UNIVERSITE DE CERGY-PONTOISE

THÈSE

présentée pour obtenir le grade de

Docteur de l'Université de CERGY-PONTOISE

Spécialité : *Sciences et Ingénierie*

par

PIERRICK ABRIAL

Sujet :

Transformations multi-échelles sur la sphère et applications

Soutenue le 18 Décembre 2008 devant la Commission d'examen :

M.	A.. BIJAOUI	Rapporteur, Président du Jury
M.	F. MURTAGH	Rapporteur
M.	J. DELABROUILLE	Examinateur
M.	J.-L. STARCK	Directeur de thèse
Mme.	M. NGUYEN	Co-directrice de thèse

Équipes Traitement de l'Image et du Signal (ETIS)
CNRS UMR 8051 / ENSEA / Université de Cergy-Pontoise
F-95000, Cergy-Pontoise Cedex, France

Laboratoire CEA-IRFU/SEDI-Sap
CEA/Saclay F - 91191 Gif sur Yvette Cedex

Résumé

Le fond diffus cosmologique (ou CMB) est d'une grande importance en cosmologie car c'est le vestige le plus ancien que l'on peut observer de l'Univers. Le satellite Planck qui sera lancé début 2009 permettra d'effectuer les observations les plus précises jamais obtenues sur tout le ciel de ce rayonnement. Le traitement de ces données va demander une amélioration considérable des techniques d'analyse existantes sur la sphère.

Sur le plan, il existe un grand nombre de représentations multi-échelles qui permettent de transformer un signal sous une forme qui permet de mieux l'analyser. Sur la sphère, l'analyse de données multi-échelles en est à ses premiers balbutiements. La transformée la plus utilisée reste la transformée en harmoniques sphériques, l'analogue de la transformée de Fourier sur la sphère.

Dans cette thèse, nous introduisons de nouvelles transformées multi-échelles sur la sphère qui possèdent la propriété très importante d'être inversibles : une transformée en ondelettes isotropes décimée et non-décimée, une transformée en ridgelets et une transformée en curvelets.

Nous utilisons ensuite ces nouvelles transformées pour améliorer l'étude des données CMB. Dans cette thèse, plusieurs applications sont présentées : le débruitage, l'interpolation des données manquantes et la détection de non-Gaussianité.

Abstract

The analysis of the slight fluctuations in the Cosmic Microwave Background radiation field (CMB) is a major issue in modern cosmology as these are strongly related to the cosmological scenarios describing the properties and evolution of our Universe. Indeed, according to current models, the map of CMB fluctuations is an imprint of primordial fluctuations in matter density around 30000 years after the Big Bang. Full-sky multi-spectral observations of the CMB with unprecedented sensitivity and angular resolution are expected from the ESA's Planck mission, which is to be launched early 2009. The statistical analysis of these data requires new data processing tools on the sphere and it is precisely the goal of this thesis to extend existing euclidean 2D multiscale decompositions to the spherical topology.
We developed an invertible isotropic wavelet transform on the sphere based on the spherical harmonics transform much in the spirit of the *à trous* wavelet transform. Thanks to the geometry of the Healpix pixelization scheme, the proposed undecimated wavelet transform is easily extended to a pyramidal transform.
Then, building on the isotropic wavelet transform on the sphere and Healpix, we describe an implementation of digital ridgelets and curvelets on the sphere, extending the ridgelet and curvelet transforms which had a major impact in flat image processing. Several applications to CMB data analysis are then described. The proposed multiscale methods on the sphere are shown to be valuable for data restoration and denoising. Other issues such as testing for non gaussianity in the CMB, or gap filling in incomplete data maps are also investigated.

Table des matières

Résumé . . . . **III**

Abstract . . . . **V**

Table des figures **XI**

Introduction **1**

1 Prologue **5**
- 1.1 Fond diffus cosmologique 5
- 1.2 Mission spatiale Planck 6
- 1.3 Reconstruction du fond diffus cosmologique 11
- 1.4 Analyse du fond diffus cosmologique 13
- 1.5 Les ondelettes et le CMB 13

2 Pixélisation **17**
- 2.1 Introduction 17
- 2.2 ECP : Equidistant Coordinate Partition 19
- 2.3 Icosaedre 19
- 2.4 COBE Quad Cube 21
- 2.5 IGLOO 21
- 2.6 HEALPix 22
- 2.7 HTM - Hierarchical Triangular Mesh 27
- 2.8 Gauss-Legendre Sky Pixelisation (GLESP) 27
- 2.9 Conclusions 30

3 Transformée en ondelettes sur la sphère **33**
- 3.1 Introduction 33
- 3.2 Transformées en ondelettes sur la sphère existantes 34
  - 3.2.1 Transformée en ondelettes de Haar sur la sphère . . . 34
  - 3.2.2 Les ondelettes axisymétriques par projection stereoscopique inverse. 35
  - 3.2.3 Ondelette directionelle sur la sphère. 39
  - 3.2.4 Conclusions 43

3.3 Nouvelles transformées/reconstructions en ondelettes sur la sphère . . . 43
3.3.1 Transformée en ondelettes isotrope non décimées sur la sphère . . . 43
3.3.2 Transformée en ondelettes isotrope pyramidale sur la sphère . . . 49
3.4 Conclusion . . . 52

4 Ridgelets et Curvelets sur la sphère 53
4.1 Introduction . . . 53
4.2 Transformée en ridgelets sur la sphère . . . 54
4.2.1 Transformée en ridgelets sur le plan . . . 54
4.2.2 Transformée en ridgelets sur la sphère. . . . 58
4.3 Transformée en curvelets sur la sphère . . . 59
4.3.1 Transformée en curvelets sur le plan . . . 59
4.3.2 Transformée en curvelets sur la sphère . . . 63
4.3.3 Transformée en curvelets pyramidale sur la sphère . . 64
4.4 Conclusions . . . 66

5 Débruitage multi-échelles sur la sphère 67
5.1 Introduction au débruitage . . . 68
5.2 Coefficients significatifs . . . 68
5.2.1 Définition . . . 68
5.2.2 Modélisation du bruit . . . 69
5.2.3 Estimation des caractéristiques d'un bruit gaussien . . 70
5.3 Méthodes de seuillage . . . 71
5.3.1 Le seuillage *dur* et le seuillage *doux* . . . 71
5.3.2 La méthode de seuillage FDR . . . 72
5.4 Méthode de filtrage combinée sur la sphère . . . 75
5.4.1 Principe . . . 76
5.4.2 Algorithme . . . 77
5.4.3 Résultats . . . 77
5.5 Conclusion . . . 79

6 Analyse en Composantes Morphologiques et Inpainting sur la Sphère 81
6.1 MCA sur la sphère . . . 82
6.1.1 Représentations parcimonieuses . . . 82
6.1.2 MCA : Principe et Algorithme . . . 85
6.1.3 Expériences numériques et applications . . . 88
6.2 *Inpainting* sur la sphère . . . 93
6.2.1 Principe et Algorithme . . . 94
6.2.2 Experiences Numériques - Application en Astrophysique 97
6.3 Conclusion . . . 104

7 Détection de non-Gaussianités dans le CMB sur la sphère 107

7.1 Introduction 107

7.2 Les ondelettes en action 108

7.3 Discrimination de non-Gaussianité 110

7.3.1 Sources ponctuels 110

7.3.2 Analyse statistique des coefficients ondelettes et curvelets 111

7.4 Débruitage et Non-Gaussianité 114

7.5 WMAP et le *Cold Spot* 115

7.6 Conclusion 120

Conclusion 123

Annexe 125

Bibliographie 127

Table des figures

1.1 Carte du fond diffus cosmologique obtenue par le satellite WMAP. 6
1.2 Spectre de puissance théorique du CMB. 7
1.3 Puissance des différentes composantes du ciel en fonction de leur fréquence. 11

2.1 Une pixélisation ecp de la sphère. 20
2.2 La pixelisation icosahèdre de Tegmark. 20
2.3 La pixélisation COBE Quad Cube. 21
2.4 La pixélisation IGLOO 3 :6 :3. 22
2.5 Visualisation des pôles de la pixélisation 12116 pixels. 23
2.6 Projection cylindrique de différentes pixélisations possibles. . 24
2.7 Projection orthographique de différentes pixélisations possibles. 24
2.8 Différentes échelles de la pixélisation HEALPix. 25
2.9 Comparaison de la séparation en θ, ϕ pour différentes pixélisations. 26
2.10 Passage d'une pixélisation à une autre par la numérotation en chaînes. 27
2.11 Différentes résolutions de la pixélisation HTM. 28
2.12 Comparaison de l'erreur de calcul des C_l avec HEALPix et GLESP. 28
2.13 Comparaison des shémas de pixélisation entre Healpix et Glesp. 29
2.14 Comparaison du nombre de pixels par anneau avec la pixélisation Healpix (à gauche) et la pixélisation Glesp (à droite). 30
2.15 Comparaison de la position angulaire des anneaux avec la pixélisation Healpix (à gauche) et la pixélisation Glesp (à droite). 30

