées restantes. Les masques galactiques servent principalement à éliminer un premier jeu de patchs afin de limiter les calculs relativement lourds de l'ensemble de la chaîne d'analyse. On peut vérifier sur la figure 8.8 que les patchs situés dans les régions du petit nuage de Magellan et de la galaxie d'Andromède ont été éliminés sur la carte à 217GHz. Pour la carte à 143GHz, la zone correspondant à M31 n'a pas été éliminée mais ne semble pas être fortement contaminée.

Enfin, nous utilisons un troisième masque galactique, moins restrictif, utilisé pour l'étude de la carte issue de la séparation de composantes. En effet, après l'application de l'algorithme de séparation, on obtient une carte de CMB sur tout le ciel. Cependant, dans les zones les plus contaminées par l'émission de la poussière, principalement dans la partie centrale du plan galactique, des résidus galactiques sont toujours présents dans la carte du CMB. Le groupe de travail chargé de la séparation de composantes dans PLANCK, a généré un masque d'émission galactique de petite taille qui peut être utilisé comme masque des résidus galactiques après séparation de composantes. Nous nous y référons sous le nom de MGALGMCA.

### 8.3.4 Masques combinés

À partir des masques décrits précédemment, nous avons réalisé des combinaisons particulières, adaptées aux différentes étapes de l'analyse. La carte issue de la séparation des composantes, notamment, nécessite un jeu de masques issus de la combinaison des masques en fréquence.

Dans un premier temps nous générons un masque de couverture pour la carte GMCA (masque MCOVGMCA) en combinant l'ensemble des masques de couverture aux six fréquences de HFI. De plus, comme GMCA n'inclut pas de composantes de sources ponctuelles, il faut également réaliser un masque de sources ponctuelles adapté à l'analyse de la carte issue de la séparation des composantes. Sa construction est un peu délicate. En effet, on obtient la carte GMCA à partir de la combinaison linéaire des six cartes en fréquence de HFI. Elle contient donc des sources provenant de toutes les fréquences. Une première approche simple serait de faire un masque combiné des masques de sources compactes à chaque fréquence. Cependant, à haute fréquence le nombre de sources est très important est cela conduit à une large fraction de pixels masqués. Nous avons utilisé un masque, préparé par Laurence Perotto, qui tient compte du poids des contributions de chaque fréquence à la carte finale dans la combinaison linéaire. Pour cela, dans un premier temps il faut convertir le flux des sources en température CMB, puis appliquer le pseudo-inverse de la matrice de mélange, afin d'estimer le flux équivalent des sources dans la carte finale. Pour finir, on réalise un seuil à  $0.15mK_{CMB}$  et on élimine toutes les sources dont le flux est plus faible que le seuil, ce qui permet de ne conserver que les sources dont le flux résiduel dans la carte GMCA est trop important.

Finalement, pour l'estimation de l'effet de lentille gravitationnelle sur les cartes combinées à 143GHz et 217GHz nous avons utilisé le masque formé par l'union du masque galactique, du masque d'objets compacts et du masque de couverture, associés à chacune des deux fréquences. Pour l'analyse de la carte issue de la séparation des composantes, nous utilisons les masques combinés des sources compactes et de la couverture, décrits ci-dessus et du petit masque galactique MGALGMCA. Nous nous référons à ces masques comme les masques MCOMB143, MCOMB217 et MCOMBGMCA pour les masques combinés associés respectivement aux cartes à 143GHz, 217GHz et à la carte issue de la séparation des composantes avec GMCA.

Selection	Carte 143GHz	Carte 217GHz	Carte GMCA
masque combiné	490	423	486
critère de gaussianité	399	308	X
critère d'isotropie	399	308	X

TABLE 8.3 – **Tableau récapitulatif du nombre de patchs restant après sélection.** Trois sélections successives des patchs sont appliquées afin de ne garder que les patchs dont le signal est dominé par l'émission du CMB. La première sélection est chargée d'éliminer les patchs dont le nombre de pixels masqués représente plus de 5% du nombre total de pixels du patch. Il permet d'éliminer principalement les patchs tombant dans le plan galactique. Le second critère est une ensemble de tests de gaussianité issus d'une analyse Monte-Carlo, visant à supprimer les patchs dont le signal est fortement non gaussien. Ce sont souvent les patchs, possédant de forts résidus de poussière galactique. Enfin un troisième test est chargé d'éliminer les patchs en fonction de leur isotropie. C'est un test beaucoup moins contraignant que les précédents et qui ne conduit au rejet d'aucun patch étant donnée la qualité actuelle du traitement des données.

## 8.4 Sélection des patchs

L'avantage majeur de l'analyse par patchs, réside dans la liberté du choix des patchs. Par exemple, dans le cas où la couverture du ciel n'est pas complète, il est possible de choisir uniquement les patchs tombant dans la zone couverte par les observations, ce qui est bien adapté à des missions qui n'observent pas le ciel complet, ou pour faire une analyse préliminaire en début de mission avant d'avoir une couverture complète. Par ailleurs, même avec des expériences comme PLANCK qui observent la totalité du ciel, il est toujours intéressant de pouvoir éviter les régions les plus contaminées ou problématiques, ou de pouvoir choisir les régions dont la couverture est la plus homogène ou encore de sélectionner des régions couvertes par d'autre expérience pour faire des études croisées... En isolant des groupes particuliers de patchs, il est également possible de faire des comparaisons entre différentes zones d'observations. C'est une technique très puissante pour mettre en évidence des effets systématiques par exemple, ou des effets liés aux contaminations dont la répartition spatiale n'est pas homogène. Finalement, l'étude par patch permet de tester les propriétés statistiques locales à deux dimensions, notamment l'isotropie, ce qu'il est impossible de faire à partir d'une décomposition en harmoniques sphériques qui donnent les propriétés sur la sphère complète.

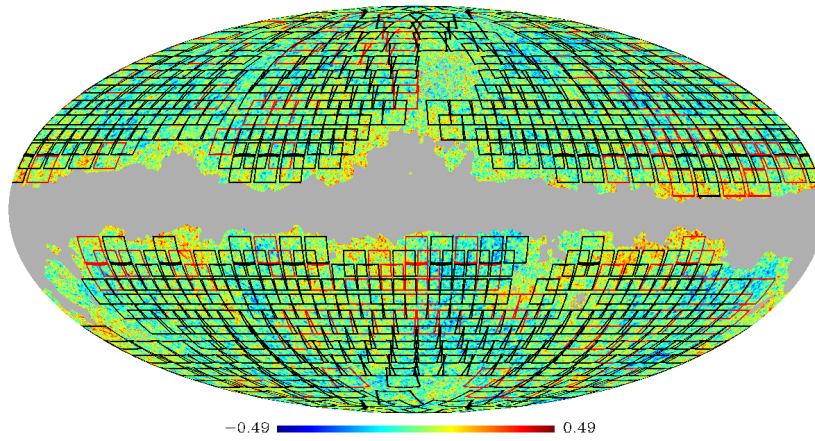
Dans cette section nous détaillons les trois critères de sélection des patchs que nous avons utilisé pour réaliser l'analyse de l'effet de lentille gravitationnelle.

### 8.4.1 Sélection sur le nombre de pixels masqués

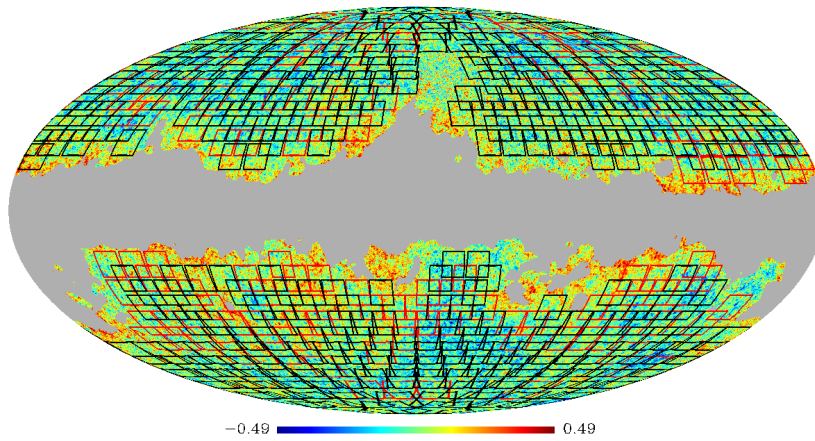
Nous avons réalisé plusieurs mesures sur trois cartes différentes, possédant chacune des masques différents, dans le but de caractériser l'influence de la contamination et des masques notamment. Dans tous les cas nous avons commencé l'analyse sur une sélection de patchs obtenue à partir des masques combinés associés aux cartes de départ. Nous avons rejeté systématiquement les patchs pour lesquels plus de 5% des données étaient masquées afin de s'assurer que l'algorithme d'inpainting local soit efficace. La méthode de sélection est très simple et peut être résumée de la façon suivante, pour chaque patch :

- récupération de la liste des pixels tombant dans le patch
- détermination du nombre de pixels du patch tombant dans le masque combiné
- ajout du patch à la liste si le pourcentage de pixels masqués est inférieur à 5%.

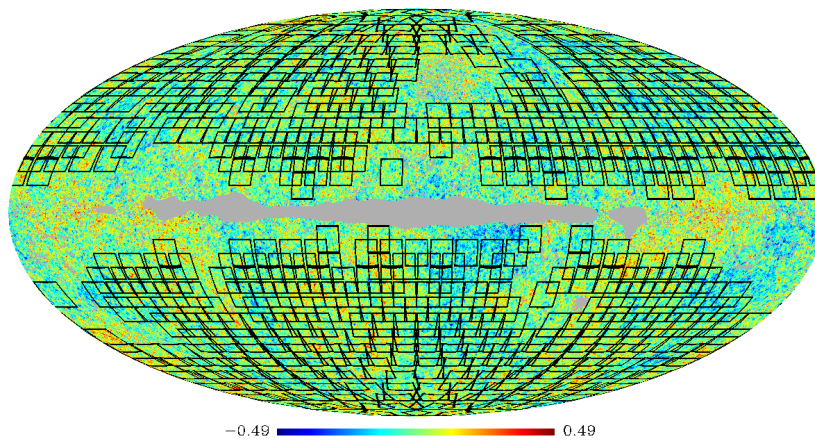
Les listes ainsi obtenues, pour les différentes cartes, constituent le point de départ de chaque analyse. À partir des cartes à 143GHz, à 217GHz et de la carte GMCA, en se basant respectivement sur les masques combinés MCOMB143, MCOMB217 et MCOMBGMCA, on obtient un ensemble de 490, 425 et 486 patchs respectifs satisfaisant le critère précédent (voir le tableau 8.3). La répartition de ces patchs sur la sphère est représentée sur la figure 8.8.



(a) Sélection des patchs sur la carte des observations combinées à 143GHz

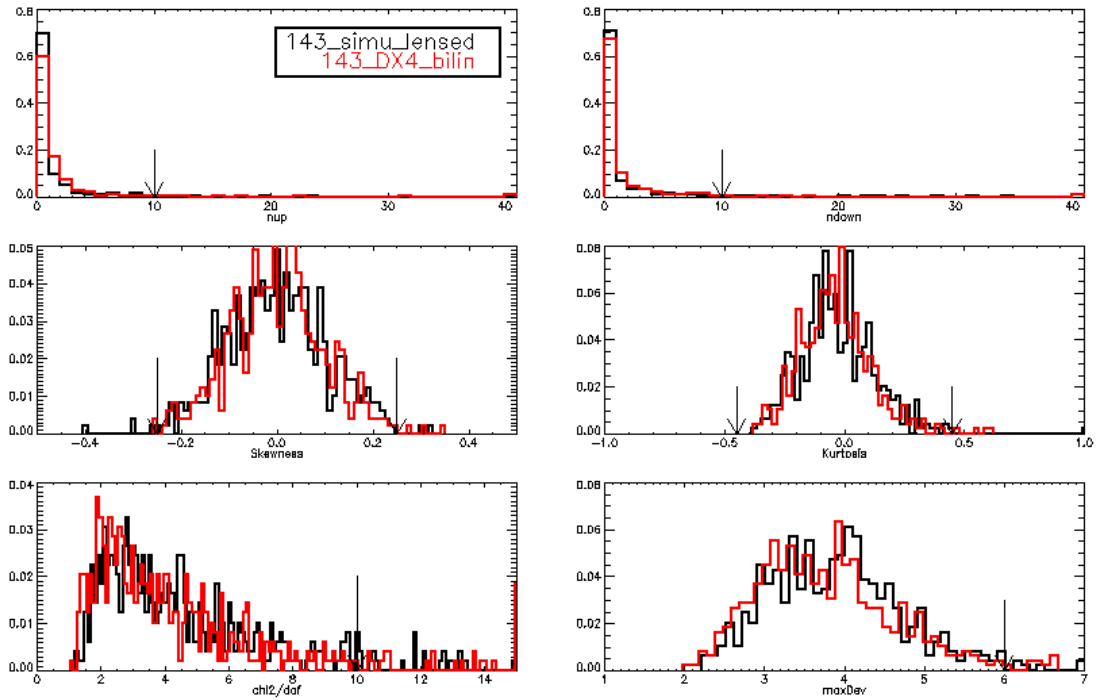


(b) Sélection des patchs sur la carte des observations combinées à 217GHz

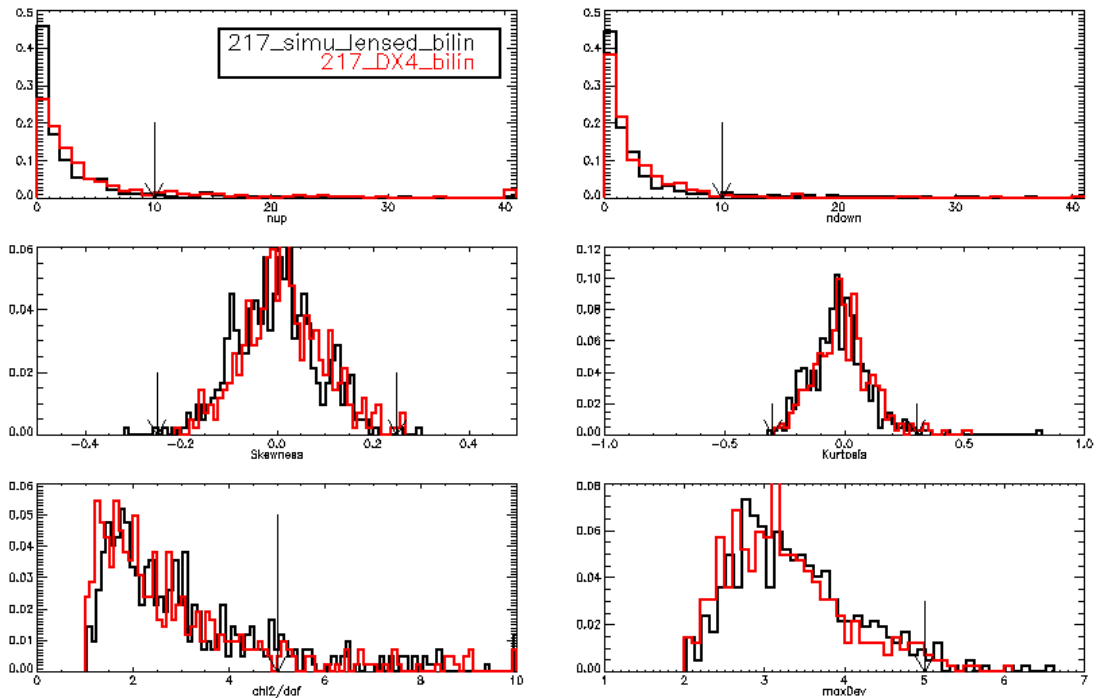


(c) Sélection des patchs sur la carte issue de la séparation des composantes avec GMCA

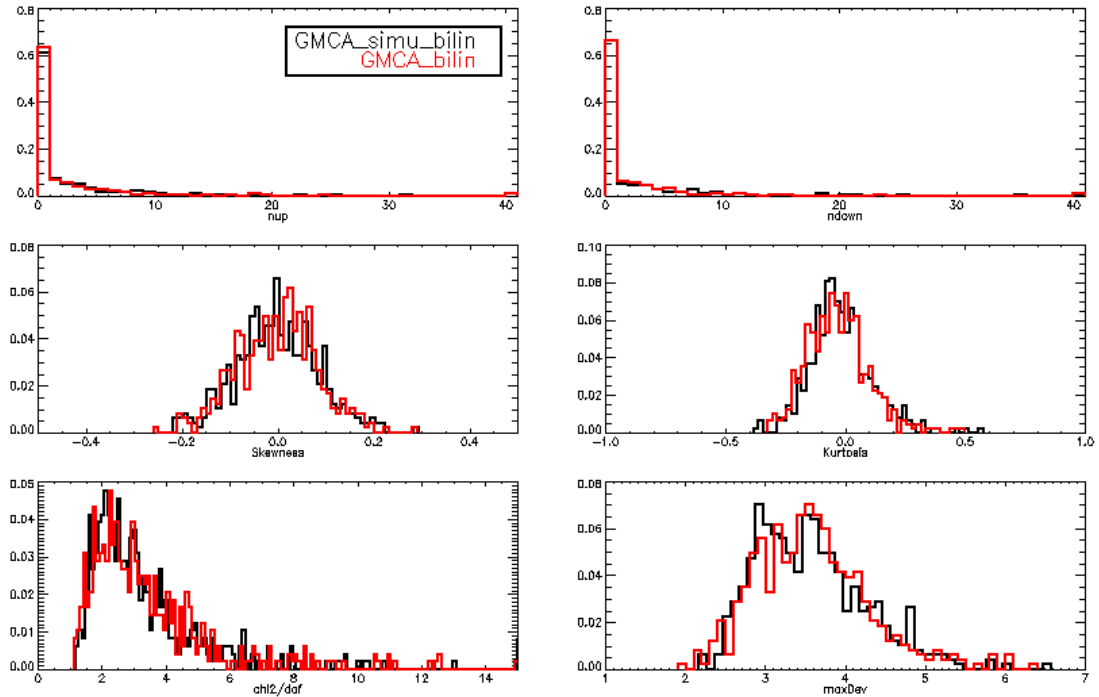
FIGURE 8.8 – **Sélection des patchs.** Trois sélections successives sont appliquées afin d'obtenir la liste finale des patchs pour l'analyse qui correspond aux patchs noirs. En rouge on peut voir les patchs éliminés par la seconde sélection (sélection sur la gaussianité) sur l'ensemble des patchs issus de la première sélection (patchs rouges et noirs). La troisième sélection ne conduit à l'élimination d'aucun patch. On voit que ce sont principalement des patchs provenant des zones proches des régions les plus contaminées, plus une partie de faux positifs.