3.1 Passage d'une échelle à la suivante par une transformation de Haar. 35
3.2 Coefficients d'échelle d'une transformation de Haar. 36
3.3 Projection stéréographique inverse d'une fonction radiale du plan sur la sphère. 36

3.4 Ondelette chapeau mexicain pour des coefficients de dilatation $a = \{1, 2, 4, 8\}$. 38
3.5 Simulation d'une carte de CMB et sa transformation ondelette chapeau mexicain. 39
3.6 Projection stéréographique inverse d'une ondelette directionnelle sur la sphère. 40
3.7 Ondelette chapeau mexicain 41
3.8 Ondelette de Morlet sur la sphère. 42
3.9 À gauche, la fonction d'échelle $\hat{\phi}$ et à droite la fonction ondelette $\hat{\psi}$. 45
3.10 Carte du CMB obtenue par WMAP et sa transformée ondelette sur la sphère 46
3.11 Filtre passe-haut et filtre passe-bas associés 47
3.12 Projection inverse d'un coefficient ondelette à différentes echelles. 48
3.13 Carte de CMB obtenue par WMAP et sa transformée ondelette pyramidale sur la sphère. 50

4.1 Exemples de fonctions ridgelets. 54
4.2 A gauche, image contenant deux lignes et du bruit Gaussien. A droite, sa transformée de Radon. 55
4.3 Diagramme de l'algorithme de la transformée en ridgelets. . . 57
4.4 Diagramme de la transformée en ridgelets sur la sphère. . . . 59
4.5 Projection inverse de coefficients ridgelets à différentes échelles et à différentes orientations. 60
4.6 Diagramme de la transformée en curvelets discrète. 62
4.7 Quelques exemples de curvelets. 63
4.8 Diagramme de l'algorithme de la transformée en curvelets sur la sphère. 64
4.9 Projection inverse de coefficients curvelets à différentes échelles et différentes orientations 65

5.1 Débruitage par les ondelettes et par les curvelets 73
5.2 Détermination graphique du seuil FDR 74
5.3 Débruitage par la méthode de filtrage combiné 78
5.4 Méthode de filtrage combinée 80

6.1 Photo de Barbara séparée en contours et textures en utilisant MCA avec un dictionnaire de cosinus locaux et de curvelets. . 85
6.2 Lignes et sources ponctuelles séparées en utilisant MCA avec un dictionnaire d'ondelettes et de ridgelets. 86
6.3 Séparation d'une harmonique sphérique et d'un profil gaussien avec MCA. 89
6.4 Séparation d'une harmonique sphérique et d'un profil gaussien avec MCA. 90

6.5 Structures en surface d'une capsule sphérique utilisée pour la fusion à confinement inertiel. 91
6.6 Plan lissé de la transformation ondelette isotrope non-décimée sur la sphère de la carte des défauts surfaciques d'une cible. . 91
6.7 Etude des structures en surface d'une capsule sphérique . . . 92
6.8 Exemple d'inpainting numérique sur une oeuvre d'art 94
6.9 Interpolation d'une sinusoide tronquée 95
6.10 Application de l'algorithme MCA-inpainting sur la sphère . . 98
6.11 Courbe représentant l'erreur d'approximation non-linéaire . . 100
6.12 Inpainting de la carte de CMB. 102
6.13 Décomposition en ondelettes de la carte inpaintée. 103
6.14 Spectre de puissance du CMB estimé. 104
6.15 Skewness et kurtosis par échelle des cartes de CMB originale et inpaintée 105

7.1 Cartes simulées du CMB, de l'effet SZ et de cordes cosmiques 110
7.2 Images contenant des gaussiennes et des droites avec et sans bruit. 112
7.3 Kurtosis normalisé en fonction du rapport signal-sur-bruit pour les coefficients ondelettes et curvelets 113
7.4 Séparation de composantes : CMB et composantes galactiques. 116
7.5 Trois échelles de la transformée en ondelettes de la carte de CMB-WMAP, avec la position du *cold spot* indiquée par une ellipse. 117
7.6 Débruitage de la carte WMAP à 4σ. 118
7.7 En haut, *cold spot* extrait de la carte WMAP ; en bas, carte WMAP où on a ajouté un second *cold spot*. 119
7.8 Décomposition de la carte WMAP+*cold spot* en deux composantes : une Gaussienne et une non-Gaussienne 120

Introduction

Cette thèse porte sur l'application des méthodes multi-échelles sur la sphère en particulier pour l'analyse du fond diffus cosmologique. De nombreux outils pour le traitement des données sur la sphère sont introduits.

Le fond diffus cosmologique souvent appelé CMB (pour Cosmic Microwave Background) est un rayonnement électromagnétique quasi-uniforme qui témoigne de l'époque dense et chaude qu'a connue l'Univers il y a environ 13 milliards d'années, conformément aux predictions faites par le modèle du Big Bang. Il correspond à la période de découplage entre la matière et le rayonnement au moment où les électrons se sont combinés avec les noyaux atomiques par un processus de recombinaison, laissant les photons traverser librement l'Univers. Ce rayonnement fût observé pour la première fois par Penzias et Wilson en 1964 grâce à une antenne radio.

L'observation du fond diffus cosmologique permet de comprendre l'histoire de l'Univers depuis ses premiers instants en nous donnant une vision de l'Univers à ses débuts. Ainsi, afin de sonder les débuts de l'Univers, un grand nombre d'expériences au sol et dans l'espace ont été effectuées afin de mesurer ce rayonnement avec précision. De nombreuses expériences ont été menées au sol, dans des ballons stratosphériques comme BOOMERanG (lancé en 1998) et Archeops (lancé en 1999) ou dans l'espace grâce à des satellites : COBE (lancé en 1989), WMAP (lancé en 2001) et bientôt Planck.

Les données Planck qui seront bientôt disponibles vont permettre d'etudier le fond diffus cosmologique avec une précision largement meilleure que celle de WMAP. Ces données comme celle de COBE ou de WMAP couvrent la totalité du ciel. Habituellement, les traitements de données sur la sphère sont basés uniquement sur la transformée en harmoniques sphériques. Afin d'exploiter au mieux ces nouvelles données de nouveaux outils de traitement doivent être développés.

Dans cette thèse, nous proposons d'utiliser des méthodes multi-échelles afin de représenter le signal sous une forme qui permet de mieux l'analyser. Nous développons ainsi de nouvelles transformées sur la sphère que nous utilisons pour améliorer la reconstruction et l'analyse du fond diffus cosmologique.

*
* *

Cette thèse est organisée de la manière suivante :

⋆ Le premier chapitre est consacré à la description du fond diffus cosmologique et du projet Planck. Nous introduisons notamment les bases théoriques nécessaires pour la reconstruction et l'analyse du fond diffus cosmologique à partir des données Planck.
⋆ Dans le chapitre II, nous présentons différentes méthodes de pixélisation existantes dans la littérature. Deux méthodes ont été sélectionnées pour leur précision : HEALPix, le standard en traitement de données sur la sphère, et GLESP, plus précis mais moins documenté. HEALPix et GLESP seront utilisés tout au long de cette thèse.
⋆ Dans le chapitre III, nous faisons, tout d'abord, l'état-de-l'art des différentes méthodes d'analyse multi-échelles, que l'on peut trouver pour traiter les données sur la sphère. Avant notre contribution, les seules représentations multi-échelles disponibles sur la sphère sont des transformées en ondelettes continues, ne disposant pas d'algorithme de reconstruction. Ensuite, nous décrivons une nouvelle transformée en ondelettes sur la sphère que nous avons développée. Cette transformée en ondelettes à des similarités avec l'algorithme "à trous". Elle est isotrope et possède une reconstruction simple et exacte.
⋆ Dans le chapitre IV, nous introduisons de nouvelles représentations multi-échelles : la transformée en ridgelets et la transformée en curvelets. Nous proposons des algorithmes permettant une transformation et une reconstruction exacte sur la sphère de ces représentations.
⋆ Le chapitre V est consacré à la description des méthodes de filtrage sur la sphère. Plusieurs méthodes de filtrage multi-échelles sont présentées qui pourront notamment être utilisées pour l'exploitation des composantes galactiques.
⋆ Dans le chapitre VI, nous présentons la méthode d'analyse en composantes morphologiques (MCA) qui permet de séparer le fond diffus cosmologique des autres composantes galactiques. Nous présentons une nouvelle méthode d'interpolation des données manquantes, dérivée de cette méthode MCA, qui permet de résoudre les problèmes liés à la présence du masque galactique.
⋆ Dans le chapitre VII, nous présentons plusieurs statistiques qui ont pour but de tester la non-Gaussianité dans les données de CMB. Nous faisons notamment une étude sur les données WMAP.

- ⋆ Dans le dernier chapitre, nous donnons les conclusions du travail qui a été mené pendant cette thèse ainsi que les perspectives.
- ⋆ L'annexe liste les publications et autres contributions qui reposent sur le travail présenté dans ce manuscrit.

Chapitre 1
Prologue

Sommaire

1.1 Fond diffus cosmologique 5
1.2 Mission spatiale Planck 6
1.3 Reconstruction du fond diffus cosmologique 11
1.4 Analyse du fond diffus cosmologique 13
1.5 Les ondelettes et le CMB 13

1.1 Fond diffus cosmologique

Le fond diffus cosmologique (Cosmological Microwave Background - CMB) est considéré par les cosmologistes comme l'une des clefs de voûte des théories actuelles de l'expansion de l'Univers (théorie du Big Bang). Le CMB[1] est le reliquat du rayonnement le plus ancien (d'où son appellation traditionnelle de "rayonnement fossile"). Selon la théorie du Big Bang, le fond diffus cosmologique est la trace du passage rapide de l'état opaque (lié à une forte ionisation de l'Univers) à transparent de l'Univers. Cet événement dont le déroulement serait daté de $3.10^5 - 4.10^5$ années après le Big Bang constitue donc la "lumière" la plus ancienne observable actuellement. Elle revêt une importance capitale pour les cosmologistes puisque le comportement très particulier du CMB est riche d'enseignements.
La théorie du Big Bang prévoit que le fond diffus cosmologique soit un champ d'émission isotrope gaussien caractérisé par son spectre de puissance. Spectralement[2], le CMB se comporte comme un corps noir dont la température est de l'ordre de $2,735K$.

[1]Observé la première fois par Penzias et Wilson en 1964, cette découverte leur a valu le prix Nobel de physique en 1978.
[2]Comportement par bande de fréquences.

Cette prévision fait du CMB un signal apparemment simple mais riche d'informations pour les cosmologistes. Le rayonnement CMB observé par Planck sera mesuré sur une sphère dont l'observateur est le centre. Dans ce cas, le CMB est un champ gaussien caractérisé par son spectre de puissance en harmoniques sphériques[3]. Mathématiquement, la matrice de covariance du CMB dans l'espace des harmoniques sphériques est diagonale ; celle-ci est appelée spectre de puissance. Ce dernier est une fonction de l'échelle angulaire l (équivalent de la fréquence sur la sphère) : $\mathcal{C}(\ell)$.

La figure 1.1 montre le fond diffus cosmologique obtenu à partir des différents canaux du satellite WMAP, et la figure 1.2 montre le spectre théorique du CMB normalisé ainsi que différentes mesures effectuées par les dernières missions de mesure du CMB. On peut observer sur la figure 1.2 que le spectre de puissance du CMB présente une série d'oscillations (appelées "modes acoustiques"). Les caractéristiques de ces différents modes (position et amplitude) permettent en particulier d'apporter des contraintes sur les paramètres cosmologiques intervenant dans les modèles d'expansion (voir Lachièze-Rey 2005). L'estimation du CMB et en particulier de son spectre de puissance est ainsi cruciale pour contraindre les modèles de Big Bang qui décrivent l'évolution de l'Univers.

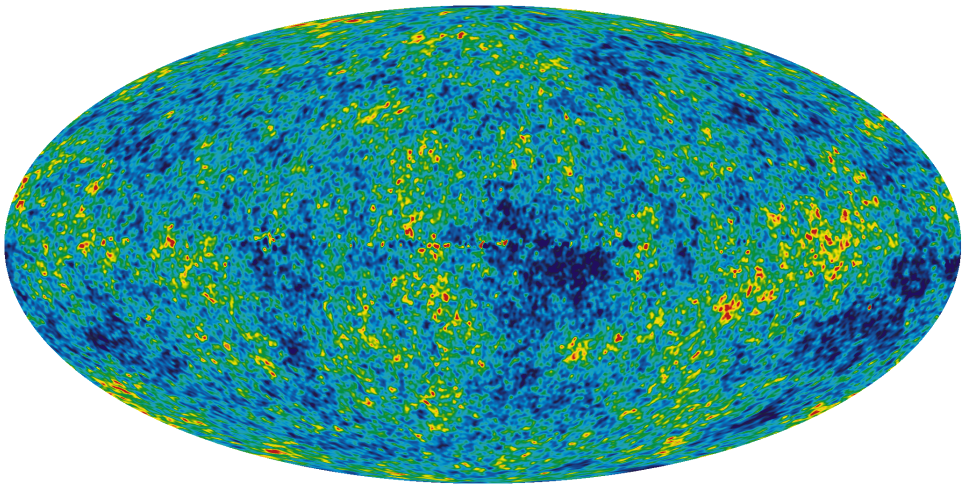

FIG. 1.1 – Carte du fond diffus cosmologique obtenue par le satellite WMAP. Une combinaison linéaire a été utilisée pour obtenir cette carte à partir des différents canaux.

1.2 Mission spatiale Planck

Au cours du premier semestre de l'année 2009, une nouvelle mission d'observation du CMB sera lancée : la mission Planck de l'Agence Spatiale

[3]Equivalent de l'espace de Fourier pour la sphère.

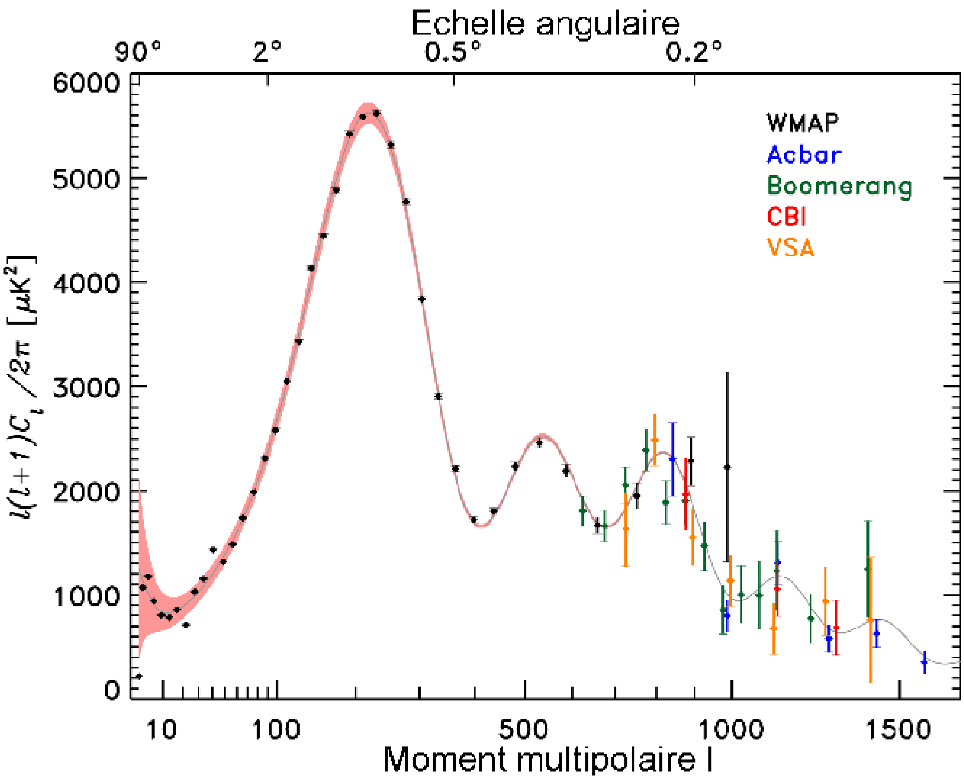

FIG. 1.2 – Spectre de puissance théorique du CMB et mesures issues des dernières missions d'observation.