(a) Distribution attendue, comparée à la distribution de la carte des données 143GHz



(b) Distribution attendue, comparée à la distribution de la carte des données 143GHz



(c) Distribution attendue, comparée à la distribution de la carte GMCA

FIGURE 8.9 – **Distributions des tests de gaussianité.** Les six paramètres sont estimés pour chaque patch du ciel, sur des simulations Monte-Carlo, avec un bruit et un lobe instrumental réalistes dans les trois cas. La distribution de ces paramètres, normalisée à 1, est tracée en rouge. Elle est dominée par la variance cosmique. Les coupures sont données par les flèches. Elles ont été choisies afin de s'assurer de ne pas laisser de patches fortement non gaussiens, tout en limitant le nombre de faux positifs. Pour la carte issue de la séparation des composantes ((c)), on ne voit pas de différences significatives de la distribution (en noir), par rapport à la distribution estimée sur les simulations Monte-Carlo. Nous n'avons donc pas appliqué de coupure. Pour les cartes combinées à 143GHz et 217GHz, les coupures conduisent à la suppression de quelques dizaines de patches dans les deux cas.

### 8.4.2 Sélection sur des critères de gaussianité

Les régions tombant dans des zones fortement contaminées par l'émission galactique ou de sources ponctuelles sont masquées. Cependant les contours du masque galactique sont définis par une coupure en amplitude, l'émission de la poussière s'étend donc au delà du masque, même si l'amplitude attendue dans ces régions est faible. Par ailleurs un ensemble de sources à la limite du seuil de détection peut également produire un effet non négligeable s'il n'est pas mis en évidence. À partir de la liste de patches restant après le test précédent, nous réalisons donc une sélection plus fine, basée sur les propriétés observées dans les données de chaque patch individuellement.

On sait que la présence de résidus galactiques ou de sources compactes, modifie les propriétés statistiques locales du signal, comparées à celles de l'émission du CMB seule, notamment la gaussianité du signal. En partant de cette constatation, nous avons cherché une procédure permettant d'éliminer les patches les plus contaminés. Nous nous basons pour cela sur une série de six tests, permettant de s'assurer que le niveau de gaussianité des données contenues dans chaque patch est comparable à celui attendu pour le CMB.

Parmi les tests effectués, quatre se basent sur les propriétés de l'histogramme des pixels contenus dans un patch. La procédure que nous avons employée, consiste alors dans un premier temps à calculer l'histogramme de chaque patch puis à déterminer l'écart type  $\sigma$  du meilleur ajustement d'un modèle gaussien par la méthode des  $\chi^2$ . À partir de ces éléments, on détermine ensuite les propriétés suivantes :

- nombre de points supérieurs à  $4\sigma$
- nombre de points inférieurs à  $-4\sigma$
- $\chi^2$  par degré de liberté
- déviation maximale des points de l'histogramme par rapport à l'ajustement gaussien

Dans un second temps nous calculons également les propriétés suivantes sur l'ensemble des pixels de chaque patch :

- kurtosis de la population
- coefficient de dissymétrie de la population

Nous utilisons finalement un jeu de simulations Monte-Carlo afin d'évaluer la distribution des valeurs des différentes propriétés pour un signal du CMB non contaminé, avec un lobe et bruit instrumental comparables à ceux des données. Nous en déduisons un ensemble de coupures permettant d'exclure les patches les plus contaminés (voir figure 8.9).

Afin de ne pas modifier la valeur du signal, nous travaillons directement sur les pixels de la sphère, sans projection. Nous récupérons dans un premier temps la liste des pixels tombant dans chaque patch, puis nous calculons l'histogramme du signal à l'intérieur du patch. La première sélection, décrite dans la section 8.4.1, ne supprime que les patches possédant plus de 5% des données masquées. Un grand nombre de patches restants possèdent donc, tout de même, une certaine fraction de pixels masqués qui ne sont pas pris en compte dans le calcul de l'histogramme.

Il est important de souligner que la sélection mise en œuvre dans cette section ne vise pas la détection de non gaussianités intrinsèques à l'émission du CMB, son but est simplement d'éliminer les zones du ciel dont la contamination est la plus importante, grâce à leur signature sur l'histogramme des données. On peut voir sur la figure 8.8 (en rouge) que les patches rejetés se situent majoritairement en périphérie du masque galactique et correspondent donc à des zones contaminées par les émissions galactiques.

Les coupures déterminées par Monte-Carlo assurent d'éliminer les patches fortement non gaussiens. Cependant, il faut garder à l'esprit que pour un ensemble de patches donné, même issus de cartes idéales, on s'attend à avoir une certaine distributions des valeurs des propriétés (voir figure 8.9), ce qui donne lieu à une certaine probabilité d'obtenir un patch gaussien mais dont les valeurs dépassent les coupures. Il faut donc trouver un compromis entre l'efficacité du test à rejeter un patch contaminé et le nombre de patches non contaminés qu'on s'autorise à perdre. Dans tous les cas, parmi les patches rejetés par ce test, quelques pourcents sont des faux positifs et ne correspondent donc pas à des patches contaminés. Il est toutefois impossible de les distinguer. Cependant, la conséquence sur l'analyse de la perte d'un patch non contaminé se réduit à une faible diminution du rapport signal sur bruit qui est négligeable étant donné le nombre de patches et comparée aux effets induits par un patch contaminé qui passerait les tests. La répartition des patches restants après cette sélection est visible sur la figure 8.8.

Pour la carte issue de la séparation des composantes, la contamination en dehors du masque est

si faible que la distribution des mesures de gaussianité sur les données est similaire à celle déduite du Monte-Carlo. Nous n'avons donc pas défini de coupure dans ce cas et nous conservons l'ensemble des patchs issus de la sélection précédente. Pour les cartes à 143GHz et à 217GHz, la sélection précédente ayant éliminé les zones les plus contaminées, on s'attend à ce que les régions les moins contaminées aient des propriétés proches de celle des cartes simulées. C'est ce qu'on peut voir sur les figures 8.9(a) et 8.9(b) où on remarque que la distribution globale est proche de celle du Monte-Carlo. Les coupures sont toutefois nécessaires dans ce cas et conduisent à la suppression de plusieurs dizaines de patchs comme on peut le voir dans le tableau 8.3.

### 8.4.3 Sélection sur des critères d'isotropie

On sait que l'émission des autres composantes que le CMB peut être, localement, fortement anisotrope. Par ailleurs, des problèmes dans le traitement des données peuvent produire des artefacts, notamment de larges stries qui brise l'isotropie localement. Nous avons donc développé une méthode de sélection s'appuyant sur les propriétés d'isotropie du CMB.

Le test d'isotropie est effectué dans l'espace de Fourier à partir du spectre de puissance angulaire à deux dimensions. En effet des directions privilégiées dans une image ressortent en général fortement dans le spectre de puissance dans une direction perpendiculaire à celle de l'espace réel. Ainsi, après avoir calculé le spectre de puissance angulaire à deux dimensions, on applique une décomposition en ondelettes afin de supprimer toutes les variations à haute fréquence. L'isotropie est finalement évaluée en comparant l'amplitude dans huit quadrants réguliers.

Ce test est appliqué après la sélection sur la gaussianité qui est plus efficace pour détecter les contaminations. Finalement étant donnée la qualité des données actuelles, cette sélection ne conduit à aucun rejet de patchs. Nous la conservons tout de même car cela permet de détecter d'éventuelles régressions dans la procédure de traitement des données de la mission.

## 8.5 Analyse préliminaire

Dans cette section nous présentons les traitements préliminaires que nous avons réalisés afin de pouvoir estimer l'effet de lentille gravitationnelle. Ce travail se résume à trois étapes principales, la découpe des patchs, la déconvolution et l'estimation du spectre de puissance des anisotropies de température nécessaire à l'estimateur de la deflexion.

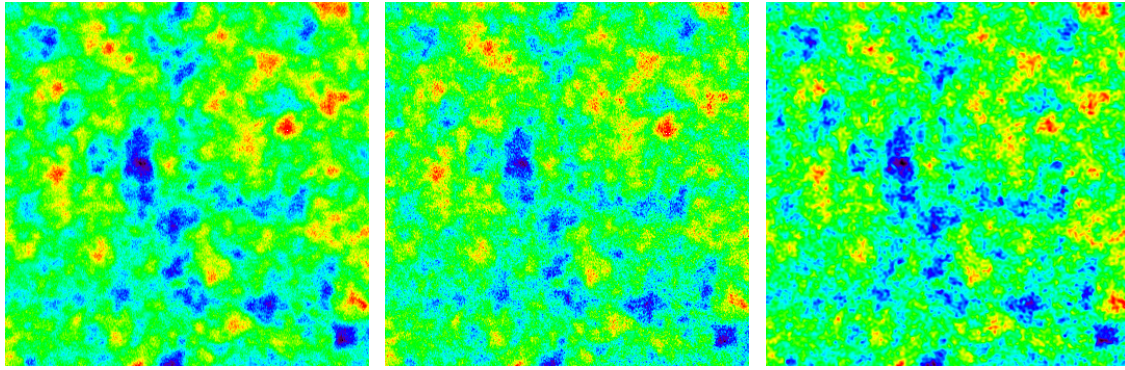
### 8.5.1 Découpe des patchs

Nous avons utilisé l'outil développé dans la section 7.1 pour découper les cartes du ciel en patchs carrés, de 10 degrés de côté. Pour chacune des trois cartes, la carte des détecteurs combinés à 143GHz et à 217GHz ainsi que la carte issue de la séparation de composantes avec GMCA, nous avons utilisé la même procédure avec les mêmes arguments. Seule la liste initiale de patchs, obtenue après les différentes sélections, diffère. Les principales étapes peuvent être résumées de la façon suivante :

- choix d'un patch dans la liste après sélection
- détermination du centre du patch
- projection gnomonique des pixels de la sphère dont la distance angulaire par rapport au centre du patch est inférieure à la distance de la demi-diagonale du patch.
- sélection des pixels tombant dans le patch
- apodisation
- estimation des coefficients de Fourier sur une grille régulière de 120 par 120 à partir des points projetés, tombant sur une grille irrégulière, à l'aide de la méthode des gradients conjugués (voir B).

Les centres des patchs sont donnés par les centres des pixels d'une carte HEALPIX avec un paramètre *nside* de 8. Seuls les patchs restants après les différentes sélections ont été découps. Comme le niveau de contamination dépend de la fréquence d'observation, nous n'avons pas le même nombre de patchs sur chaque carte. On obtient alors, pour chacune des trois cartes de départ, un ensemble de patchs qui sont la base de l'analyse. Sur la figure 8.10 on peut voir une zone du ciel découpée dans les différentes cartes (avec interpolation bilinéaire et sans apodisation). Elle





(a) Patch découpé dans la carte des observations combinées à 143GHz (b) Patch découpé dans la carte des observations combinées à 217GHz (c) Patch découpé dans la carte issue de la séparation des composantes avec GMCA

FIGURE 8.10 – **Exemple de patch découpé dans les données.** Les patches font 10 degrés de coté. Ils sont découpés dans une zone centrée sur le centre du pixel 147 d'une carte HEALPIX de paramètre *nside* égal à 8, ce qui correspond à  $35.7^\circ$  de latitude et  $33.75^\circ$  de longitude. La découpe a été effectuée sans apodisation afin de pouvoir distinguer les différences entre les cartes initiales. Par ailleurs, la méthode par interpolation linéaire a été utilisée, afin d'accélérer la découpe. Les plus grandes structures visibles correspondent au CMB et sont très semblables d'une carte à l'autre. Les différences visibles sont produites par les différences de taille de lobe et de niveau de bruit entre les détecteurs aux différentes fréquences.

correspond au patch centré sur le pixel HEALPIX numéro 147, ce qui correspond à  $35.7^\circ$  de latitude et  $33.75^\circ$  de longitude.

Pris deux à deux, les patches ont un recouvrement moyen proche de 25%. Pour les cartes combinées à 143GHz et 217GHz et la carte GMCA, la surface totale pour tous les patches découpés correspond respectivement à 62.64%, 49.82% et 70.82% de la surface de la sphère soit environ 25841, 20552 et 29215 degrés carrés respectivement. Pour comparaison, une expérience comme ACT couvre une surface de quelques centaines de degrés carrés seulement.

## 8.5.2 Déconvolution des effets du lobe de l'instrument

Les cartes du ciel produites par PLANCK sont des cartes convoluées par la réponse impulsionnelle spatiale de l'instrument. Comme on l'a vu dans la section 6.4.1 l'effet du lobe de l'instrument entraîne une perte de puissance aux petites échelles angulaires. L'estimateur du potentiel gravitationnel intégré nécessite une carte déconvoluée. Il est possible de déconvoluer la carte du ciel avant la découpe des patches. Cependant, comme on l'a vu dans la section 6.4 le bruit issu de l'électronique de lecture n'est pas convolué car il apparaît après dans la chaîne de détection. Ainsi lorsqu'on déconvolue une carte, le bruit se trouve également déconvolué, ce qui a pour effet de faire exploser sa puissance aux petites échelles angulaires.

Nous utilisons la détermination de la taille du lobe de l'instrument effectuée par Antoine Chamballu<sup>4</sup>, basée sur la première observation de la planète Mars. Les chiffres donnés représentent la meilleure estimation en supposant que le lobe de l'instrument est un lobe gaussien circulaire. On sait que cela représente une bonne approximation, même si dans l'idéal, des corrections non gaussiennes et une ellipticité devraient être prises en compte. Cependant, l'hypothèse précédente facilite grandement la déconvolution des données. En effet dans le cas d'un lobe gaussien circulaire, on est capable de calculer de façon analytique la réponse impulsionnelle spatiale de l'instrument. Par ailleurs, l'effet du lobe sur le spectre de la déflection, se manifeste principalement aux très bas  $\ell$  (voir Hanson et al. (2010b)), auxquels nous n'avons pas accès avec une analyse par patch à cause de l'approximation plane. De plus, la combinaison des observations de plusieurs détecteurs et de plusieurs passages sur le ciel tend à symétriser le lobe.

Une déconvolution, idéale tenant compte des non-gaussianités du lobe et de son élongation, est une opération très lourde et qui va au delà du cadre de cette thèse. Elle nécessite de tenir compte de

4. Voir note interne <http://wiki.planck.fr/index.php/Proc/101017BeamEstimates>

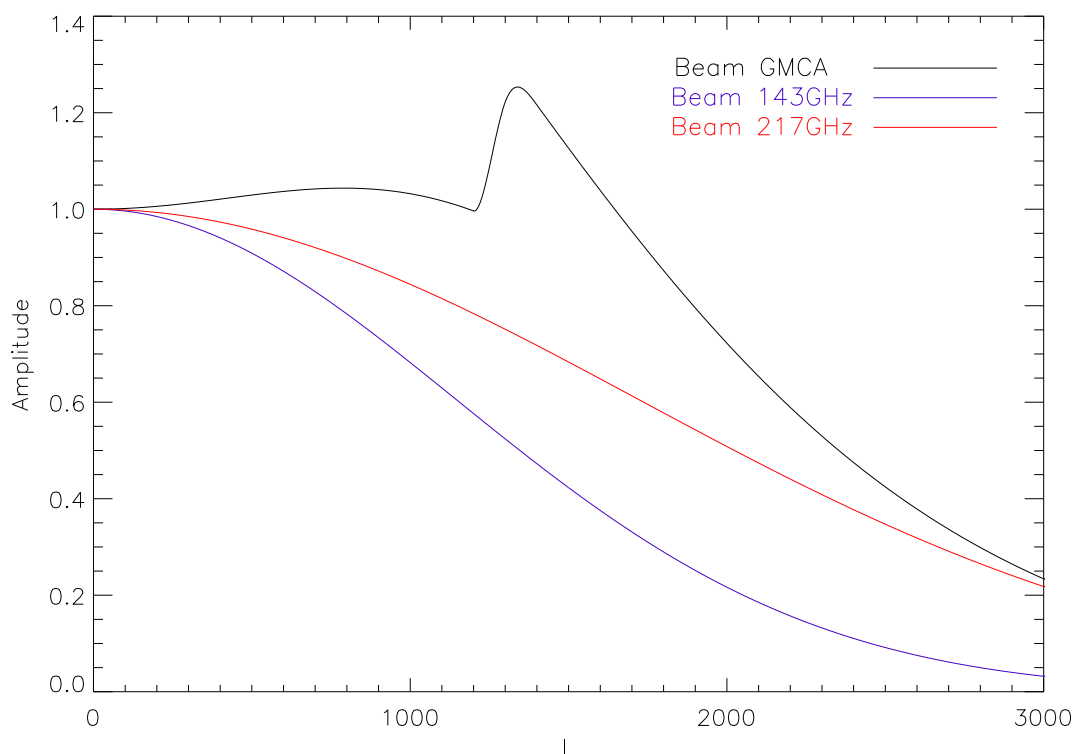


FIGURE 8.11 – **Profil des lobes des instruments dans l'espace de Fourier.** Les lobes pour les cartes à 143GHz et 217GHz sont des lobes gaussiens circulaires. Pour la carte issue de GMCA c'est une combinaison linéaire des lobes aux différentes fréquences pondérée par la matrice de mélange issue de la séparation des composantes. On voit bien dans tous les cas, la perte de puissance aux petites échelles induite par l'optique. Elle est d'autant plus importante que le lobe est grand.

la direction d'observation en chaque point et de réaliser la déconvolution échantillon par échantillon. C'est une opération qui n'est pas envisagée pour PLANCK sans approximations. Les résultats des mesures effectuées par Antoine Chamballu sont résumés dans le tableau 8.4 qui récapitule la meilleure estimation de la largeur à mi-hauteur du lobe de l'instrument pour chaque fréquence.

La formule de déconvolution d'un lobe gaussien circulaire de déviation standard  $\sigma$ , est donnée dans la section 6.4.1. Les calculs peuvent être effectués simplement dans l'espace de Fourier. Nous réalisons ainsi cette opération pour chaque patch découpé dans la sphère, afin de pouvoir faire l'analyse de l'effet de lentille gravitationnelle.

Pour la carte issue de la séparation des composantes, le lobe effectif est plus complexe que celui des cartes à 143GHz et 217GHz. En effet, la carte issue de GMCA est une combinaison des six cartes en fréquence de HFI, qui possèdent chacune un lobe différent. Connaissant la matrice de mélange et son pseudo-inverse (voir section 8.1.4), il est toutefois possible d'estimer le lobe effectif résultant de la combinaison des cartes. Étant donnée la linéarité du système permettant d'obtenir les cartes des différentes composantes, le lobe effectif s'obtient par une simple combinaison linéaire des lobes effectifs aux différentes fréquences :

$$B_{\text{GMCA}}^{\text{CMB}}(\ell) = \sum_{\nu} A_{\nu i}^+ A_{\nu i} B_{\nu}(\ell) \quad (8.5)$$

où  $A_{\nu i}$  et  $A_{\nu i}^+$  représentent respectivement la matrice de mélange du système et son pseudo-inverse,  $\nu$  correspond à l'indice de la fréquence considéré et  $i$  à l'indice de la composante qui nous intéresse, c'est-à-dire le CMB dans notre cas. On déconvolue finalement la carte issue de GMCA, à partir du lobe effectif  $B_{\text{GMCA}}^{\text{CMB}}(\ell)$ , de façon similaire aux cartes en fréquence. Les lobes effectifs utilisés pour déconvoluer les trois cartes utilisées pour notre analyse, sont tracés sur la figure 8.11. On remarque

Fréquence (Giga Hertz)	100	143	217	353	545	857
Taille du lobe (fwhm [minute d'arc])	9.53	7.08	4.71	4.50	4.72	4.42

TABLE 8.4 – **Tableau récapitulatif de la taille du lobe de l'instrument en fonction de la fréquence.** Ces mesures représentent la meilleure estimation d'un lobe gaussien elliptique obtenue à partir des données de la première observation de la planète Mars par Antoine Chamballu et al. La taille du lobe est la plus importante à basse fréquence et la plus faible pour les voies à haute fréquence.

que la forme du lobe effectif de la carte issue de la séparation de composantes est relativement complexe, cependant le lobe reste toujours gaussien car il correspond à une combinaison linéaire de lobes gaussiens. Sa forme particulière en fonction de  $\ell$ , ne dépend que des coefficients de la matrice de mélange mais ne dépend pas du choix de l'algorithme utilisé pour déterminer cette matrice.

### 8.5.3 Estimation du spectre de puissance de température

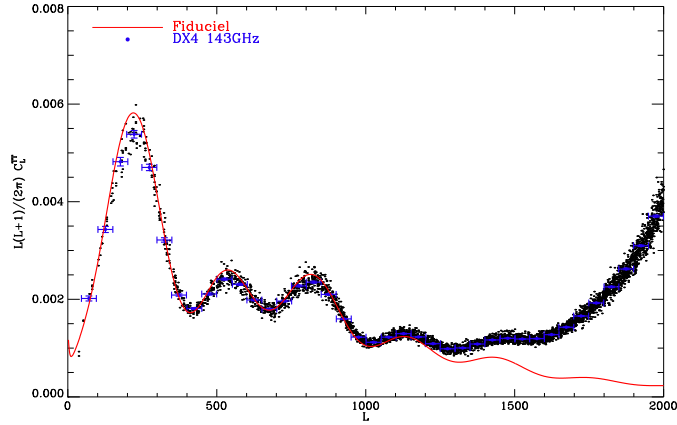
Dans la section 6.6 nous avons présenté l'estimateur du potentiel de lentille de Okamoto et Hu. Nous avons vu qu'il nécessite en paramètre, le spectre de puissance observé sur une carte déconvoluée du lobe de l'instrument. Nous estimons ce spectre à partir de l'ensemble des patchs découpés puis déconvolués des effets du lobe de l'instrument. Nous employons pour cela la méthode expliquée dans la section 6.5.2. Dans un premier temps on commence par estimer le spectre de puissance à deux dimensions, la coaddition des modules carré des coefficients de Fourier de l'ensemble des patchs par analogie avec le périodogramme de Welsh à une dimension. Le spectre de puissance est alors moyenné sur des couronnes en fonction du module du vecteur de Fourier afin d'obtenir l'estimation finale du spectre de puissance angulaire. On peut voir sur la figure 8.12 les spectres obtenus sur les trois cartes que nous avons analysées, comparés au spectre calculé avec CAMB à partir de l'estimation des paramètres cosmologiques donnés par WMAP, après sept ans d'observations. Comme aucun traitement du bruit n'a été effectué la mesure est dominée à haut  $\ell$  par l'explosion du bruit déconvolué. Pour chacune des trois mesures, les barres d'erreur sont déduites de la dispersion sur un ensemble d'estimations du spectre de puissance sur 100 cartes, possédant un bruit et un lobe réalistes (voir section 8.7.1) et représentent l'erreur attendue pour une réalisation.

La mesure précédente est uniquement adaptée à une analyse dédiée à l'estimation du spectre de puissance de déflexion à l'aide de l'estimateur quadratique de Okamoto et Hu, qui nécessite une mesure du signal plus du bruit décorrélé. Si on s'intéresse à une mesure du spectre de puissance des anisotropies de température elles-mêmes, il est nécessaire de choisir une méthode plus adaptée. Il est notamment plus judicieux, dans ce cas, de ne pas déconvoluer les données afin d'éviter l'explosion du bruit aux petites échelles angulaires. Dans ces conditions, on peut par exemple estimer le spectre de puissance sur les cartes convoluées, puis soustraire une estimation du bruit avant de prendre en compte les effet du lobe de l'instrument au niveau du spectre. Le spectre de puissance du bruit peut alors être obtenu dans une première approximation par *Jack-Knife* (voir section 8.7.1).

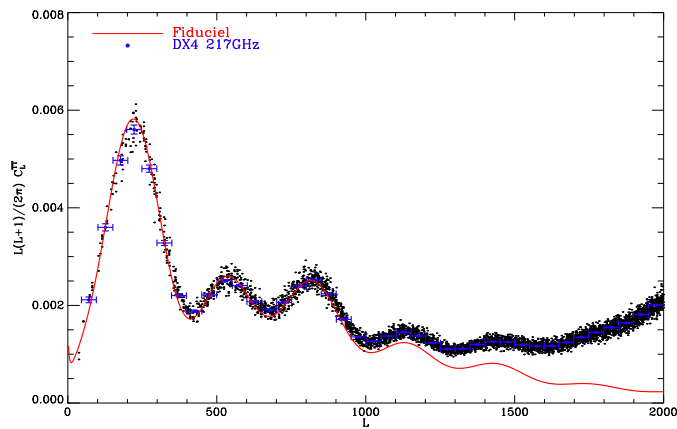
## 8.6 Estimation du spectre de la déflexion

### 8.6.1 Estimation du potentiel de lentille

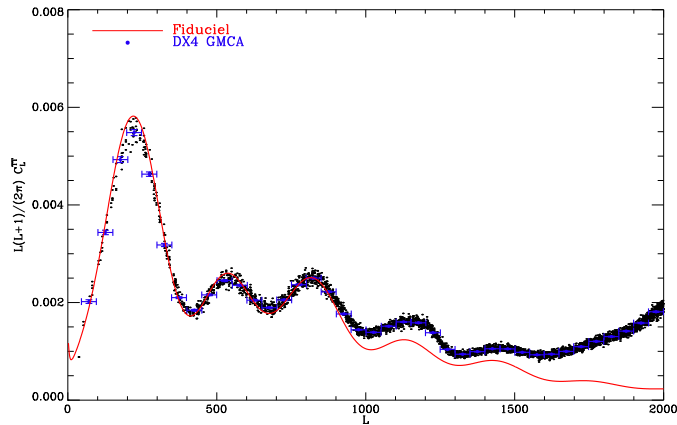
L'estimation du spectre de puissance de la déflexion se fait en plusieurs étapes. Dans un premier temps, nous estimons le potentiel de lentille, ou champ  $\phi$ , sur chaque portion du ciel, grâce à l'estimateur quadratique dans le plan, basé sur les travaux de Hu et Okamoto, que nous avons développé (voir section 6.6.2). Cette estimation est basée sur un produit de différents modes de Fourier des cartes de température, associé à un filtrage en fréquence. La borne de l'intégrale de l'équation 6.32, est limitée à  $\ell = 2000$ . Il n'est pas nécessaire d'aller au-delà car étant donnée la taille du lobe de l'instrument, on commence alors à être dominé par le bruit déconvolué. On ne gagne donc pas en rapport signal sur bruit pour l'estimation du spectre de la matière. Par ailleurs, les bornes de l'intégrale pour le calcul de la normalisation et des biais de l'estimateur doivent être identiques au choix précédent et sont donc également fixées à 2000.



(a) Spectre de puissance de température par patch, sur la carte des observations combinées à 143GHz



(b) Spectre de puissance de température par patch, sur la carte des observations combinées à 217GHz



(c) Spectre de puissance de température par patch, sur la carte GMCA

FIGURE 8.12 – **Spectre de puissance de température brut, par patch, estimé sur les données déconvoluées.** Le spectre de puissance est estimé par la coaddition de l'ensemble des spectres de puissance angulaire à deux dimensions des patches découpés dans une carte de température. On peut observer l'explosion exponentielle de la puissance du bruit instrumental à haut  $l$ , causée par la déconvolution des données. Plus le lobe de l'instrument est important plus l'explosion apparaît tôt. La carte à 143GHz est donc celle où l'effet est le plus marqué. Les barres d'erreurs sont estimées, à partir de la variance de la mesure sur 100 cartes issues du Monte-Carlo décrit dans la section 8.7 et représentent les variations attendues pour une réalisation. En noir on a tracé les points estimés un à un pour donner une idée de la variance. Les valeurs en bleu sont moyennées dans des intervalles de 50, on les compare au spectre fiduciel en rouge qui est calculé avec CAMB à partir des paramètres cosmologiques estimés par WMAP, intégrant sept ans de données.

Comme on peut le voir dans l'équation 6.40, le filtre dépend du spectre de puissance de température du CMB non lentillé, ainsi que le spectre de puissance de la carte observée, avec son bruit et déconvolué des effets du lobe de l'instrument. Le spectre de puissance de température non lentillé, correspond à celui que l'on pourrait mesurer en l'absence de tous les autres effets extérieurs, c'est-à-dire le spectre de puissance des anisotropies à l'époque de la recombinaison. Bien évidemment, la mesure faite à l'heure actuelle inclue l'ensemble des effets modifiant les propriétés statistiques, incluant l'effet de lentille gravitationnelle. Nous utilisons donc le spectre théorique, calculé avec CAMB, à partir de la meilleure estimation des paramètres cosmologiques par l'équipe de WMAP, après sept ans d'observation (Larson et al. (2010)). Pour le spectre de puissance des données nous utilisons le spectre mesuré dans la section précédente, à l'aide de la méthode par patch, qui correspond au spectre du signal et du bruit, déconvolué du lobe de l'instrument.

Finalement, l'estimateur du potentiel de lentille est appliqué à chaque patch découpé sur la sphère (voir section 8.5.1). On obtient alors pour chacune des trois cartes que nous analysons (les cartes combinées à 143GHz et 217GHz et la carte GMCA) un ensemble de cartes du potentiel de lentille à deux dimensions.

### 8.6.2 Estimation du spectre de puissance angulaire de la déflexion

La seconde étape de l'analyse consiste à estimer le spectre de puissance brut de la déflexion. À partir de l'ensemble des cartes à deux dimensions du potentiel de lentille, estimées sur chaque morceau découpé de la sphère, nous calculons le spectre de puissance angulaire de la même manière que celui des anisotropies de température. Nous réalisons donc la coaddition des cartes d'amplitude des coefficients de Fourier à deux dimensions par analogie avec le périodogramme de Welch. À cause de la couverture inhomogène du ciel par PLANCK, le rapport signal sur bruit varie avec la direction du ciel considérée. Ainsi, le rapport signal sur bruit des différentes estimations du potentiel de lentille varie d'une zone découpée à l'autre mais nous n'en tenons pas compte dans le calcul de la moyenne. Les effets du bruit inhomogène sont pris en compte par les corrections estimées par Monte-Carlo. On obtient finalement un spectre de puissance à deux dimensions qu'on moyenne dans des couronnes de largeur 50 afin de diminuer les erreurs statistiques sur l'estimation du spectre de puissance angulaire ainsi obtenue.