Européenne (ESA)[4]. Cette mission fait suite à une série de missions spatiales (COBE, WMAP) qui ont déjà apporté de nombreuses informations sur le CMB (Jungman 1996). La dernière de ces missions, WMAP, a observé la totalité du ciel (sphère entière) sur 5 canaux de fréquences variant de 20 à 90 GHz avec une résolution de 15 minutes d'arc et a permis de caractériser avec précision les deux premiers modes du spectre de puissance du CMB.

Le défi porté par Planck est l'observation du ciel à une résolution plus fine (allant jusqu'à 5 minutes d'arc) sur 9 canaux de 30 à 857 GHz. Planck est composé de deux instruments : *Low Frequency Instrument* (LFI) et *High Frequency Instrument* (HFI).

L'instrument LFI

Cet instrument est composé de trois récepteurs radio accordés opérant aux fréquences de 30, 44 et 70 Ghz. Le beam effectif correspondant est de 33, 24 et 14 minutes d'arc. Ces récepteurs radios sont capables de détecter l'intensité totale ainsi que la polarisation des photons reçus. Les détecteurs

[4]Nous invitons le lecteur à se reporter au site *http ://planck.esa.int*.

de cet instrument sont basés sur des transistors HEMT (Heavy Expanded Mobility Tactical Truck).

L'instrument HFI

L'instrument HFI est composé de 54 bolomètres travaillant entre 90 et 130 mK. Ces bolomètres sont configurés pour opérer dans les bandes : 100, 143, 217, 353, 545 et 857 GHz. La résolution sera respectivement de 10, 7.1, 5, 5, 5, 5 minutes d'arc. Les quatre premiers canaux seront sensibles à l'intensité et à la polarisation des photons, les deux derniers seront uniquement sensibles à l'intensité des photons.

Outre le fond diffus cosmologique, chaque canal Planck capturera différentes composantes astrophysiques (Bouchet and Gispert 1999) que nous allons décrire par la suite.

Fond diffus cosmologique

Les connaissances actuelles du fond diffus cosmologique sont les suivantes :

- **Loi d'émission :** sous l'angle de la modélisation de données multicanales, l'émission de fond diffus cosmologique à la fréquence ν, $\mathbf{x}_\nu^{cmb}$ est le produit de la carte des fluctuations du CMB, ΔT^{cmb} (indépendante de ν) et d'une loi d'émission définie comme la dérivée de celle d'un corps noir à la température du CMB[5] ($\bar{T} \simeq 2,735K$) :

$$\mathbf{x}_\nu^{cmb} = \Delta T^{cmb} \left[\frac{\partial B_\nu(T)}{\partial T} \right]_{T=\bar{T}} \tag{1.1}$$

  De sorte que la composante CMB, $\mathbf{X}^{cmb}$ contenue dans les données $\mathbf{X}$ peut s'écrire sous la forme d'une composante de rang 1, dans un modèle de fluctuations linéaires autour de $\bar{T}$. En d'autres termes, le CMB est présent dans les données $\mathbf{X}$ sous la forme d'un mélange parfait dont la colonne de mélange sera supposée connue :

$$\mathbf{a}^{cmb} = \left[\left[\frac{\partial B_\nu(T)}{\partial T} \right]_{T=\bar{T}} \right]_{\nu=\{30,\cdots,857\mathrm{GHz}\}} \tag{1.2}$$

  La source correspondante, $\mathbf{s}^{cmb} = \Delta T^{cmb}$ sera définie comme la carte des fluctuations de température du fond diffus cosmologique. Notons que cette loi d'émission prévue par la théorie est validée par les observations du CMB antérieures : spatiales telles que COBE et WMAP ou des sondes atmosphériques telles que Boomerang et Archeops.

[5] Par la suite, les données seront manipulées en température d'antenne.

– **Stationnarité et caractère gaussien :** la carte des fluctuations de température que l'on notera $\mathbf{s}^{cmb}$ est, en théorie, un champ gaussien stationnaire entièrement défini par son spectre de puissance $C(\ell)$ dans l'espace des harmoniques sphériques. En réalité, la recherche de non-gaussianités dans les données du CMB est également réalisée pour tester certains modèles d'inflation de l'Univers. En l'état actuel des connaissances sur le CMB, le caractère gaussien de $\mathbf{s}^{cmb}$ peut être considéré comme vérifié. Les observations Planck, de part leur résolution, pourraient permettre d'apporter de nouvelles données pour tester le caractère gaussien du champ de fluctuation de température du fond diffus cosmologique.

– **Spectre de puissance :** comme nous l'avons souligné, sous hypothèse du caractère gaussien du fond diffus cosmologique, le champ $\mathbf{s}^{cmb}$ est statistiquement déterminé par son spectre de puissance dans l'espace des harmoniques sphériques $C(\ell)$. Ce spectre de puissance (angulaire puisque les données sont définies sur la sphère), est connu avec une faible erreur pour des valeurs de ℓ pouvant atteindre $\ell = 800$ notamment grâce aux données WMAP.

Composantes galactiques

Les composantes galactiques sont directement liées aux émissions de la voie lactée :

– **Emissions Synchrotron :** les émissions synchrotron sont provoquées par le mouvement en spirale de particules chargées dans un champ magnétique. Dans notre galaxie, de tels effets sont particulièrement intenses au niveau du centre galactique et peuvent s'étendre sur de plus hautes latitudes (par convention, l'équateur de la sphère d'observation). Le cube de données synchrotron est modélisé comme le produit d'un champ avec une loi d'émission suivant une loi de puissance $\nu^{-\beta}$ pour chaque pixel i de la sphère :

$$\forall \nu; \quad x_\nu^{sync}[i] = s^{sync}[i]\nu^{-\beta^{sync}[i]} \tag{1.3}$$

Le cube de données synchrotron, $\mathbf{X}^{sync}$ est donc entièrement déterminé par les champs[6] $\mathbf{s}^{sync}$ et β^{sync}. Ce dernier champ d'indice spectral, β^{sync} fluctue typiquement entre 2.5 et 3.1. Malheureusement, les émissions synchrotron ne peuvent être modélisées comme une composante de rang 1 : il n'intervient donc pas comme un mélange linéaire.

[6] Au sens d'une carte sur la sphère.

- **Emission Free-Free :** l'émission "Free-Free" provient de l'interaction d'électrons libres avec des ions dans le milieu galactique ionisé. Au cours de ces intéractions, l'énergie perdue par un électron engendre l'émission d'un photon. En première approximation, le cube de données "Free-Free" peut être modélisé comme une composante de rang 1 issue du produit d'un champ de température et d'une loi d'émission en loi de puissance d'indice spectral $\beta^{FF} \simeq 2,15$. De sorte qu'à la fréquence ν, la composante liée au "Free-Free" intervient comme suit :

$$\forall \nu; \quad \mathbf{x}_\nu^{FF} = \mathbf{s}^{FF} \nu^{-\beta^{FF}} \tag{1.4}$$

- **Emissions thermiques de poussières :** différents mécanismes peuvent être à l'origine des émissions de poussières galactiques. La description du rayonnement lié aux poussières est, à notre connaissance, uniquement effectuée à l'aide d'une loi empirique. La contribution liée aux émissions de poussières est modélisée pour chaque pixel i de la sphère comme suit :

$$\forall \nu; \quad x_\nu^P[i] \propto \frac{\nu^{\beta^P[i]}}{exp(h\nu/ks^P[i]) - 1} \tag{1.5}$$

  où β^P est le champ d'indices spectraux de la loi d'émission supposée varier aux alentours de 2-3. Le champ $\mathbf{s}^P$ est le champ de température dont les fluctuations sont comprises entre 16 et $18K$. Tout comme la composante synchrotron, la composante de poussière n'intervient pas sous la forme d'un mélange linéaire.

Cette liste de composantes est non exhaustive. Par exemple, certains astrophysiciens supposent l'existence d'une composante d'émissions de poussières dite "anormale" (non liée à des émissions thermiques).

Effet "Sunyaev Zel'Dovich"

L'effet Sunyaev Zel'Dovich (effet SZ) est la diffusion Compton inverse des photons du fond diffus cosmologique avec les électrons libres d'un milieu ionisé. Il y a plus précisément deux types d'effet SZ :

- **Effet SZ thermique :** l'effet SZ thermique est lié à la diffusion des photons dans un gaz d'électrons à haute température tels que ceux des amas de galaxies. En première approximation, les données de SZ thermique peuvent être modélisées comme une composante de rang 1 (mélange linéaire) :

$$\mathbf{X}^{sz} = \mathbf{a}^{sz}\mathbf{s}^{sz} \tag{1.6}$$

– **Effet SZ cinétique :** l'effet SZ cinétique est lié à la diffusion des photons dans un gaz d'électrons en mouvement par rapport au CMB. Le SZ cinétique suit une loi d'émission similaire à celle du CMB. De faible amplitude, il est de fait difficile à distinguer des fluctuations du fond diffus cosmologique. Étant fortement corrélé aux amas de galaxies, l'estimation du SZ cinétique peut s'effectuer en post-traitement sur la carte de CMB.

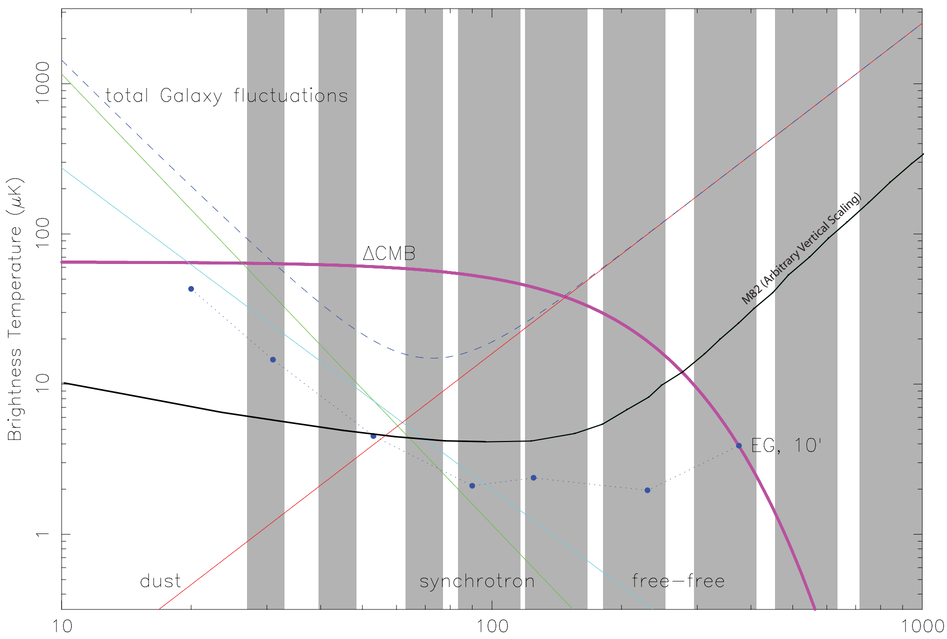

FIG. 1.3 – Puissance des différentes composantes du ciel en fonction de leur fréquence, superposée aux bandes de fréquences des récepteurs du satellite Planck.

La figure 1.3 montre la puissance des différentes composantes du ciel en fonction de leur fréquence.

1.3 Reconstruction du fond diffus cosmologique

Pour obtenir une carte du fond diffus cosmologique, de nombreux prétraitements sont nécessaires. Tout d'abord il s'agit de construire une carte de toute la sphère à partir des séries temporelles en tenant compte des problèmes liés au bruit basse fréquence et à l'impact des rayons cosmiques. Cette étape, appelée *map making* (Ashdown et al. 2007), ne sera pas abordée dans cette thèse.

Le *map making* nous impose de choisir une pixélisation sur la sphère comme par exemple la pixélisation Healpix (Gorski et al. 1999; Górski et al. 2005). Le chapitre suivant présente en détails différentes approches de pixélisation de la sphère.

Disposant d'une carte sur la sphère pour chacune des fréquences observées, la seconde étape est la séparation des composantes. En première approximation, les différentes composantes astrophysiques s'ajoutent linéairement au niveau des données :

$$\mathbf{X} = \mathbf{X}^{cmb} + \mathbf{X}^{SZ} + \mathbf{X}^{FF} + \mathbf{X}^{sync} + \mathbf{X}^{P} + \mathbf{X}^{PS} + \mathbf{N} \tag{1.7}$$

Le nombre total de pixels par canal est $t = 50331648$. Les 9 canaux Planck fournissent donc un jeu de données colossal contenant près de 450 millions de pixels. Les effets instrumentaux modélisés au niveau des données WG2/CH2 sont les suivants :

- **Bruit instrumental non-corrélé et non-stationnaire :** le bruit instrumental au niveau de la carte du ciel reconstruite à partir d'une année d'observation est lié à la manière dont le ciel est balayé par Planck. Plus précisément, pour le pixel k du canal i, la contribution $\mathbf{N}[i,k]$ a une variance inversement proportionnelle au nombre de fois où la zone du ciel correspondante a été balayée. En première approximation, le bruit est considéré gaussien, de moyenne nulle et tel que :

$$\forall i \text{ et } k \neq k'; \quad \mathbb{E}\left\{\mathbf{N}[i,k]\mathbf{N}[i,k']\right\} = 0 \tag{1.8}$$
$$\forall k \text{ et } i \neq i'; \quad \mathbb{E}\left\{\mathbf{N}[i,k]\mathbf{N}[i',k]\right\} = 0 \tag{1.9}$$
$$\forall k \text{ et } i; \quad \mathbb{E}\left\{\mathbf{N}[i,k]^2\right\} = \sigma^2_{\mathbf{N}i,k} \tag{1.10}$$

  où $\sigma^2_{\mathbf{N}i,k}$ est la variance effective du bruit au pixel k du canal i. Étant lié à l'instrument et au balayage du ciel par le satellite, cette variance effective est connue.

- **Résolution et convolution :** le ciel est observé par Planck dans différentes bandes de fréquence. L'ouverture de l'instrument étant fixe, chaque canal possède donc une résolution spatiale différente. L'ouverture circulaire de l'instrument D étant finie, il est équivalent de considérer que les données sont convoluées avec un noyau gaussien dont la largeur à mi-hauteur est la résolution de l'instrument.

Chaque observation effective $\mathbf{y}_i$ peut être modélisée somme suit :

$$\forall i = 1, \cdots, m; \quad \mathbf{y}_i = (\mathbf{h}_i \star \mathbf{x}_i) + \mathbf{n}_i \tag{1.11}$$

où $\mathbf{x}_i$ est l'observation au canal i idéale : non-convoluée et sans bruit. Le terme $\mathbf{h}_i$ correspond au noyau de convolution pour le canal i.

De nombreuses méthodes ont été proposées ces dernières années pour résoudre ce problème (Cardoso and Souloumiac 1996; Delabrouille et al.

2003; Moudden et al. 2005), notamment, la méthode GMCA (Bobin et al. 2007a, 2006b) qui utilise la transformée en ondelettes sur la sphère que nous présenterons dans la suite.

1.4 Analyse du fond diffus cosmologique

Pour analyser la carte de CMB reconstruite, on commence en général par masquer (c'est-à-dire mettre les pixels à la valeur zéro) les zones où la séparation n'a pas été efficace. Ces zones correspondent au plan galactique et aux sources ponctuelles. La séparation est relativement mauvaise sur le plan galactique car le modèle de linéarité du mélange est trop approximatif. Pour les sources ponctuelles, étant donné qu'elles n'ont pas le même spectre, il est impossible de les modéliser sous la forme d'une composante avec une loi d'émission. En pratique, la plupart des méthodes détectent d'abord les sources ponctuelles dans les différentes cartes, et masquent les zones correspondantes avant de lancer l'algorithme de séparation de composantes.

C'est donc à partir d'une carte de CMB masquée qu'il faut estimer le spectre de puissance du CMB et analyser les données. Une approche possible pour gérer ces données manquantes est d'utiliser l'algorithme d'*inpainting* (Abrial et al. 2007). Cette méthode que nous décrivons en détails au chapitre 6 est basée sur le principe de parcimonie de la solution, et permet de combler les zones masquées par des valeurs qui obéissent à certaines contraintes. La carte CMB *inpaintée* peut ensuite être utilisée comme s'il n'y a pas de données manquantes. Ceci permet donc de simplifier très fortement la manière d'analyser les données.

1.5 Les ondelettes et le CMB

Les ondelettes en astronomie

Les ondelettes sont très populaires en astronomie (Starck and Murtagh 2006) et ont conduit à des résultats impressionnants dans les applications de filtrage de bruit et de détection de sources. Par exemple, les deux centres de données Chandra et XMM utilisent des ondelettes pour détecter les sources étendues dans les images à rayon X. En cosmologie, les ondelettes ont été utilisées dans de nombreuses études telles que l'analyse de la distribution spatiale des galaxies (Slezak et al. 1993; Escalera and MacGillivray 1995; Starck et al. 2005; Martínez et al. 2005), l'estimation de la topologie de l'univers (Rocha et al. 2004), la détection de non-Gaussianité dans les cartes de CMB (Aghanim and Forni 1999; Barreiro et al. 2001; Vielva et al. 2004; Starck et al. 2004a), la reconstruction du spectre de puissance primordial (Mukherjee and Wang 2003), ou encore la reconstruction de cartes de matière noire à partir des mesures de cisaillement gravitationnel (Starck et al. 2006; Massey et al. 2007).