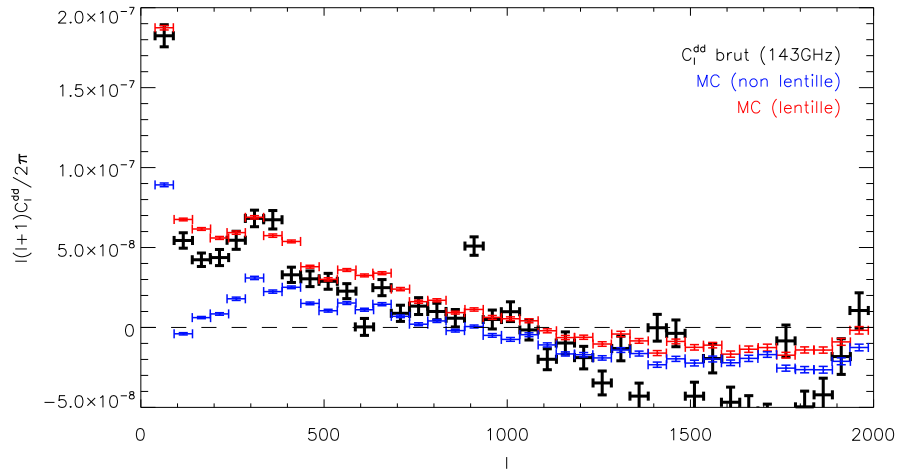
Les travaux de Okamoto et Hu ont montré la présence d'un bruit dans l'estimation du potentiel de lentille produit par la réalisation du CMB elle-même. Le biais introduit par ce bruit sur le spectre de puissance ( $N_K^{(0)}$ ) a été calculé et peut être soustrait (voir section 6.6). Après soustraction de ce biais on obtient une estimation brute du spectre de puissance du potentiel de lentille  $C_\ell^{\phi\phi}$ . Cette mesure ne tient pas compte des bruits additionnels corrélés au champ  $\phi$ , ni des effets systématiques de l'ensemble de la chaîne d'analyse.

En général on ne représente pas directement le spectre du potentiel de lentille mais le spectre de la déflexion  $C_\ell^{dd} = \ell * (\ell + 1) C_\ell^{\phi\phi}$ . L'estimation brute du spectre de la déflexion pour chacune des trois cartes analysées, est présentée sur la figure 8.13. Les erreurs sont estimées à partir de la dispersion entre 100 réalisations issues du Monte-Carlo détaillé dans la section 8.7.4 et représentent l'erreur attendue pour une réalisation. À ce niveau, le signal est dominé par les effets systématiques qui sont pris en compte dans la section 8.7. Les résultats sont comparés aux spectres mesurés à partir de la moyenne des spectres des 100 cartes du Monte-Carlo avec un signal de lentille gravitationnelle et avec le jeu de 100 cartes sans signal. Pour les multipôles inférieurs à 500, les données sont mieux reproduites par les simulations contenant un signal.

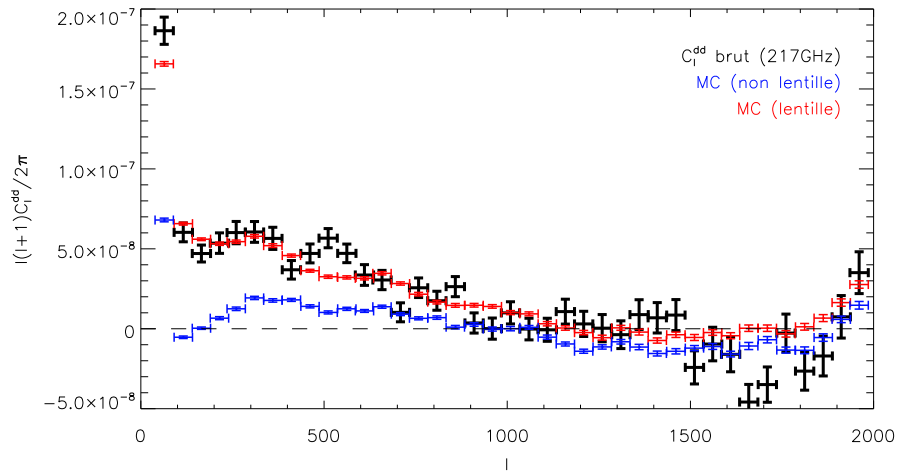
### 8.6.3 Correction des biais de l'estimateur

Les spectres de déflexion estimés dans la section précédente, sont des spectres bruts. Afin d'obtenir une mesure fiable, dans un premier temps, il faut corriger des biais de l'estimateur qui sont connus et qui ont été prédits analytiquement (voir section 6.6.3).

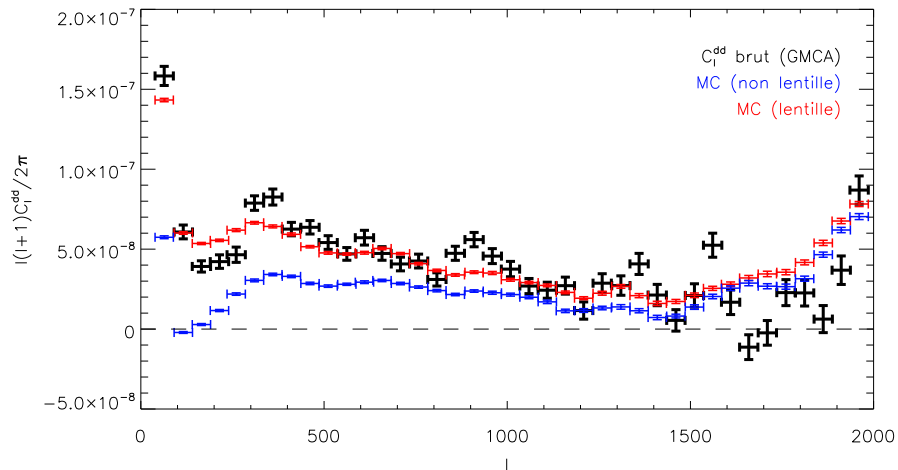
D'après les équations 6.45 et 6.46, le calcul des biais  $N_K^{(1)}$  et  $N_K^{(2)}$  nécessite de connaître le spectre non lentillé mais également le spectre du signal des cartes sur lesquelles sont faites les estimations, c'est-à-dire le spectre lentillé avec bruit déconvolué. Nous utilisons les mêmes mesures que pour l'estimation du potentiel de lentille : le spectre non lentillé calculé à partir des données de WMAP et le spectre lentillé mesuré par la méthode par patch sur les données dans la section 8.5.3.



(a) Spectre de la déflexion brut pour la carte à 143GHz



(b) Spectre de la déflexion brut pour la carte à 217GHz



(c) Spectre de la déflexion brut pour la carte GMCA

FIGURE 8.13 – **Estimation brute du spectre de puissance de la déflexion.** Sur les trois graphiques ci-dessus, on compare le spectre de puissance de la déflexion, brut, sans les corrections (en noir), aux spectres estimés à partir de la moyenne sur un jeu de 100 simulations contenant un signal de lentille gravitationnelle (en bleu) et sur le jeu équivalent ne contenant pas de signal (en rouge). Les erreurs sur les spectres mesurés sur les données de PLANCK correspondent aux erreurs statistiques attendues pour une réalisation et ont été estimées à partir de la dispersion sur les 100 estimations sur les cartes lentillées.

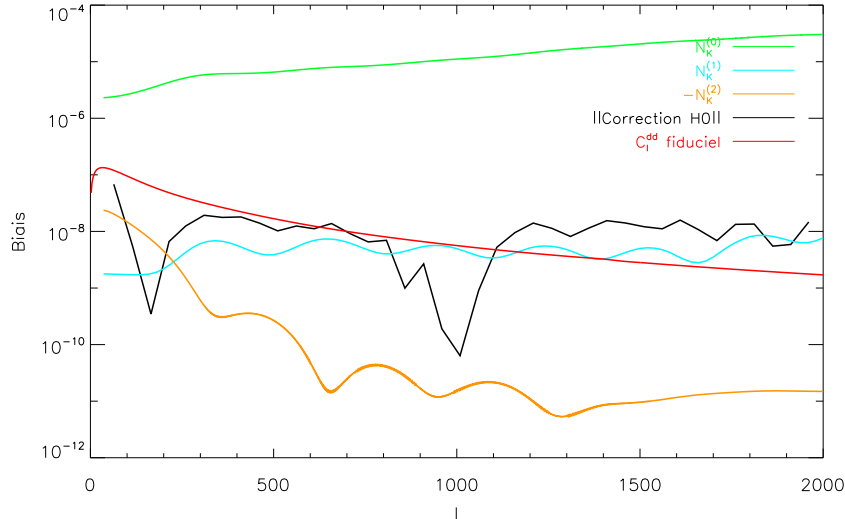


FIGURE 8.14 – **Biais de l'estimateur pour la carte à 217GHz.** Les biais dominants de l'estimateur du spectre de puissance de la déflexion ont été calculés analytiquement. Le terme de biais principal est  $N_K^{(0)}$  (en vert). Il est bien connu et peut être facilement retiré. Dans un second temps, on peut voir que  $N_K^{(2)}$  (en orange) est dominant en terme d'amplitude aux petits  $l$ .  $N_K^{(1)}$  (en bleu) domine le terme précédent aux grands  $l$  et représente un biais important en pourcentage du signal (en rouge) à ces échelles. Le biais additionnel introduit par les effets systématiques que nous avons pris en compte (en noir) est du même ordre de grandeur que  $N_K^{(1)}$ .

Enfin,  $N_K^{(2)}$  dépend également du spectre de la déflexion ce qui peut poser un problème car nous ne disposons à l'heure actuelle d'aucune mesure de ce spectre. Nous utilisons donc le spectre calculé avec CAMB dans un modèle  $\Lambda\text{CDM}$ , à partir des paramètres cosmologiques déterminés par l'équipe de WMAP avec sept ans de données. C'est une bonne approximation car  $N_K^{(2)}$  a une amplitude relativement faible et contribue surtout à bas  $l$ . De plus les variations attendues pour un spectre de la déflexion légèrement différent sont petites. Finalement on a également une justification à posteriori avec la mesure finale qui est fortement compatible avec celle attendue par WMAP étant données les barres d'erreurs de la mesure. Pour la mesure finale, on peut imaginer une méthode itérative permettant de recalculer  $N_K^{(2)}$  à partir des valeurs estimées, cependant l'erreur attendue sur  $N_K^{(2)}$  est à l'heure actuelle dominée par les autres effets systématiques. Sur la figure 8.14 on peut voir l'amplitude des biais de l'estimateur  $N_K^{(0)}$ ,  $N_K^{(1)}$  et  $N_K^{(2)}$  obtenus pour la carte des données combinées à 217GHz, comparée à la valeur du spectre de puissance attendu à partir des paramètres cosmologiques déterminés par WMAP. On a également représenté l'amplitude du biais introduit par les effets systématiques que nous avons pris en compte. Ce biais est calculé à partir de simulations Monte-Carlo dans la section suivante. On peut voir qu'il est du même ordre de grandeur que le terme  $N_K^{(1)}$ .

## 8.7 Correction des effets systématiques par Monte-Carlo

Dans le cas d'une expérience idéale, dont le signal est constitué d'une émission du CMB purement gaussienne convoluée par un lobe gaussien circulaire, avec un bruit instrumental blanc, l'estimateur développé par Okamoto & Hu (2003b) permet de reconstruire le spectre de puissance angulaire de la déflexion avec une bonne précision si on tient compte des biais additionnels calculés dans Kesden et al. (2003) et Hanson et al. (2011) (Voir section 6.6.3). C'est le travail réalisé dans la section précédente.

Dans le cas d'une expérience réelle, Pour obtenir une estimation précise du spectre de la déflexion, il faut corriger des effets systématiques. En revanche, il n'existe pas de méthode analytique simple pour prendre en compte l'effet de l'apodisation, du bruit inhomogène, de l'asymétrie du lobe, du

masque, de l'inpainting, de la séparation des composantes, des résidus de rayons cosmiques, du *TOI-processing*... On peut toutefois s'affranchir en grande partie, des effets connus, par une correction adhoc, estimée sur un ensemble de simulations Monte-Carlo. C'est la démarche que nous avons adoptée afin de corriger l'effet du masque des données, de l'inpainting, de l'apodisation et du bruit inhomogène qui sont les effets dominants. et qui est détaillée dans la suite.

### 8.7.1 Jeux de simulations

Nous avons vu dans la section 7.5.3 qu'il était possible de corriger certains effets systématiques à l'aide de simulations Monte-Carlo. Nous utilisons l'hypothèse  $H_0$  qui implique de déterminer les corrections sur un jeu de données ne contenant pas de signal, c'est-à-dire des cartes non lentillées. Nous avons montré qu'il était nécessaire de générer un grand nombre de simulations afin d'obtenir une bonne estimation de la correction. Pour notre étude, nous avons choisi de nous concentrer sur la correction des effets suivants :

- un bruit inhomogène
- un lobe instrumental gaussien et circulaire
- le masquage des données
- l'inpainting
- l'apodisation
- la projection

qui sont les effets dominants pour la mesure du spectre de la matière.

La simulation est réalisée en plusieurs étapes. Dans un premier temps nous utilisons le spectre de bruit de chaque fréquence déterminé à l'aide des cartes *Jack-Knives*. Nous mesurons le spectre de la différence entre les cartes produites avec la première moitié des cercles d'un *ring* et celles produites avec la seconde moitié. Le signal étant supposé constant, seul le bruit est présent dans cette différence, ce qui permet d'estimer le spectre de puissance du bruit. Nous utilisons pour cela le code anafast.

Dans un second temps, on simule un ensemble de réalisations de cartes de bruit, possédant le spectre déterminé précédemment. On utilise ensuite la carte de *hit* associée à chaque fréquence pour produire une carte de bruit inhomogène comme décrit dans la section 6.4.2.

La troisième étape consiste à réaliser un ensemble de tirages de cartes du CMB non lentillées. Nous générons en plus un ensemble de réalisations de cartes de  $\phi$  afin de simuler les cartes lentillées associées. Ces dernières ne sont pas directement utilisées pour la correction des effets systématiques car nous nous basons sur l'hypothèse  $H_0$ , cependant elles permettent de valider la correction obtenue (voir section 8.7.3. Les spectres de puissance des cartes du CMB et de  $\phi$  utilisés sont ceux obtenus avec CAMB à partir de la meilleure estimation des paramètres cosmologiques par l'équipe WMAP, intégrant 7 ans de données. Les cartes sont générées avec le code LENSPIX comme expliqué dans la section 6.3.1. Elles sont convoluées avec un lobe gaussien circulaire dont la taille est tirée du tableau 8.4 et dépend de la fréquence.

Finalement les cartes de signal et de bruit sont additionnées pour obtenir le jeu de cartes final. Nous répétons cette opération pour chacune des six fréquences. Pour la correction de la carte GMCA, nous avons intégré en plus, dans la simulation, une étape de combinaison des cartes en fréquence tirée de la méthode de séparation de composantes. C'est pour cela que nous avons généré des simulations aux six fréquences de l'instruments HFI. Il suffit alors d'utiliser la matrice de mélange déterminée sur les données et de l'appliquer aux cartes simulées pour produire un jeu de simulations de cartes GMCA issu d'une combinaison linéaire réaliste des cartes en fréquence.

À ce stade, la simulation ne tient compte que du lobe de l'instrument et du bruit inhomogène. Afin d'obtenir une simulation reproduisant le mieux possible la procédure appliquée aux données, il faut également tenir compte de l'effet du masquage et appliquer aux simulation la même procédure que sur les données. L'ensemble des cartes précédemment simulées sont alors masquées puis inpaintées avec la même méthode que celle employée sur les données. On utilise donc les masques combinés associés à chaque fréquence et la technique d'inpainting local par patch avec les mêmes paramètres que dans la section 8.2.1.

Dans chaque cas nous avons simulé 100 cartes, ce qui permet d'obtenir des erreurs statistiques faibles et donne accès à une bonne estimation des effets systématiques. On obtient donc trois jeux de simulations associés aux cartes à 143GHz et 217GHz et la carte GMCA.



### 8.7.2 Détermination de la correction

Pour estimer la correction Monte-Carlo, il faut mesurer le spectre de puissance de la déflexion sur les simulations non lentillées en suivant la même procédure que sur les cartes de PLANCK :

- masquage des données
- inpainting local par patch
- découpe des patches
- déconvolution
- estimation du spectre de puissance des anisotropies de température
- estimation des cartes du potentiel de lentille
- estimation du spectre de la déflexion
- correction des bruits de l'estimateur

Pour chacun des trois jeux de simulations on obtient alors une série de 100 estimations du spectre de la déflexion. La correction est finalement obtenue en moyennant ces estimations. Étant donné le nombre de mesures, cela permet de diminuer les fluctuations statistiques par un facteur 10.

Les cartes analysées ne sont pas lentillées et si la procédure était parfaite le spectre moyen devrait être nul. L'intérêt des simulations Monte-Carlo est justement de mesurer le signal introduit par la méthode afin de le soustraire sur les données réelles. Le signal restant correspond donc aux effets systématiques introduits par les effets instrumentaux et la méthode d'analyse. Cette correction prend en compte l'effet du bruit inhomogène, d'un lobe gaussien circulaire, de la projection, de l'apodisation, du masquage et de l'inpainting.

### 8.7.3 Validation de la correction Monte-Carlo

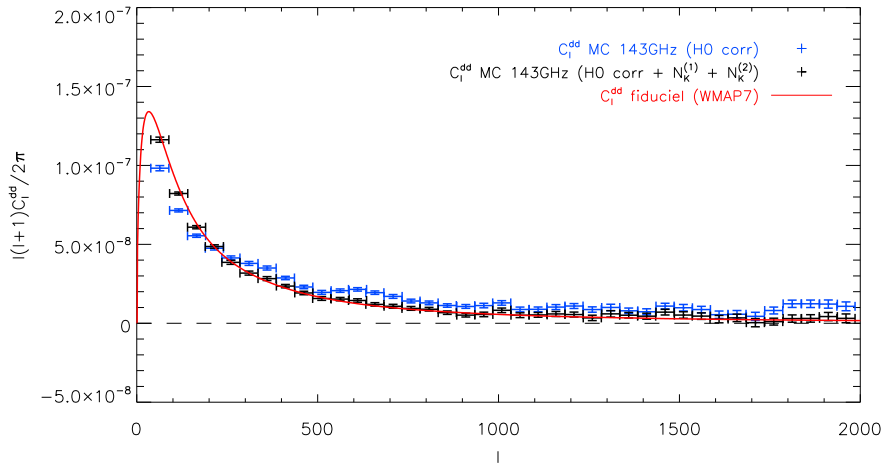
Afin de valider la correction Monte-Carlo nous avons mesuré le spectre de puissance de la déflexion sur les simulations lentillées. Nous appliquons pour cela la même procédure que sur les cartes de PLANCK. Les cartes lentillées contiennent un bruit inhomogène, issu du spectre *Jack-Knife* et de la carte de *hit* associés à chaque fréquence, elles sont masquées avec le même masque que les données puis inpaintées avec la même méthode et les mêmes paramètres que les cartes de PLANCK correspondantes. Dans ce test elles correspondent donc à des simulations des données mesurées par HFI.

Finalement, les spectres estimés sont corrigés à l'aide de la correction Monte-Carlo déterminée dans la section 8.7.2. Afin d'augmenter la statistique et d'obtenir une mesure précise du spectre de puissance de la déflexion, l'opération est répétée sur les 100 cartes lentillées, pour chacun des trois jeux de simulation. Encore une fois, on moyenne les 100 mesures afin de diminuer les fluctuations statistiques par un facteur 10. Si la correction est parfaite l'ensemble des points devrait être compatible avec le spectre fiduciel utilisé dans les simulations, aux incertitudes statistiques près, mesurées par les barres d'erreurs à  $1\sigma$ , estimées par la dispersion sur les 100 réalisations. On peut observer sur la figure 8.15 qu'on obtient un très bon accord. Il reste cependant un faible effet systématique qui reste cependant bien inférieur aux erreurs statistiques attendues pour la mesure sur une seule carte.

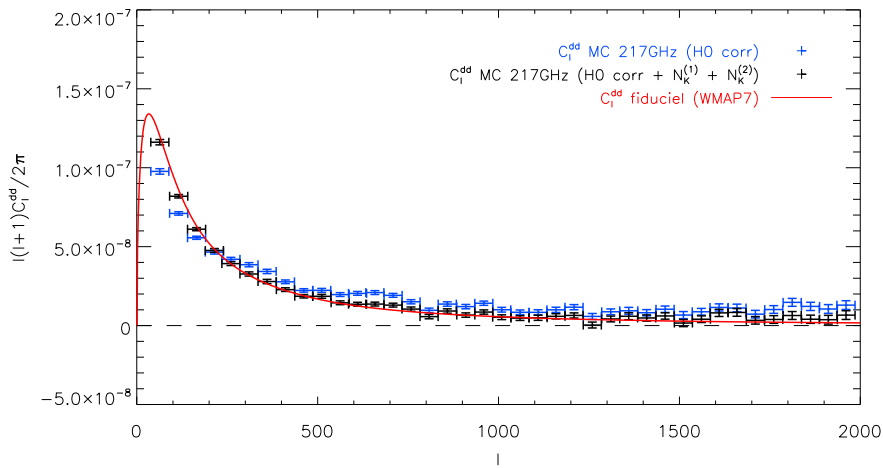
### 8.7.4 Correction du spectre de puissance de la déflexion sur les données

Nous avons montré dans la section précédente que les corrections Monte-Carlo permettaient de bien corriger des principaux effets systématiques introduits par les effets instrumentaux et la méthode d'analyse. Nous avons donc appliqué les corrections déterminées dans la section 8.7.2 sur les spectres bruts mesurés dans la section 8.6 ce qui nous a permis de réaliser la première mesure du spectre de puissance de la déflexion.

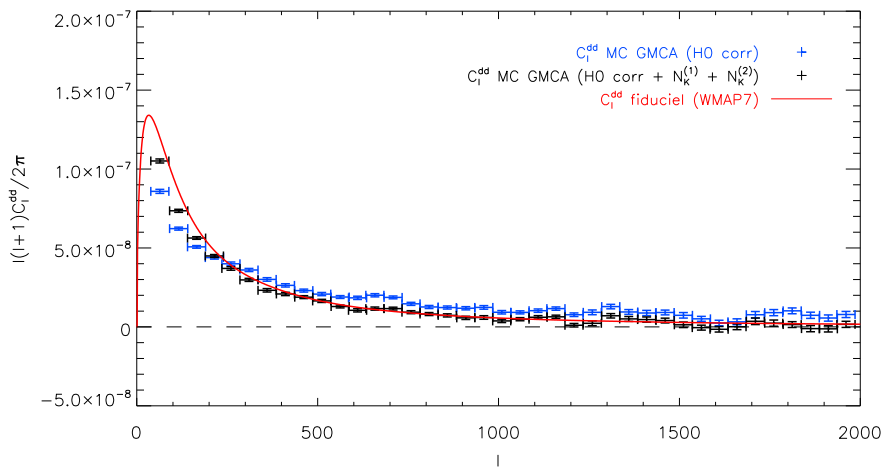
On peut voir sur la figure 8.16 les spectres obtenus pour les cartes combinées à 217GHz et 143GHz. Les observations à 143GHz et à 217GHz sont effectuées avec des jeux de détecteurs différents. Ce sont donc des mesures indépendantes. Afin de diminuer les erreurs statistiques, on peut donc combiner ces deux observations. Ce résultat est présenté sur la figure 8.17. Un test d'hypothèse à partir des 10 points de mesure inférieurs à 500, donne une incompatibilité avec l'absence de signal de déflexion à  $17\sigma$ . La comparaison au spectre calculé dans un modèle  $\lambda$ CDM à partir des paramètres cosmologiques déterminés par WMAP intégrant 7 ans d'observation (WMAP7), donne un  $\chi^2$  par degré de liberté de 1.26.



(a) Spectre de la déflexion pour le jeu de cartes Monté Calro à 143GHz

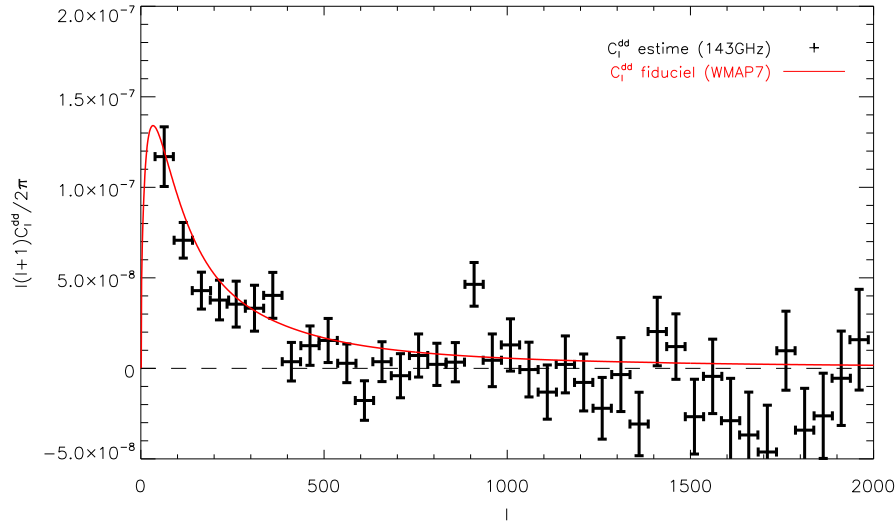


(b) Spectre de la déflexion pour le jeu de cartes Monté Calro à 217GHz

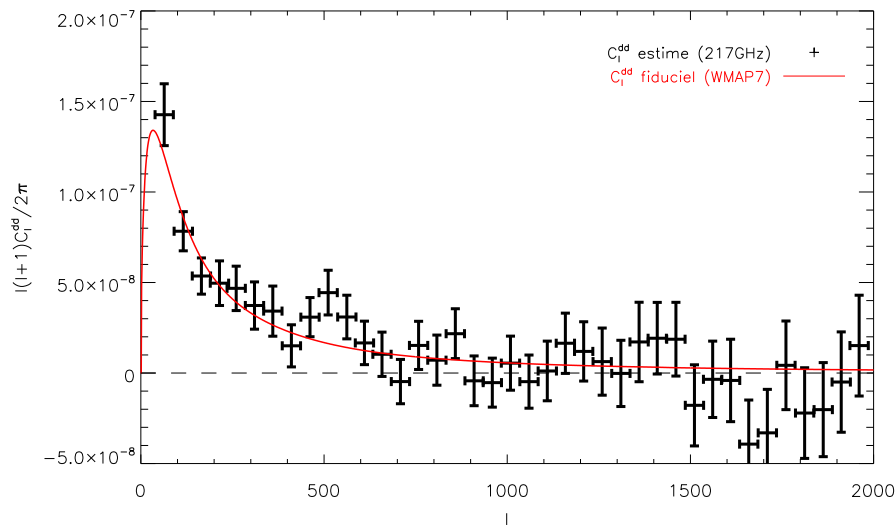


(c) Spectre de la déflexion pour le jeu de cartes Monté Calro GMCA

FIGURE 8.15 – **Validation de la correction du spectre de la déflexion par Monte-Carlo.** Sur les trois graphiques ci-dessus, on a représenté en bleu, le spectre estimé sur les 100 cartes lentillées du Monte-Carlo, corrigé uniquement des effets systématiques, comparé au spectre fiduciel tracé en traits plans rouge. Le spectre final, intégrant en plus, la correction des biais de l'estimateur  $N_K^{(1)}$  et  $N_K^{(2)}$ , est visible en noir. Dans les trois cas, les spectres sont obtenus par la moyenne des mesures effectuées sur les 100 cartes lentillées issues du Monte-Carlo et corrigées par la correction estimée sur les cartes non lentillées. L'accord avec le spectre attendu est très bon. Les erreurs systématiques restantes sont négligeables devant les erreurs statistiques pour une réalisation. 175



(a) Spectre de puissance de la déflexion de la carte des observations combinées à 143GHz



(b) Spectre de puissance de la déflexion de la carte des observations combinées à 217GHz

FIGURE 8.16 – **Spectres de puissance de la déflexion des cartes combinées, corrigés par Monte-Carlo.** L'estimation brute du spectre de puissance de la déflexion, déterminée dans la section 8.6.2, est corrigée des effets systématiques à l'aide d'une correction estimée sur un jeu de 100 réalisations Monte-Carlo. Les spectres des cartes à 143GHz et 217GHz sont visibles respectivement sur les figures (a) et (b) (en noir). Étant donné le niveau de précision de notre simulation, on peut considérer qu'on a un bon accord global avec le spectre fiduciel (en rouge) calculé à partir des paramètres cosmologiques déterminés par WMAP avec sept années d'observation. Les données sont tout de même compatibles avec l'absence de signal au-delà de  $l = 500$ .

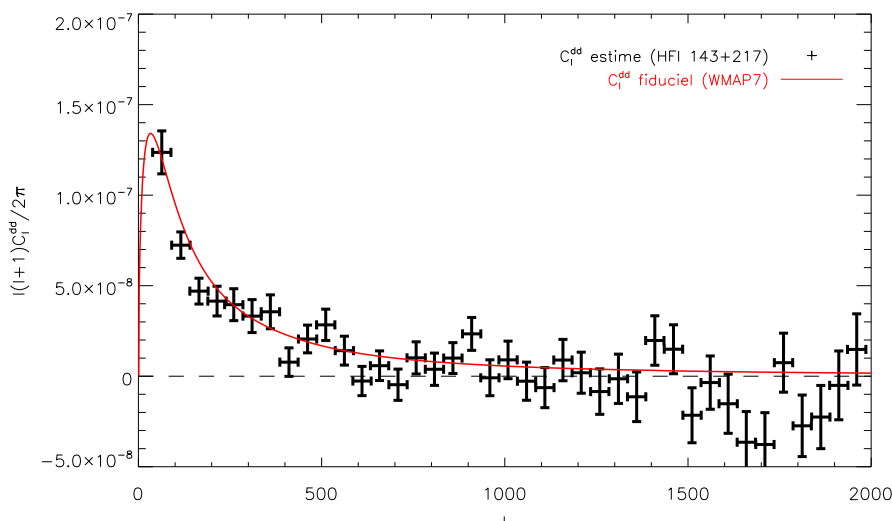


FIGURE 8.17 – **Spectre de puissance de la déflexion combiné.** Les mesures à 143GHz et 217GHz sont indépendantes. Elles peuvent donc être combinées pour diminuer les erreurs statistiques. Pour les multipôles inférieurs à  $\ell = 500$ , les résultats (en noir) sont toujours compatibles avec le spectre de puissance attendu à partir des paramètres cosmologiques déterminés par WMAP avec 7 ans de données (en rouge).

De façon générale, on constate qu'on a un bon accord global avec les prédictions de WMAP7, étant donnée la taille des barres d'erreur à  $1\sigma$ . On observe toutefois, pour certains intervalles, des déviations supérieures aux déviations attendues pour un simple effet statistique. Ces déviations sont principalement présentes dans la carte à 143GHz. Il faut garder à l'esprit que ces analyses ont été effectuées sur les cartes en fréquence, sans séparation de composantes. Malgré les différents tests utilisés pour éliminer les zones les plus problématiques, il est possible que certaines zones restent contaminées. Cette hypothèse est toutefois à nuancer par le fait que les effets systématiques sont plus importants à 143GHz. Or, d'après la figure 2.9, on s'attend à une contamination plus forte de la poussière galactique notamment à 217GHz, comme on peut le voir également sur les cartes de PLANCK de la figure 8.3. On peut également penser que la contamination à l'origine de ces déviations par rapport au spectre attendu par WMAP7, est présente uniquement à 143GHz et pas à 217GHz. C'est le cas notamment de l'effet SZ dont la signature est presque nulle à 217GHz. On notera toutefois que la taille du lobe est plus faible à 217GHz ce qui devrait permettre d'obtenir une meilleure estimation.

Pour tester l'influence de la contamination, il est intéressant de comparer les mesures par fréquence à la mesure effectuée sur la carte GMCA. Les résultats sont présentés sur la figure 8.18. Sur les 10 premiers points de mesure, un test d'hypothèse donne une incompatibilité avec l'absence de signal à  $14\sigma$  dans ce cas, tandis que la comparaison au donnée de WMAP7 conduit à  $\chi^2$  par degrés de liberté de 1.76. Ce qui permet de dire qu'on a un bon accord global. On observe cependant encore une fois, la présence de déviations importantes par rapport à la prédiction obtenue à partir des paramètres cosmologiques déterminés par WMAP7. Les points affectés ne sont pas les mêmes que pour les cartes en fréquence, ce qui ne permet de conclure. L'impact de la séparation des composantes avec GMCA a été étudié dans Perotto et al. (2010). Dans le cadre du modèle d'émission utilisé, il a été montré que la procédure était bien adaptée à l'analyse de l'effet de lentille gravitationnelle. Cependant, c'est la première fois que des observations sur tout le ciel sont réalisées aux plus hautes fréquences de PLANCK. Il est possible que certaines composantes soient mal modélisées et produisent une contamination dans les données réelles qui n'a pas pu être détectée dans l'étude de Perotto et al. (2010). Par ailleurs, on a également vu dans la section 8.7.3 que le Monte-Carlo était capable de corriger des effets systématiques introduit par une combinaisons de cartes de CMB pur, avec bruit et lobe de l'instrument, comparable à celle de la séparation des composantes sur les données réelles. Cependant, l'impact de la combinaison d'effets instrumentaux différents du modèle idéalisé pris en compte doit aussi être étudié avant de conclure.