L'algorithme le plus utilisé est certainement la transformée en ondelettes dyadique et isotrope (aussi appelé algorithme à trous) (Starck and Murtagh 2006). L'isotropie de la fonction ondelette permet une détection très efficace des sources isotropes. Cette décomposition est non-décimée (le nombre de coefficients dans la décomposition est égal au nombre de pixels dans les données multiplié par le nombre d'échelles). Elle ne génère donc pas les artefacts liés à la décimation que l'on aurait avec une transformée en ondelettes (bi-)orthogonale dans des applications de restauration. Une autre propriété intéressante de cet algorithme est que la reconstruction est obtenue de manière extrêmement simple, par une simple addition de toutes les échelles et de la carte à basse résolution.

Ridgelet et Curvelet

Quand les données contiennent des structures anisotropes, les ondelettes ne sont plus optimales pour les analyser, et ceci a motivé la construction de nouvelles transformées comme la transformée en ridgelets et la transformée en curvelets (Donoho and Duncan 2000; Starck et al. 2002). Il a été montré (Starck et al. 2004a; Jin et al. 2005) que le transformée en curvelets pouvait être un outil extrêmement intéressant pour la détection et la discrimination de non-Gaussianités dans le CMB. Pour que ces nouvelles constructions géométriques multi-échelles soient utilisables dans le cadre de PLANCK, il est évidemment nécessaire de les développer sur la sphère.

Transformées multi-échelles et dictionnaire d'analyse

Plus généralement, chaque transformée est caractérisée par un dictionnaire donné, $\Phi = \{\phi_1, ..., \phi_K\}$, et la transformation consiste à représenter les données y comme une combinaison linéaire : $y = \sum_k \phi_k \alpha_k$, où α sont les coefficients. Dans le cas d'une transformée en ondelettes, le dictionnaire est composé de fonctions ondelettes, et α sont les coefficients d'ondelettes. En fonction du contenu des données, un dictionnaire est plus ou moins bien adapté pour la détection du signal d'intérêt. Un dictionnaire donné est considéré comme un bon choix pour une classe de signaux, si la transformation d'un signal appartenant à cette classe conduit à une représentation parcimonieuse, c'est à dire que la plupart des coefficients sont nuls ou proches de zéro tandis que peu de coefficients ont une grande amplitude. Si la signature morphologique des structures à détecter n'est pas connue, il est préférable de ne pas se restreindre à une analyse en ondelettes ou en curvelets, mais au contraire d'exploiter tous les outils disponibles (Starck et al. 2004a).

Transformées multi-échelles sur la sphère

Plusieurs transformées en ondelettes sur la sphère ont été proposées ces dernières années. Schröder and Sweldens (1995) ont développé une transformée en ondelettes orthogonale basée sur l'ondelette de Haar. Son intérêt est toutefois très limité à cause des propriétés de l'ondelette de Haar (manque de régularité), et des problèmes relatifs à une décomposition orthogonale pour toutes les applications de restauration de données. De nombreux papiers ont décrit de nouvelles transformées en ondelettes continues sur la sphère (Antoine 1999; Tenorio et al. 1999; Cayón et al. 2001a; Holschneider 1996; Wiaux et al. 2006a) et la détection de non-Gaussianités dans le CMB a été obtenue par Vielva et al. (2004, 2006) en utilisant une transformée continue sur la sphère avec une ondelette *chapeau mexicain* (Cayón et al. 2001a). Ces travaux ont été étendus aux ondelettes directionnelles sur la sphère (Antoine et al. 2002; McEwen et al. 2005; Wiaux et al. 2005, 2006a, 2008). Toutes ces transformées sont intéressantes pour analyser les données, mais ne peuvent pas être utilisées pour des applications de restauration, car elles n'ont pas d'algorithme de reconstruction.

Notre objectif

L'objectif de cette thèse est de développer une panoplie d'outils multi-échelles sur la sphère, afin que les scientifiques du projet PLANCK puissent disposer de la même qualité de méthodes d'extraction d'information que ce nous avons actuellement pour le dépouillement des images astronomiques. Après un bref survol dans les deux prochains chapitres de la pixélisation de la sphère et de l'état-de-l'art en ce qui concerne les ondelettes sphériques, nous présentons des nouvelles transformées (ondelettes isotropes sur la sphère, ridgelets, curvelets, etc) qui sont toutes inversibles. Cette propriété d'inversibilité nous permet alors d'utiliser ces outils non seulement pour des applications de détection de non-Gaussianités, comme avec les ondelettes continues, mais aussi pour d'autres applications comme le débruitage, la déconvolution, ou encore l'inpainting (Abrial et al. 2007; Abrial et al. 2008).

CHAPITRE 2
Pixélisation

Sommaire

2.1 Introduction 17
2.2 ECP : Equidistant Coordinate Partition 19
2.3 Icosaedre 19
2.4 COBE Quad Cube 21
2.5 IGLOO 21
2.6 HEALPix 22
2.7 HTM - Hierarchical Triangular Mesh 27
2.8 Gauss-Legendre Sky Pixelisation (GLESP) 27
2.9 Conclusions 30

2.1 Introduction

Il existe un certain nombre de pixélisations sur la sphère, elles ont toutes leurs avantages et leurs inconvénients. En règle générale, on est intéressé par un certain nombre de caractéristiques lorsque l'on choisit une pixélisation. On peut citer les points suivants :

– choix du nombre de pixels et de leur taille ;
– compacité des pixels pour un nombre de pixel donné ;
– simplicité des transformations en harmoniques sphériques Y_{lm} ;
– égalité de surface des pixels ;
– polyèdre régulier ;
– séparabilité des variables en latittude et en longitude $(\theta,\ \varphi)$;
– hiérarchie ;
– disponibilité de librairies, si possible performantes, robustes et éventuellement parallélisables.

Un certain nombre de ces points sont liés. D'autres, par contre, sont contradictoires. Par conséquent, aucune pixélisation ne peut satisfaire tous ces points. Un certain nombre de ces propriétés peut s'exprimer simplement en termes de contraintes sur la pixélisation. La compacité se traduit par la volonté d'obtenir des formes de pixels les plus proches possible du cercle. Cela peut aussi s'exprimer par une minimisation du diamètre des pixels c'est-à-dire $\max(d(p_1, p_2))$, où p_1 et p_2 sont 2 points dans le même pixel. D'autres contraintes découlent des propriétés que l'on souhaite que la transformée en harmoniques sphériques vérifie. Définissons rapidement la transformée en harmoniques sphériques. Toute fonction sur la sphère S^2 peut se décomposer sur la base des harmoniques sphériques :

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{-l}^{l} a_{lm} Y_{lm}(\theta, \phi) \tag{2.1}$$

où les Y_{lm} sont les harmoniques sphériques définis par :

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_{lm}(\cos\phi) e^{im\theta}, \tag{2.2}$$

Les P_{lm} sont les polynômes de Legendre associés. Ceux-ci sont définis par l'équation différentielle suivante :

$$(1-x^2)\frac{d^2}{dx^2}P_{lm} - 2x\frac{d}{dx}P_{lm} + \Big(l(l+1) - \frac{m^2}{1-x^2}\Big)P_{lm} = 0 \tag{2.3}$$

Ces polynômes associés sont reliés aux polynômes de Legendre classiques par :

$$P_{lm} = (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x) \tag{2.4}$$

où les P_l sont définis par :

$$P_l = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2-1)^l \tag{2.5}$$

De plus, les harmoniques sphériques ont comme propriété d'être orthogonaux deux à deux, ce qui s'écrit de la manière suivante :

$$\sum_{l \in \mathbb{N}} \sum_{|m| \leqslant l} Y^*_{lm}(\omega')\ Y_{lm}(\omega) = \delta(\omega' - \omega) \tag{2.6}$$

Ils forment également une base complète :

$$\int_{S^2} d\Omega\ Y^*_{lm}(\omega)\ Y_{l'm'}(\omega) = \delta_{ll'}\ \delta_{mm'} \tag{2.7}$$

$$\text{avec } d\Omega(\theta, \varphi) = sin\theta \; d\theta \; d\varphi$$

On peut ainsi souhaiter que la transformée en harmoniques sphériques possède un algorithme rapide. On peut également vouloir qu'elle soit inversible pour des images à bande limitée. Pour le premier point, il est indispensable que les points soient équirépartis sur les méridiens et les parallèles. Pour le deuxième point, il suffit que l'échantillonage en φ soit suffisant et que les points d'échantillonage en latittude θ appartiennent, soit aux zéros du polynôme de Legendre ou qu'ils soient équirépartis.