Les résultats obtenus sont toutefois encourageants et ont permis la première mise en évidence

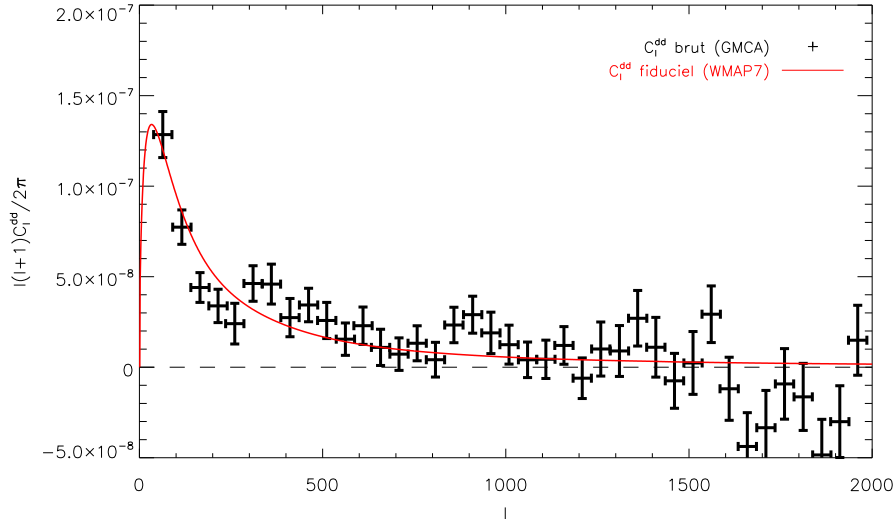


FIGURE 8.18 – **Spectre de puissance de la déflexion de la carte GMCA, corrigé par Monte-Carlo.** La correction est effectuée sur un jeu de 100 simulations non lentillées. Des effets systématiques sont visibles. Leur origine reste indéterminée et leur compréhension nécessite d’améliorer la simulation.

de l’effet de lentille gravitationnelle. Sur les simulations Monte-Carlo, on a vu qu’on était capable de corriger avec une bonne précision les principaux effets systématiques. L’application aux données conduit à un résultat satisfaisant étant donné le niveau des effets systématiques avant la correction. Cependant, à l’heure actuelle, aucune simulation ne reproduit les effets systématiques observés. Il faut garder à l’esprit que la procédure ne tient pas compte de tous les effets instrumentaux. L’impact de l’ensemble de la chaîne sur l’effet de lentille gravitationnelle n’a encore jamais été testé et des effets importants restent certainement à identifier et prendre en compte, notamment l’impact des rayons cosmiques, le *TOI-processing*, la génération des cartes, des lobes elliptiques, de la séparation de composantes en présence de tous ces effets... Cela demande de produire des nouvelles simulations prenant en compte l’ensemble des étapes de la mission. C’est un travail long qu’il est impératif de réaliser avant de pouvoir conclure.

L’outil que nous avons développé ouvre également la voie à de nombreuses possibilités permettant de distinguer les contributions des effets instrumentaux, de celles produites par une éventuelle contamination. L’étude locale est capable notamment d’analyser des zones particulières du ciel individuellement. C’est une approche originale et très pertinente notamment pour la compréhension des contaminations.

Pour finir, étant donné les bons résultats des validations sur Monte-Carlo et au vu des mesures sur les données, on peut être relativement confiant sur le niveau global de la détection. Il est alors difficile de ne pas être tenté d’effectuer la comparaison aux prédictions faite à partir des paramètres cosmologiques estimés par WMAP. La courbe rouge des figures 8.17 et 8.18 donne la prédiction attendue par WMAP7. Dans une première approximation on peut affirmer que les résultats donnent confiance dans les modèles et que de fortes déviations par rapport aux prédictions sont exclues. Il serait cependant prématuré de pousser plus loin la comparaison ou les conclusions qu’on peut en tirer. Il faut finalement garder à l’esprit la présence de fortes incertitudes systématiques qui n’ont pas encore été évaluées. C’est le principal défi qui devra être relevé par la collaboration afin d’obtenir un résultat scientifiquement exploitable.

# Conclusions

Moins de 18 mois après le début des observations, PLANCK vient de livrer ses premiers résultats scientifiques à l'ensemble de la communauté. Les analyses effectuées jusqu'à aujourd'hui montrent que la qualité des données sera à la hauteur des objectifs ambitieux de la mission. La compréhension fine des effets systématiques qui les affectent constitue toutefois un des principaux défis restant à relever. Ma contribution au sein de la collaboration, s'est inscrite dans l'effort collectif qui a été mené jour après jour depuis le lancement, dans le but de s'assurer de la qualité des données délivrées et du bon fonctionnement de l'ensemble de la chaîne de mesure, depuis les détecteurs jusqu'au centre de traitement des données. Mon travail s'est également porté sur l'analyse des données dans un but scientifique reposant sur le développement d'une méthode originale d'analyse de l'effet de lentille gravitationnelle.

Dans le cadre des efforts de surveillance de la qualité des données de l'instrument HFI, j'ai eu en charge le suivi régulier du débit de données et du bon fonctionnement de la compression ainsi que l'optimisation des paramètres de l'algorithme de compression pour l'instrument HFI. Après une période de développement des outils nécessaires, j'ai effectué le suivi régulier de façon parallèle aux analyses de l'effet de lentille gravitationnelle, à l'exception des trois mois suivant le début des observations qui ont été consacrés uniquement à la compréhension des problèmes de compression rencontrés et à l'optimisation des paramètres de compression dans les conditions réelles de fonctionnement en vol.

Mon suivi de la compression a permis de mettre en évidence, dès le premier jour d'observation, la présence d'un faible nombre d'erreurs de compression, ce qui n'avait pas été prédit par les simulations simples effectuées avant le lancement. Le travail systématique de recherche et de classement que j'ai effectué dans les premiers mois, a permis de comprendre leur l'origine. Elles sont provoquées par l'impact de certains rayons cosmiques qui déposent, sur les bolomètres, une quantité d'énergie beaucoup plus importante que celle des populations les plus courantes, conduisant à une modification trop brutale du signal. Dans ces conditions, on peut atteindre une limite de sécurité, introduite pour ne pas saturer le débit de données, qui conduit à des erreurs de compression. Ces erreurs de compression correspondent toutefois à un fonctionnement normal et attendu de l'algorithme de compression, pour des situations extrêmes comme c'est le cas avec certains rayons cosmiques ou plus généralement avec tout phénomène capable de produire une variation du signal de plusieurs ordres de grandeur sur un temps de l'ordre de la seconde. On observe en moyenne un maximum de 10 impacts de ce type par jour et par voie, conduisant, dans le pire des cas, à la perte de 190 échantillons, ce qui représente une fraction de données perdues de quelques  $10^{-5}$ . Parallèlement, la recherche d'échantillons manquants issus d'une éventuelle saturation de la bande passante, m'a conduit à identifier plusieurs erreurs inattendues liées à un problème de traitement au niveau L1, provoquant la perte de certains paquets télémétriques. Suite à mon rapport, entre autre, un correctif a été introduit. Mon suivi régulier du débit des données, a aussi permis de mettre en évidence une augmentation anormale au cours du mois de septembre 2009. Son origine a vite été attribuée à une diminution de l'efficacité de la compression sur les voix à haute fréquence lors du passage sur le centre galactique. Dans cette zone du ciel, les variations du signal sont très importantes pour les canaux à haute fréquence, sur de courtes échelles de temps ce qui ne correspond plus à la condition optimale de fonctionnement de l'algorithme de compression. Pour satisfaire aux contraintes sur le débit de données, afin de ne pas saturer la bande passante et l'espace de stockage alloués à HFI dans le satellite, nous avons procédé à un changement des paramètres de compression sur l'ensemble des voies à 857Ghz, ce qui a permis de ne pas atteindre un débit trop élevé au cours des passages suivants. En outre, les fortes variations sur de courtes échelles de temps, rencontrées lors

du passage dans la région du centre galactique ont produit des erreurs de compression sur l'ensemble des voies à 857GHz, au cours de la première couverture du ciel. Le changement des paramètres de compression de ces voies a par la suite, aussi résolu ce problème. Par ailleurs, des erreurs de nature similaire se sont produites lors de l'observation de la planète Jupiter sur certaines voies à 857GHz, toujours provoquées par la variation brutale d'intensité rencontrée. Enfin, j'ai également contribué à l'optimisation des paramètres de l'algorithme de compression grâce à un outil que j'ai développé, qui permet de déterminer avec précision le rapport entre le pas de quantification et l'écart type du signal qui fixe totalement la qualité et l'efficacité de la compression. Ce travail a été utilisé pour déterminer les nouvelles valeurs des paramètres pour les voies à haute fréquence et a permis de s'assurer que les réglages effectués au début du vol, étaient optimaux sur les autres voies. Depuis le changement des paramètres de compression à haute fréquence, aucun problème important n'a été détecté. Même si un suivi régulier reste nécessaire, le système doit pouvoir fonctionner sans problème jusqu'à la fin de la prise de données, prévue en janvier 2012.

L'objectif principal de ma thèse est l'étude de l'effet de lentille gravitationnelle dans les données de l'instrument HFI de PLANCK. Grâce à la résolution et au niveau de bruit des instruments de PLANCK et au développement d'un estimateur statistique, la mesure du spectre de la déflexion des photons du CMB est accessible pour la première fois. Sa mesure devrait apporter de nouvelles contraintes sur les modèles d'Univers et devrait notamment permettre d'améliorer la limite supérieure sur la somme des masses des neutrinos. L'application concrète de cet estimateur sur des données réelles, n'avait jamais été conduite auparavant. Elle a fait apparaître des problèmes qui n'avaient pas été pris en compte lors du développement théorique de l'estimateur, qui se plaçait dans le cadre d'observations idéales, avec un bruit blanc gaussien parfaitement homogène, un lobe instrumental gaussien symétrique, une couverture idéale et complète du ciel, sans contamination. La variante de l'estimateur pour des cartes locales existait seulement avec des conditions aux bords périodiques. Les problèmes liés à l'estimation sur une carte possédant des régions masquées, ou sur une observation partielle du ciel, l'influence de la contamination par d'autres composantes astrophysiques ou les effets instrumentaux complexes, n'avaient pas été pris en compte. La complexité du problème m'a conduit à orienter mon analyse vers une méthode originale basée sur un traitement des données par petites régions du ciel.

Un des intérêts principal de cette méthode, est la liberté dans le choix des régions du ciel à étudier qui donne la possibilité de s'affranchir des zones les plus contaminées et permet de réaliser des analyses locales. Cette approche est particulièrement adaptée à l'étude des effets systématiques qui varient spatialement, à la mise en évidence de contaminations localisées sur les cartes en fréquences ou de résidus d'autres composantes sur une carte du CMB, issue de la séparation de composantes. Elle facilite, par ailleurs, la comparaison de différentes zones du ciel, permettant de vérifier l'isotropie de certaines propriétés ou de faire des études « Jack-Knife » spatiales sur une carte (par exemple la comparaison de la partie Nord et Sud, ou des différents quadrants...). Enfin, l'étude permet de tirer partie de la plus grande homogénéité du bruit sur des cartes locales.

Je me suis d'abord concentré sur le développement des outils nécessaires à cette analyse, notamment la découpe du ciel en morceaux de petite taille qui est au cœur de la méthode employée. J'ai développé l'ensemble des programmes en C++ sous une forme modulaire, facilement réutilisable. Ils peuvent ainsi être adaptés à d'autres problématiques. Une inclusion dans l'architecture PLANCK sous forme de pipeline est envisageable grâce à une séparation stricte des entrées/sortie du reste des programmes et à des conventions identiques sur les paramètres entre les différents exécutables. Le développement a été guidé par deux objectifs : l'application à la mesure du spectre de la déflexion et la mise à disposition d'un outil général utilisable dans d'autres contextes. La découpe de cartes en morceaux est une approche rencontrée dans de nombreuses analyses en particulier la détection et la caractérisation des sources compactes. Le travail sur des morceaux se fait presque exclusivement sur des grilles régulières et le passage de la pixelisation de la sphère à celle du plan nécessite d'évaluer le signal sur cette grille. De nombreuses analyses se basent sur une simple interpolation pour estimer le signal sur une grille régulière. La précision requise pour la mesure de l'effet de lentille gravitationnelle ne permet pas d'employer ces méthodes simples et m'a conduit vers une méthode dont j'ai prouvé l'efficacité, basée sur un ajustement des coefficients de Fourier.

Dans un second temps, la compréhension et la caractérisation du comportement de l'estimateur et le développement d'une méthode permettant de corriger ses limites afin d'obtenir une procédure applicable aux données de PLANCK, a constitué une grande part de mon travail. Les résidus galactiques peuvent être évités facilement grâce à la technique employée qui permet de sélectionner

les zones du ciel à analyser et de s'affranchir ainsi des régions contaminées par les avant-plans. La contamination par les sources compactes qui sont distribuées sur tout le ciel, ne peut pas être évitée sur les zones du ciel de plusieurs degrés de côté nécessaires à l'étude de l'effet de lentille gravitationnelle, ce qui m'a conduit à adapter une méthode « d'inpainting », afin de réaliser un traitement local. Le problème des conditions aux bords périodiques a été résolu par l'apodisation de la carte. Cette solution produit, en contre partie, un effet systématique qui a été traité à l'aide de simulations Monte-Carlo que j'ai effectuées. Ces simulations intègrent également l'effet du bruit inhomogène et de l'inpainting. En plus des effets systématiques précédents, les simulations employées pour corriger les résultats de la mesure sur la carte issue de la séparation des composantes incluent aussi les effets de la combinaison des cartes en fréquence de la méthode.

Une analyse préliminaire sur les données disponibles après moins de cinq mois d'observation a conduit à des résultats préliminaires, incompatibles avec l'absence de signal. Ils ont démontré à l'époque, la qualité déjà fort satisfaisante des données et la validité de la méthode de mesure. Par la suite, cette méthode a été appliquée au jeu de données à la base des premiers résultats scientifiques de PLANCK. Dans un premier temps, j'ai effectué l'analyse des cartes en fréquence à 217GHz puis à 143GHz démontrant pleinement la faisabilité de l'analyse et la présence d'un signal de déflexion. L'analyse finale que j'ai réalisée, s'est appuyée sur la carte issue de la séparation des composantes avec l'algorithme GMCA. Dans chaque cas, la correction des effets systématiques s'est appuyée sur un ensemble de 100 simulations Monte-Carlo réalisées au centre de calcul de l'IN2P3 à Lyon et m'a demandé plusieurs semaines de traitement, nécessitant plusieurs dizaines de milliers d'heure de calcul. La manipulation de ces données a nécessité au total plusieurs To de données. Les mesures obtenues, à partir des résultats combinés des cartes à 143GHz et 217GHz et de la carte issue de la séparation de composantes montrent une incompatibilité avec l'absence de signal de déflexion à  $17\sigma$  et  $14\sigma$  respectivement. Par ailleurs, ces observations sont en accord avec le spectre de la déflexion théorique attendu pour un modèle  $\lambda$ CDM avec les paramètres cosmologiques déterminés par WMAP et constitue la première mesure de ce spectre réalisée à partir du CMB. En se limitant aux points de mesure dans la gamme des  $\ell$  inférieurs à 500, on trouve un  $\chi^2$  par degré de liberté de 1.26 et 1.76 respectivement pour les spectres obtenus à partir de la combinaison des observations à 143GHz et 217GHz d'une part et avec la carte issue de la séparation de composantes avec GMCA d'autre part. Il est intéressant de mentionner tout de même que les expériences ACT et SPT qui étudient également le CMB, depuis le sol, ont également obtenu très récemment, des résultats préliminaires sur le sujet, qui montrent un spectre de la déflexion compatible avec celui attendu pour un modèle  $\lambda$ CDM, en accord avec les résultats que j'ai obtenus.

Même si elle permet déjà d'obtenir des résultats très prometteurs, plusieurs orientations sont imaginables pour améliorer ma procédure d'analyse. Tout d'abord, il est nécessaire d'estimer l'impact des effets systématiques que nous n'avons pas pris en compte, notamment l'effet du traitement des données ordonnées en temps, du *destriping* ou encore de la génération des cartes. Il est également important de tester l'effet des résidus de séparation de composantes en se basant sur des simulations complètes intégrant les autres émissions. Par ailleurs, le code d'estimation des coefficients de Fourier sur une grille régulière qui est au cœur de la méthode employée pourrait également être amélioré. Dans un premier temps on pourrait tenir compte de l'erreur associée à chaque pixel de la carte du ciel afin de prendre en compte l'effet du bruit inhomogène directement dans l'estimation des coefficients de Fourier. Dans un second temps on pourrait intégrer un préconditionnement afin de faciliter la convergence, ce qui devrait permettre de s'affranchir totalement de l'inpainting des sources ponctuelles, en ne prenant plus du tout en compte les pixels à masquer. Enfin, cette méthode générale peut être utile à toutes les analyses à petite échelle, notamment l'étude des sources ponctuelles, des amas de galaxies, des sources compactes en général... En effet, la découpe du ciel en morceaux est un problème récurrent pour ce type d'analyse. On peut constater que dans la plupart des cas, les méthodes employées pour extraire des cartes du ciel, se basent sur une découpe du ciel utilisant une estimation de la valeur des pixels basique, basée sur une interpolation bilinéaire ou même par plus proche voisin. Il serait intéressant de développer des collaborations avec les équipes travaillant sur ce sujet afin de tester si la méthode que j'ai développé reposant sur un ajustement des coefficients de Fourier sur une grille régulière peut apporter des améliorations.

Finalement, la détection dans les données de PLANCK du spectre de la déflexion, ouvre la voie à des développements scientifiques très prometteurs. L'effet de lentille gravitationnelle représente une observable de plus pour la détermination des paramètres cosmologiques qui permet d'améliorer la



précision de leur mesure mais également de lever certaines dégénérescences. Un des résultats les plus attendus reste sans doute la contrainte sur la somme des masses des neutrinos qui devrait permettre de mettre une limite supérieure de  $0.3eV$ . Enfin, la carte du potentiel intégré peut être corrélée aux grandes structures, notamment les grands relevés de galaxies. Un projet de corrélation croisée avec le fond diffus infrarouge s'apprête également à être lancé. Dans ce contexte l'étude par morceaux est particulièrement adaptée car les mesures précédentes ne sont accessibles que sur des petites portions du ciel.

## Annexe A

# Champ aléatoire en cosmologie

Un champs aléatoire scalaire  $T$ , dans un espace de dimension  $N$  est un ensemble de variables aléatoires avec une fonction de distribution associée. Si l'on suppose que ce champ est *homogène*, la fonction de distribution est invariante par translation et le processus est stationnaire au second ordre :

- sa valeur moyenne (d'ensemble)  $m(\mathbf{x}) = \langle T(\mathbf{x}) \rangle$  en une position donnée est constante
  - sa fonction d'autocorrélation est donnée par  $\xi(\mathbf{x}_1, \mathbf{x}_2) = \langle (T(\mathbf{x}_1) - m(\mathbf{x}_1))(T(\mathbf{x}_2) - m(\mathbf{x}_2))^* \rangle$
- Le théorème de décomposition spectrale montre alors que ses composantes de Fourier

$$\tilde{T}(\mathbf{k}) = \int T(\mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}} d^N \mathbf{x} \quad (\text{A.1})$$

deviennent des variables aléatoires *décorrélées* :

$$\langle \tilde{T}(\mathbf{k}_1) \tilde{T}^*(k_2) \rangle = P(\mathbf{k}_1) (2\pi)^N \delta(\mathbf{k}_1 - \mathbf{k}_2) \quad (\text{A.2})$$

où  $P(\mathbf{k}_1)$  désigne le spectre de puissance du processus.

L'hypothèse d'*isotropie*, indique que la fonction d'autocorrélation ne doit pas dépendre de la direction :

$$\xi(\mathbf{x}, \mathbf{x} + \mathbf{r}) = \xi(\mathbf{x}, r) \quad (\text{A.3})$$

où  $r = |\mathbf{r}|$  ne dépend que du module. Ainsi le spectre de puissance ne dépend aussi que du module de  $k$  :

$$\langle \tilde{T}(\mathbf{k}_1) \tilde{T}^*(k_2) \rangle = P(k_1) (2\pi)^N \delta(\mathbf{k}_1 - \mathbf{k}_2) \quad (\text{A.4})$$

En particulier les fluctuations de températures juste avant le découplage matière-radiation sont décrites comme un champ aléatoire de dimension 3 avec les hypothèses du principe cosmologique.

Dans la pratique on n'a accès qu'à une seule réalisation d'Univers : on utilise alors le théorème d'*ergodicité* qui est démontré dans Lyth D. (2009) dans le cas où la « boîte » contenant les modes de Fourier est suffisamment grande devant les échelles de perturbations considérées. Il permet alors de remplacer la moyenne d'ensemble par une moyenne spatiale.



## Annexe B

# Approximation par séries de Fourier de fonctions irrégulièrement échantillonnées

### B.1 Présentation à une dimension

#### B.1.1 Position du problème

Soit une fonction  $f$  de la variable  $t$ , échantillonnée sur un ensemble de points  $\{t_i\}$ , arbitrairement répartis dans l'intervalle  $[0, T]$ . On note  $f_i = f(t_i)$ . Le problème que nous voulons résoudre consiste à estimer la fonction  $f$ , sur l'intervalle  $[0, T]$ , à partir de ses mesures  $f_i$ , en estimant ses coefficients de Fourier. La méthode consiste à supposer que la fonction est périodique sur l'intervalle  $[0, T]$ . On peut alors la décomposer en série de Fourier sous la forme :

$$f(t) = \sum_{k=-\infty}^{\infty} a_k e^{2i\pi k \frac{t}{T}}$$

Si on suppose par ailleurs, que le **spectre en fréquence** de la fonction à approcher est à **support borné**, on peut limiter la série de Fourier à une somme finie d'éléments :

$$f(t) = \sum_{k=-M}^M a_k e^{2i\pi k \frac{t}{T}}$$

La solution consiste alors à calculer les coefficients  $a_k$ , étant données les valeurs d'échantillonnage  $f_i$ .

#### B.1.2 Formalisme

Pour faciliter les formules, on travaille avec la variable réduite  $u = \frac{t}{T}$ . Si le nombre des points d'échantillonnage  $u_i$  est  $N$ , on peut alors écrire les  $N$  équations :

$$f_i = \sum_{k=-M}^M a_k e^{2i\pi k u_i}$$

On obtient donc un système linéaire de  $N$  équations à  $2M+1$  inconnues  $a_k$  que l'on peut écrire sous forme matricielle. L'élément générique de la matrice  $G$  de ce système est  $g_{kl} = e^{2i\pi k u_l}$ . La solution optimale, au sens de la norme hilbertienne, qui est celui des moindres carrés, est obtenue en résolvant le système linéaire  $(G^T G)X = G^T F$ , où  $F$  est le vecteur colonne des valeurs d'échantillonnage. Le point intéressant ici est que le terme général de la matrice de ce dernier système est :

$$T_{kl} = \sum_{j=1}^N e^{-2i\pi(k-l)u_j}$$

La matrice  $T$  est donc une matrice de Tœplitz, c'est-à-dire une matrice dont chaque élément  $c_{ij}$ , vérifie la relation  $c_{ij} = c_{i-j}$ . Le terme général du second membre s'écrit lui sous la forme :

$$b_k = \sum_{j=1}^N f_j e^{-2i\pi k u_j}$$

### B.1.3 Mise en œuvre

La matrice du système étant une matrice hermitienne définie positive, on peut donc résoudre le système, en se basant sur l'algorithme décrit dans Feichtinger et al. (1995), par la méthode des gradients conjugués. C'est un algorithme itératif qui se réduit dans notre cas, à une succession de produits matrice-vecteur.

Il est possible de transformer une matrice de Tœplitz en matrice circulante, c'est-à-dire une matrice dont chaque élément  $c_{ij}$  vérifie la relation  $c_{ij} = c_{i-j} = c_{i-j+n}$ . Cette astuce permet d'utiliser les propriétés des matrices circulantes, pour accélérer les calculs (voir Chan & Ng (1996)). En effet, le produit d'une matrice circulante par un vecteur, n'est rien d'autre qu'une convolution discrète, que l'on peut effectuer par transformée de Fourier rapide (voir section B.2.1).

Une matrice de Tœplitz,  $T$ , de dimension  $n$  par  $n$ , peut être transformée en matrice circulante  $\tilde{T}$  de dimension  $((2n-1) \text{ par } (2n-1))$ , en complétant les éléments avec des zéros et les conjugués des éléments de  $T$ . Le vecteur à multiplier  $a$  doit être complété par  $(n-1)$  zéros pour obtenir le vecteur  $\tilde{a}$ . On peut alors effectuer le produit par transformée de Fourier. Le résultat apparaît comme les  $n$  premières composantes du vecteur  $\tilde{T}\tilde{a}$ . Bien que la matrice à traiter soit plus grande, le produit est très rapide par FFT.

Si l'échantillonnage de la fonction est mauvais, par exemple s'il est proche de la densité critique ou avec des écarts importants entre deux échantillons successifs, le gradient conjugué peut éprouver des difficultés à converger. On peut alors préconditionner la matrice de Tœplitz par une matrice circulante appropriée. On peut combiner ce préconditionnement avec l'usage de *poids* permettant de tenir compte de l'erreur associée à chaque mesure. Si on a les points d'échantillonnage :

$$0 \leq u_1 \leq u_2 \leq \dots \leq u_r \leq 1$$

on définit les poids suivants :

$$w_j = \frac{u_{j+1} - u_{j-1}}{2}$$

$$w_{N+1} = u_1 + 1$$

$$w_0 = u_N - 1$$

La matrice de Tœplitz devient :

$$T_{kl} = \sum_{j=1}^N w_j e^{-2i\pi(k-l)u_j}$$

et le second membre :

$$b_k = \sum_{j=1}^N f_j w_j e^{-2i\pi k u_j}$$

On peut évidemment utiliser n'importe quel jeu de poids tenant compte du degré de confiance dans les différentes mesures. Il faut noter que la meilleure qualité de l'approximation n'est garantie que si aucun intervalle entre deux échantillons consécutifs ne dépasse la période de Nyquist ( $1/2M$ ).

## B.2 Généralisation à deux dimensions

### B.2.1 Outils mathématiques

Nous commençons par présenter quelques propriétés et outils mathématiques afin de poser les notations. Ces résultats permettent de mieux comprendre le développement de la méthode et facilitent les calculs.





### B.2.4 Mise en œuvre

Comme dans le cas à une dimension on va résoudre le système par la méthode des gradients conjugués. Sachant que la matrice du système est toeplitz-bloc-toeplitz, on va étendre le procédé utilisé dans le cas à une dimension. Pour cela, on peut tout d'abord ré-écrire le produit de la matrice du système par un vecteur quelconque, donné sous la forme  $vec(X)$  :

$$\left( \sum_{j=0}^{N-1} T_j^v \otimes T_j^u \right) vec(X) = \sum_{j=0}^{N-1} vec(T_j^u X (T_j^v)^T) \quad (\text{B.9})$$

On étend, comme pour le cas à une dimension, les matrices de Toeplitz  $T_j^u$  et  $T_j^v$  à des matrices circulantes  $C_j^u$  et  $C_j^v$  en ajoutant des zéros et les termes conjugués adéquats. Dans la pratique on s'arrangera (avec les zéros) pour obtenir des dimensions égales à des puissances de 2, pour accélérer les transformées de Fourier. Il faut noter que les deux matrices circulantes n'auront pas nécessairement les mêmes dimensions. Finalement il faut compléter la matrice  $X$  avec des zéros, de manière à obtenir une compatibilité des dimensions, ce qui donne la matrice  $\tilde{X}$  :

$$\tilde{X} = \begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix}$$

Le produit  $T_j^u X (T_j^v)^T$  est la sous-matrice de dimensions  $(2M_x + 1)(2M_y + 1)$  en haut et à gauche de la matrice :

$$C_j^u \tilde{X} (C_j^v)^T$$

Mais, il est facile de vérifier que cette dernière matrice est le produit de convolution de  $\tilde{X}$  par la matrice  $W_j = \alpha_j^u \cdot (\alpha_j^v)^T$  où  $\alpha_j^u$  et  $\alpha_j^v$  sont les vecteurs premières colonnes des circulantes  $C_j^u$  et  $C_j^v$ .

On effectue donc le calcul par FFT :

$$DFT2(\tilde{X}W_j) = DFT2(\tilde{X}) \cdot * DFT2(W_j)$$

Mais, ce qu'il faut effectuer, c'est le produit (B.9), en sommant sur  $j$ , soit :

$$DFT2(\tilde{X} \sum_{j=0}^{N-1} W_j) = DFT2(\tilde{X}) \cdot * DFT2(\sum_{j=0}^{N-1} W_j)$$

Et on prend la transformée de Fourier inverse pour obtenir le produit désiré. On retiendra que la transformée  $DFT2(\sum_{j=0}^{N-1} W_j)$  se calcule une fois pour toutes.

Finalement si on introduit la possibilité d'affecter des degrés de confiance sur les données, sous la forme d'un vecteur de poids de composantes  $w_j, j = 0, N - 1$ , le système à résoudre s'écrit alors :

$$\left( \sum_{j=0}^{N-1} w_j T_j^v \otimes T_j^u \right) vec(A) = \sum_{j=0}^{N-1} vec(\bar{C}_j^u (\bar{C}_j^v)^T) w_j f_j \quad (\text{B.10})$$

La mise en oeuvre n'en est pas notablement modifiée.





## Annexe C

# Paramètres de compression

Voie	Nom	$q$	$q/\sigma$
00	100 1A	38	0.4
01	100 1B	38	0.4
02	143 1A	40	0.4
03	143 1B	44	0.4
04	217 1	43	0.4
05	353 1	55	0.4
10	143 5	45	0.4
11	217 5A	40	0.4
12	217 5B	50	0.4
13	353 2	56	0.4
14	545 1	15	0.4
15	Dark 1	40	0.4
20	100 2A	37	0.4
21	100 2B	38	0.4
22	217 2	42	0.4
23	353 3A	47	0.4
24	353 3B	44	0.4
25	857 1	18	0.5
30	143 2A	42	0.4
31	143 2B	55	0.4
32	353 4A	50	0.4
33	353 4B	43	0.4
34	545 2	14	0.4
35	857 2	18	0.5
40	100 3A	39	0.4
41	100 3B	44	0.4
42	143 6	43	0.4
43	217 6A	43	0.4
44	217 6B	45	0.4
45	353 7	46	0.4
50	143 3A	47	0.4
51	143 3B	45	0.4
52	217 3	43	0.4
53	353 5A	43	0.4
54	353 5B	43	0.4
55	545 3	17	0.4
60	143 7	45	0.4
61	217 7A	50	0.4
62	217 7B	49	0.4
63	353 6A	45	0.4

Voie	Nom	$q$	$q/\sigma$
64	353 6B	46	0.4
65	857 3	21	0.5
70	143 8	51	0.4
71	217 8A	45	0.4
72	217 8B	47	0.4
73	545 4	15	0.4
74	857 4	18	0.5
75	Dark2	46	0.4
80	100 4A	36	0.4
81	100 4B	38	0.4
82	143 4A	42	0.4
83	143 4B	48	0.4
84	217 4	46	0.4
85	353 8	48	0.4
90	Ther 0.1K 1	16	0.5
91	Ther PID2 N	16	0.5
92	Ther PID1 N	671	0.5
93	Ther PID1.6 N	24	0.5
94	Ther 1.6K 1	27	0.5
95	Ther PID4 N	23	0.5
A0	Resistor	104	0.5
A1	Capa2	261	0.5
A2	Ther 0.1K 2	28	0.5
A3	Ther PID1.6 R	27	0.5
A4	Ther 4K H 1	22	0.5
A5	Ther 4K L 1	21	0.5
B0	Ther PID2 R	16	0.5
B1	Ther PID1 R	19	0.5
B2	Ther 1.6K 2	27	0.5
B3	Ther PID4 R	21	0.5
B4	Ther 4K H 2	21	0.5
B5	Ther 4K L 2	21	0.5

TABLE C.1 – Tableau récapitulatif des paramètres de compression.