Après ce rapide rappel sur l'importance du choix de la pixélisation, nous présentons dans ce chapitre les principales méthodes de pixélisation qui ont été proposées dans la littérature.

2.2 ECP : Equidistant Coordinate Partition

Cette pixélisation est la plus ancienne et la plus naturelle. Elle est très utilisée en géophysique ainsi qu'en météorologie. Elle consiste en un découpage de la sphère par des parallèles et des méridiens régulièrement espacés (voir figure 2.1) permettant une bonne séparabilité en θ, φ (voir figure 2.9). Cette pixélisation s'appelle aussi $\Delta\theta\Delta\varphi$. L'énorme avantage de cette pixélisation est sa simplicité, tous les calculs que l'on peut effectuer avec sont extrêment simples. Cependant, cette simplicité est mise à mal par le fait que d'une part les pixels près des pôles sont extrêmement déformés et d'autre part les pixels n'ont pas une surface égale. Cela implique de plus un sur-échantillonnage non nécessaire près des pôles. Tous ces problèmes font que cette pixélisation est très peu utilisée pour l'étude du CMB.

2.3 Icosaedre

Tegmark a proposé une nouvelle pixélisation de la sphère (Tegmark 1996) très régulière. Pour cela, il utilise une projection radiale de l'icosaèdre (voir figure 2.2). Et il modifie légèrement la position des pixels afin de pouvoir imposer une surface égale des pixels.

Cette pixélisation est assez ambitieuse puisqu'elle permet d'améliorer la compacité des pixels. En effet, pour une pixélisation avec des pixels carrés, le diamètre des pixels est : $d_{cube} \approx \sqrt{(2 * \pi)/4}$, tandis qu'avec cette pixélisation, nous avons pour diamètre $d_{icosa} \approx \sqrt{(8 * \pi)/(3 * \sqrt{3} * N)}$. Cette pixélisation a pourtant deux défauts majeurs. Le premier est l'absence de séparation en θ, φ (voir figure 2.9). Ceci implique obligatoirement une très grosse lourdeur d'un point de vue numérique. En effet, les algorithmes deviennent inapplicables. Le deuxième point est l'absence de structure hiérarchique. De ce fait, il est assez difficile de passer d'une résolution à une autre.

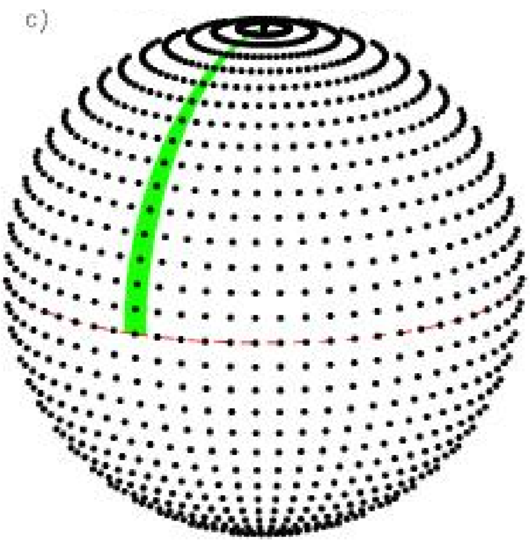

FIG. 2.1 – Une pixélisation ecp de la sphère.

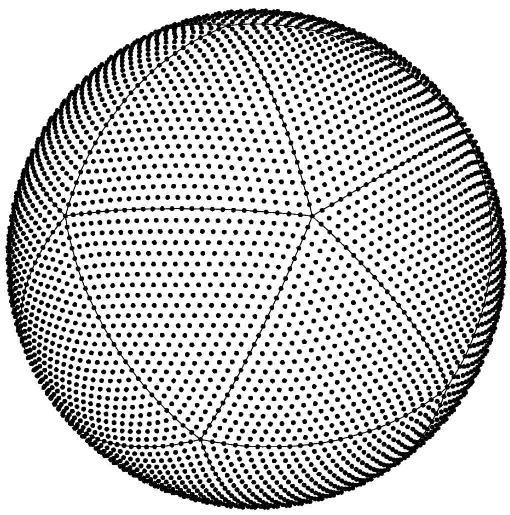

FIG. 2.2 – La pixelisation icosahèdre de Tegmark.

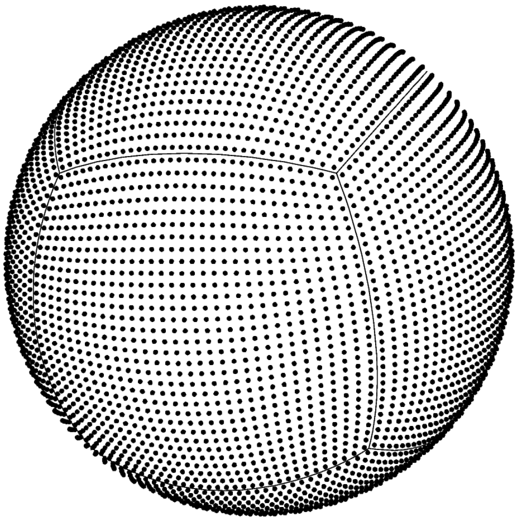

FIG. 2.3 – La pixélisation COBE Quad Cube.

2.4 COBE Quad Cube

Utilisée par COBE, cette pixélisation correspond à la projection radiale d'un cube dans la sphère (voir figure 2.3). Chaque face du cube est ensuite redécoupée de façon hiérarchique. Historiquement, cette pixélisation fut utilisée pour les données du satellite COBE. Sa strucuture simple se prêtait bien à l'analyse du petit nombre de données fournies par le satellte COBE (environ 6000 pixels). Cependant, cette pixélisation n'est pas parfaite, car elle n'est pas à symétrie azimuthale, si bien que le calcul des coefficients en harmoniques sphériques est approché et nécessite de nombreux ajustements. De plus, elle a une mauvaise séparabilité en θ, φ (voir figure 2.9). Cette pixélisation n'a plus été utilisée par la suite.

2.5 IGLOO

Suite aux nombreux problèmes introduits par la pixélisation quad cube de COBE, Critenden et Turok ont proposé plusieurs pixélisations IGLOO (Crittenden 2000). Leur idée est simple : compte tenu que la pixélisation quad cube présente un problème de symétrie azimuthale, ils mettent au point trois pixélisations entièrement à symétrie azimuthale. Pour cela, ils utilisent trois schémas de pixélisation différents.
Tout d'abord, ils proposent de couper la sphère en 12 pixels, 6 pixels triangulaires aux pôles, et 6 le long de l'équateur.(voir figure 2.4). Les pixélisations plus fines sont obtenues en divisant chaque pixel de base en quatre. Les pixels équatoriaux sont divisés de manière classique sous forme de grille. Les

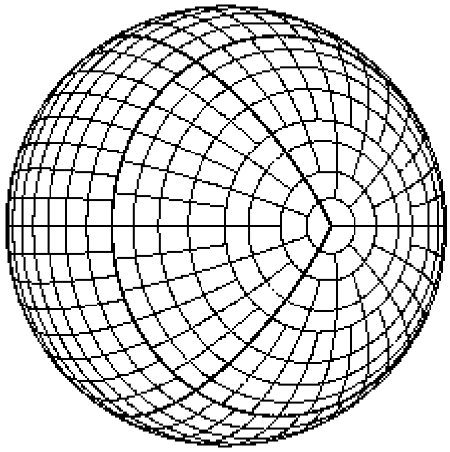

FIG. 2.4 – La pixélisation IGLOO 3 :6 :3.

pixels près des pôles sont aussi divisés en quatre mais de manière différente. Ils sont d'abord coupés en 1/4 - 3/4 par un parallèle, puis le sous-pixel de surface 3/4 est lui-même coupé en 3 suivant des méridiens. Il existe deux variantes à cette pixélisation, une privilégiant des pixels de même surface et une privilégiant une répartition égale en latitude.

Une troisième pixélisation de 12116 pixels à symétrie azimuthale a aussi été présentée dont on peut visualiser un pôle figure 2.5. La sous-division en éléments plus fins se fait de la même façon que précédemment.