# Annexe D

## Liste des acronymes

**CAMB** Code for Anisotropies in the Microwave Background  
**CMB** Cosmic Microwave Background  
**L1** Level 1  
**L2** Level 2  
**L3** Level 3  
**LS** Level Simulation  
**GMCA** Generalized Morphological Component Analysis  
**HEALPix** Hierarchical Equal Area isoLatitude Pixelization of a sphere  
**DPU** Data Processing Unit  
**ODB** Ordinateur de Bord  
**REU** Readout Electronic Unit  
**CBI** Cosmic Background Imager  
**DASI** Degree Angular Scale Interferometer  
**ACT** Atacama Cosmology Telescope  
**BOOMERanG** Balloon Observations Of Millimetric Extragalactic Radiation and Geophysics  
**MAXIMA** Millimeter Anisotropy eXperiment IMaging Array  
**ACBAR** Arcminute Cosmology Bolometer Array Receiver  
**SPT** South Pole Telescope  
**COBE** COsmic Background Experiment  
**ESA** European Space Agency  
**WMAP** Wilkinson Microwave Anisotropy Probe  
**COBRA** COsmic Background Radiation Anisotropy satellite  
**SAMBA** SAteLLite for Measurement of Background Anisotropies  
**IAS** Institut d'Astrophysique Spatiale  
**CNR** Consiglio Nazionale delle Ricerche (Centre national de la recherche italienne)  
**SREM** Standard Radiation Environment Monitor  
**MOC** Main Operation Center



# Bibliographie

- Afonso, C., Albert, J. N., Andersen, J., Ansari, R., Aubourg, É., Bareyre, P., Beaulieu, J. P., Blanc, G., Charlot, X., Couchot, F., Coutures, C., Ferlet, R., Fouqué, P., Glicenstein, J. F., Goldman, B., Gould, A., Graff, D., Gros, M., Haissinski, J., Hamadache, C., de Kat, J., Lasserre, T., Le Guillou, L., Lesquoy, É., Loup, C., Magneville, C., Marquette, J. B., Maurice, É., Maury, A., Milsztajn, A., Moniez, M., Palanque-Delabrouille, N., Perdureau, O., Prévot, L., Rahal, Y. R., Rich, J., Spiro, M., Tisserand, P., Vidal-Madjar, A., Vigroux, L., & Zylberajch, S. 2003, *A&A*, 400, 951
- Aghanim, N., Majumdar, S., & Silk, J. 2008, *Reports on Progress in Physics*, 71, 066902
- Alcock, C., Allsman, R. A., Alves, D., Axelrod, T. S., Becker, A. C., Bennett, D. P., Cook, K. H., Freeman, K. C., Griest, K., Guern, J., Lehner, M. J., Marshall, S. L., Peterson, B. A., Pratt, M. R., Quinn, P. J., Rodgers, A. W., Stubbs, C. W., Sutherland, W., & The MACHO Collaboration. 1996, *ApJ*, 471, 774
- Alpher, R. A. & Herman, R. C. 1948, *Physical Review*, 74, 1737
- Banta, E. 1965, *IEEE Trans. Inform. Theory*, 11, 114
- Barkana, R. & Loeb, A. 2001, *Phys. Rep.*, 349, 125
- Bartelmann, M. & Schneider, P. 2001, *Phys. Rep.*, 340, 291
- Bartlett, J. G. 1999, *New A Rev.*, 43, 83
- Bertin, E. & Arnouts, S. 1996, *A&AS*, 117, 393
- Blanchard, A. 1991, *Physical cosmology* (Editions Frontières) [LINK]
- Bobin, J., Moudden, Y., Starck, J.-L., Fadili, J., & Aghanim, N. 2008, *Statistical Methodology*, 5, 307, *astrostatistics* [LINK]
- Buehler, P., Desorgher, L., & Zehnder, A. 1996, in *ESA Special Publication*, Vol. 392, *Environment Modeling for Space-Based Applications*, ed. T.-D. Guyenne & A. Hilgers, 87–+
- Carvalho, P., Rocha, G., & Hobson, M. P. 2009, *MNRAS*, 393, 681
- Chan, R. H. & Ng, M. K. 1996, *SIAM Rev.*, 38, 427 [LINK]
- Clowe, D., Gonzalez, A., & Markevitch, M. 2004, *ApJ*, 604, 596
- Crittenden, R. G. & Turok, N. G. 1998, *ArXiv Astrophysics e-prints*
- Das, S. & Bode, P. 2008, *ApJ*, 682, 1
- 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
- Donoho, D. L. & Huo, X. 2001, *IEEE Trans. Inf. Theory*, 47, 2845
- Elad, M., Starck, J., Querre, P., & Donoho, D. 2005, *Applied and Computational Harmonic Analysis*, 19, 340 [LINK]

- Feichtinger, H. G., Gröchenig, K., & Strohmer, T. 1995, *Numerische Mathematik*, 69, 423, 10.1007/s002110050101 [LINK]
- Gaertner, S., Benoit, A., Lamarre, J., Giard, M., Bret, J., Chabaud, J., Desert, F., Faure, J., Jegoudez, G., Lande, J., Leblanc, J., Lepeltier, J., Narbonne, J., Piat, M., Pons, R., Serra, G., & Simiand, G. 1997, *A&AS*, 126, 151
- Gamow, G. 1948a, *Nature*, 162, 680
- . 1948b, *Physical Review*, 74, 505
- González, R. & Woods, R. 2002, *Digital image processing* (Prentice Hall of India) [LINK]
- Hamimeche, S. & Lewis, A. 2008, *Phys. Rev. D*, 77, 103013
- Hanson, D., Challinor, A., Efstathiou, G., & Bielewicz, P. 2011, *Phys. Rev. D*, 83, 043005
- Hanson, D., Challinor, A., & Lewis, A. 2010a, *General Relativity and Gravitation*, 42, 2197
- Hanson, D., Lewis, A., & Challinor, A. 2010b, *Phys. Rev. D*, 81, 103003
- Hinshaw, G., Nolta, M. R., Bennett, C. L., Bean, R., Doré, O., Greason, M. R., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Komatsu, E., Limon, M., Odegard, N., Meyer, S. S., Page, L., Peiris, H. V., Spergel, D. N., Tucker, G. S., Verde, L., Weiland, J. L., Wollack, E., & Wright, E. L. 2007, *ApJS*, 170, 288
- Hirata, C. M. & Seljak, U. 2003a, *Phys. Rev. D*, 67, 043001
- . 2003b, *Phys. Rev. D*, 68, 083002
- Hivon, E., Górski, K. M., Netterfield, C. B., Crill, B. P., Prunet, S., & Hansen, F. 2002, *ApJ*, 567, 2
- Hu, W. 2000, *Phys. Rev. D*, 62, 043007
- . 2001, *ApJ*, 557, L79
- Hu, W. & Okamoto, T. 2002, *ApJ*, 574, 566
- Kesden, M., Cooray, A., & Kamionkowski, M. 2003, *Phys. Rev. D*, 67, 123507
- Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Nolta, M. R., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2011, *ApJS*, 192, 18
- Kosowsky, A. 1996, *Annals of Physics*, 246, 49
- Larson, D., Dunkley, J., Hinshaw, G., Komatsu, E., Nolta, M. R., Bennett, C. L., Gold, B., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Page, L., Smith, K. M., Spergel, D. N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2010, *ArXiv e-prints*
- Lavaux, G. & Wandelt, B. D. 2010, *ApJS*, 191, 32
- Leach, S. M., Cardoso, J., Baccigalupi, C., Barreiro, R. B., Betoule, M., Bobin, J., Bonaldi, A., Delabrouille, J., de Zotti, G., Dickinson, C., Eriksen, H. K., González-Nuevo, J., Hansen, F. K., Herranz, D., Le Jeune, M., López-Cañiego, M., Martínez-González, E., Massardi, M., Melin, J., Miville-Deschênes, M., Patanchon, G., Prunet, S., Ricciardi, S., Salerno, E., Sanz, J. L., Starck, J., Stivoli, F., Stolyarov, V., Stompor, R., & Vielva, P. 2008, *A&A*, 491, 597
- Lewis, A. 2005, *Phys. Rev. D*, 71, 083008
- Lewis, A. & Challinor, A. 2006, *Phys. Rep.*, 429, 1

- Lineweaver, C. H. 1997, in *Astronomical Society of the Pacific Conference Series*, Vol. 126, From Quantum Fluctuations to Cosmological Structures, ed. D. Valls-Gabaud, M. A. Hendry, P. Molaro, & K. Chamcham, 185–+
- Lyth D., L. R. 2009, *The primordial density fluctuations* (Cambridge university press)
- Masnou, S. & Morel, J.-m. 1998, *ICIP*, ed. IEEE, 3, 259 [linklabel](#)
- Mather, J. C. 1982, *Appl. Opt.*, 21, 1125 [[LINK](#)]
- Melin, J., Bartlett, J. G., & Delabrouille, J. 2006, *A&A*, 459, 341
- Montier, L. 2005, PhD thesis, Centre d'Etude Spatiale des Rayonnements
- Montier, L. A., Pelkonen, V., Juvela, M., Ristorcelli, I., & Marshall, D. J. 2010, *A&A*, 522, A83+
- Okamoto, T. & Hu, W. 2003a, *Phys. Rev. D*, 67, 083002
- . 2003b, *Phys. Rev. D*, 67, 083002
- Oppenheim, A. V. & Schafer, R. 1975, *Digital signal processing*, Prentice-Hall international editions (Prentice-Hall)
- Partridge, R. 1995, *3K : the cosmic microwave background radiation*, Cambridge astrophysics series (Cambridge University Press) [[LINK](#)]
- Peacock, J. A., Cole, S., Norberg, P., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R. D., Colless, M., Collins, C., Couch, W., Dalton, G., Deeley, K., De Propriis, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Percival, W. J., Peterson, B. A., Price, I., Sutherland, W., & Taylor, K. 2001, *Nature*, 410, 169
- Peebles, P. J. E. 1982, *ApJ*, 263, L1
- Peebles, P. J. E. & Wilkinson, D. T. 1968, *Phys. Rev.*, 174, 2168
- Peebles, P. J. E. & Yu, J. T. 1970, *ApJ*, 162, 815
- Penzias, A. A. & Wilson, R. W. 1965, *ApJ*, 142, 419
- Perotto, L. 2006, PhD thesis, Université Denis Diderot - Paris VII
- Perotto, L., Bobin, J., Plaszczynski, S., Starck, J.-L., & Lavabre, A. 2010, *A&A*, 519, A4 [[LINK](#)]
- Piat, M. 2000, PhD thesis, Université de Paris-Sud (Paris XI)
- Pires, S., Starck, J., Amara, A., Teyssier, R., Réfrégier, A., & Fadili, J. 2009, *MNRAS*, 395, 1265
- Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Balbi, A., Banday, A. J., Barreiro, R. B., & et al. 2011, *ArXiv e-prints*
- Planck HFI Core Team, Ade, P. A. R., Aghanim, N., Ansari, R., Arnaud, M., Ashdown, M., Aumont, J., Banday, A. J., Bartelmann, M., Bartlett, J. G., Battaner, E., Benabed, K., Benot, A., Bernard, J., Bersanelli, M., Bhatia, R., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bradshaw, T., Brelle, E., Bucher, M., Camus, P., Cardoso, J., Catalano, A., Challinor, A., Chamballu, A., Charra, J., Charra, M., Chary, R., Chiang, C., Church, S., Clements, D. L., Colombi, S., Couchot, F., Coulais, A., Cressiot, C., Crill, B. P., Crook, M., de Bernardis, P., Delabrouille, J., Delouis, J., Désert, F., Dolag, K., Dole, H., Doré, O., Douspis, M., Efstathiou, G., Eng, P., Filliard, C., Forni, O., Fosalba, P., Fourmond, J., Ganga, K., Giard, M., Girard, D., Giraud-Héraud, Y., Gispert, R., Gorski, K. M., Gratton, S., Griffin, M., Guyot, G., Haissinski, J., Harrison, D., Helou, G., Henrot-Versillé, S., Hernandez-Monteagudo, C., Hildebrandt, S. R., Hills, R., Hivon, E., Hobson, M., Holmes, W. A., Huenberger, K. M., Jae, A. H., Jones, W. C., Kaplan, J., Kneissl, R., Knox, L., Lagache, G., Lamarre, J., Lami, P., Lange, A. E., Lasenby, A., Lavabre, A., Lawrence, C. R., Leriche, B., Leroy, C., Macas-Perez, Y. L. J. F., Maciaszek, T.,



- MacTavish, C. J., Maei, B., Mandolesi, N., Mann, R., Mansoux, B., Masi, S., Matsumura, T., McGehee, P., Melin, J., Mercier, C., Miville-Deschnes, M., Moneti, A., Montier, L., Mortlock, D., Murphy, A., Nati, F., Nettereld, C. B., Norgaard-Nielsen, H. U., North, C., Noviello, F., Novikov, D., Osborne, S., Paine, C., Pajot, F., Patanchon, G., Peacocke, T., Pearson, T. J., Perdereau, O., Perotto, L., Piacentini, F., Piat, M., Plaszczynski, S., Pointecouteau, E., Pons, R., Ponthieu, N., Prezeau, G., Prunet, S., Puget, J., Reach, W. T., Renault, C., Ristorcelli, I., Rocha, G., Rosset, C., Roudier, G., Rowan-Robinson, M., Rusholme, B., Santos, D., Savini, G., Schaefer, B. M., Shellard, P., Spencer, L., Starck, J., Stassi, P., Stolyarov, V., Stompor, R., Sudiwala, R., Sunyaev, R., Sygnet, J., Tauber, J. A., Thum, C., Torre, J., Touze, F., Tristram, M., Van Leeuwen, F., Vibert, L., Vibert, D., Wade, L. A., Wandelt, B. D., White, S. D. M., Wiesemeyer, H., Woodcraft, A., Yurchenko, V., Yvon, D., & Zacchei, A. 2011, ArXiv e-prints
- Planck Science Team. 2005, Planck : The Scientific Programme (Blue Book) (ESA-SCI)
- Ponthieu, N., Macías-Pérez, J. F., Tristram, M., Ade, P., Amblard, A., Ansari, R., Aumont, J., Aubourg, É., Benoît, A., Bernard, J., Blanchard, A., Bock, J. J., Bouchet, F. R., Bourrachot, A., Camus, P., Cardoso, J., Couchot, F., de Bernardis, P., Delabrouille, J., Désert, F., Douspis, M., Dumoulin, L., Filliatre, P., Fosalba, P., Giard, M., Giraud-Héraud, Y., Gispert, R., Grain, J., Guglielmi, L., Hamilton, J., Hanany, S., Henrot-Versillé, S., Kaplan, J., Lagache, G., Lange, A. E., Madet, K., Maffei, B., Masi, S., Mayet, F., Nati, F., Patanchon, G., Perdereau, O., Plaszczynski, S., Piat, M., Prunet, S., Puget, J., Renault, C., Rosset, C., Santos, D., Vibert, D., & Yvon, D. 2005, *A&A*, 444, 327
- Press, W. 2007, Numerical recipes : the art of scientific computing (Cambridge University Press) [LINK]
- Puget, J., Abergel, A., Bernard, J., Boulanger, F., Burton, W. B., Desert, F., & Hartmann, D. 1996, *A&A*, 308, L5+
- Rich, J. 2004, Principes de la cosmologie (École polytechnique) [LINK]
- Slipher, V. M. 1918, *Lowell Observatory Bulletin*, 3, 63
- Smith, M. S., Kawano, L. H., & Malaney, R. A. 1993, *ApJS*, 85, 219
- Smoot, G. F., Bennett, C. L., Kogut, A., Wright, E. L., Aymon, J., Boggess, N. W., Cheng, E. S., de Amici, G., Gulkis, S., Hauser, M. G., Hinshaw, G., Jackson, P. D., Janssen, M., Kaita, E., Kelsall, T., Keegstra, P., Lineweaver, C., Loewenstein, K., Lubin, P., Mather, J., Meyer, S. S., Moseley, S. H., Murdock, T., Rokke, L., Silverberg, R. F., Tenorio, L., Weiss, R., & Wilkinson, D. T. 1992, *ApJ*, 396, L1
- Sunyaev, R. A. & Zeldovich, Y. B. 1970, *Ap&SS*, 7, 3
- Tristram, M., Filliard, C., Perdereau, O., Plaszczynski, S., Stompor, R., & Touze, F. 2011, ArXiv e-prints
- Tristram, M. & Ganga, K. 2007, *Reports on Progress in Physics*, 70, 899
- Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, *Nature*, 279, 381
- White, M., Carlstrom, J. E., Dragovan, M., & Holzzapfel, W. L. 1999, *ApJ*, 514, 12
- Widrow, B. & Kollár, I. 2008, Quantization noise : roundoff error in digital computation, signal processing, control, and communications (Cambridge University Press) [LINK]
- Zaldarriaga, M. & Seljak, U. 1997, *Phys. Rev. D*, 55, 1830
- Zeldovich, Y. B. & Sunyaev, R. A. 1969, *Ap&SS*, 4, 301