Cette pixélisation a plusieurs avantages. Tout d'abord elle a une bonne séparabilité en θ, φ (voir figure 2.9) permettant un calcul rapide des coefficients en harmoniques sphériques. D'autre part, elle est hiérarchique, ce qui permet de passer facilement d'une pixélisation à une autre. Le point négatif est le fait que la forme du pixel est très peu conservée entre les pixels près de l'équateur et ceux près des pôles, ceux-ci sont d'autant plus déformés qu'ils sont situés près des pôles. Une amélioration de la pixélisation 12166 aurait peut-être permis à cette pixélisation de percer.

2.6 HEALPix

La pixélisation HEALPix[1] (Gorski et al. 1999; Górski et al. 2005) a été proposée afin de répondre au besoin en termes de pixélisation des nouvelles sondes WMAP et PLANCK. En effet, ces sondes fournissent des cartes de plusieurs millions de pixels sur la sphère et il s'agit de pouvoir traiter correctement ces données. Les différentes pixélisations précédentes, comme COBE

[1] http ://healpix.jpl.nasa.gov/

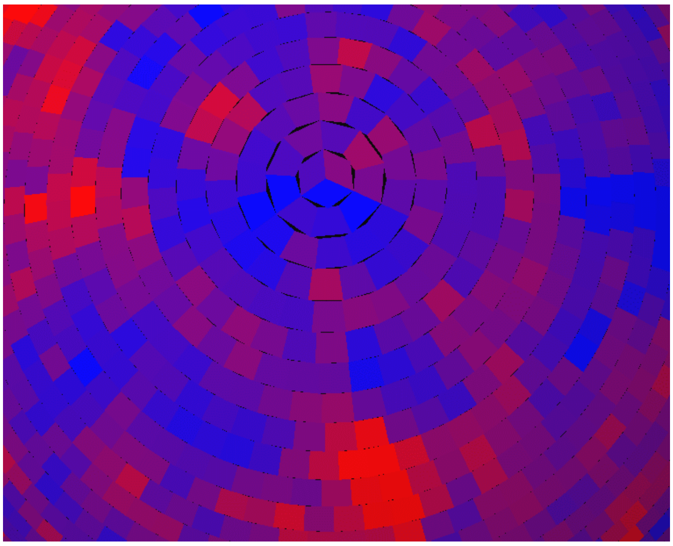

Fig. 2.5 – Visualisation des pôles de la pixélisation 12116 pixels.

quad cube, icosaèdre, ne peuvent convenir, notamment à cause de l'absence d'iso-latitude des pixels engendrant une explosion des temps de calcul des coefficients en harmoniques sphériques. En mettant au point leur pixélisation, Górski *et. al.* se sont concentrés sur trois points :

- l'iso-latitude des pixels ;
- la hiérarchie de la structure ;
- des pixels de même surface.

Pour cela, ils ont considéré la projection cylindrique de la sphère sur le plan. Cette projection permet une conservation de l'aire des surfaces de la sphère vers le plan. Ceci permet de dessiner simplement une pixélisation sur le plan sans se soucier de la projection.

Leur pixélisation s'effectue ainsi de cette façon. Ils découpent la surface formée suivant la projection cylindrique en N_θ et N_φ parties suivant respectivement la latitude et la longitude, selon les figures 2.6 et 2.7. La séparation entre les pixels près du pôle et ceux près de l'équateur est choisie de telle façon que tous les pixels aient la même surface. Du fait de ce choix de pixélisation de base, les frontières de chaque pixel sont assez simples. Elles sont de la forme $cos\theta = a + b * \varphi$ dans la zone équatoriale et de la forme $cos\theta = a + b/(\varphi)^2$ ailleurs.

Pour leur implantation, ils ont choisi $N_\theta = 3$ et $N_\varphi = 4$, si bien que la pixélisation de base comporte 4 pixels autour de chaque pôle et 4 pixels le long de l'équateur.

L'obtention de nouvelles pixélisations à partir de cette pixélisation de base est assez simple. Elle consiste simplement à diviser chaque pixel de base en quatre, comme le montre la figure 2.8 ainsi que la projection cylindrique

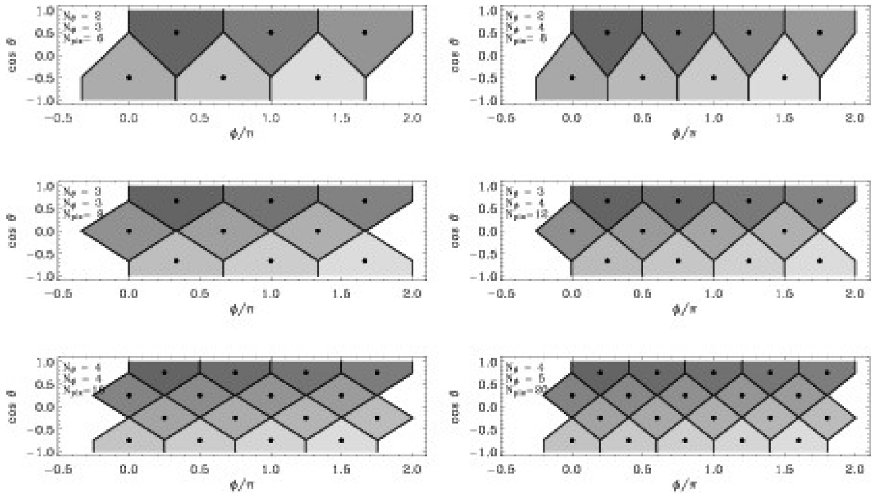

Fig. 2.6 – Projection cylindrique de différentes pixélisations possibles pour différents N_θ et N_φ. La grille HEALPix utilise $N_\theta = 3$ et $N_\varphi = 4$

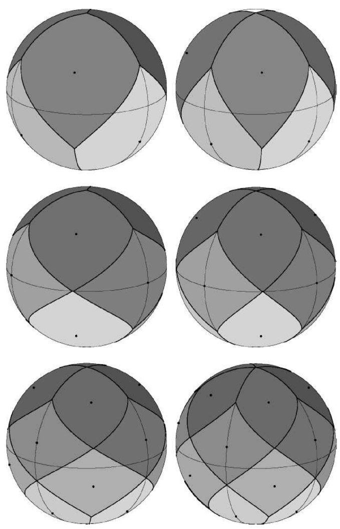

Fig. 2.7 – Projection orthographique de différentes pixélisations possibles pour différents N_θ et N_φ.

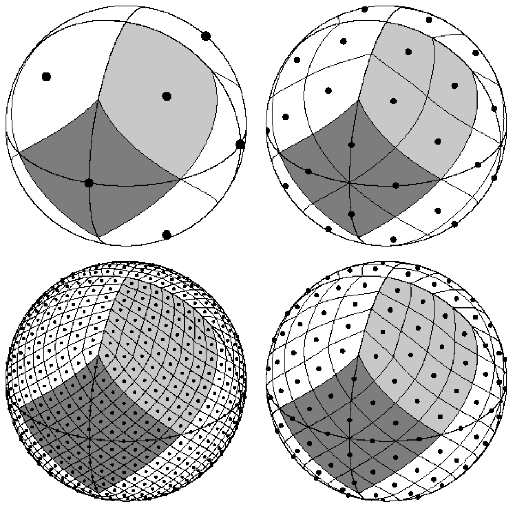

FIG. 2.8 – Différentes échelles de la pixélisation HEALPix. Chaque carte contient quatre fois de pixel que la précédente.

du découpage présentée figure 2.9. Ce type de découpage a deux avantages. D'une part, il permet une numérotation des pixels aisée. D'autre part cela garantit la conservation de l'égalité des aires des pixels. On peut noter de plus que l'ensemble des pixels se répartit sur un ensemble d'anneaux de latitude constante. Ce point est très important car il permet un calcul des coefficients en harmoniques sphériques rapide. A partir de ce découpage, deux schémas de numérotation des pixels peuvent être utilisés : la numérotation en anneaux, appelée 'ring scheme', et la numérotation en chaînes, appelée 'nested scheme'. Ces deux numérotations ont chacune leur propre utilité. La pixélisation en anneaux consiste à numéroter les pixels dans l'ordre dans lequel ils apparaissent le long des différents anneaux de la pixélisation. La numérotation en chaînes consiste à utiliser la hiérarchie de la pixélation HEALPix. Le numéro de chaque pixel est lié au numéro du pixel dont il est issu dans la pixélisation précédente (voir figure 2.10). Ainsi, grâce à cette numérotation, il devient très aisé de passer d'une pixélisation à une autre.

La pixélisation HEALPix a été choisie pour l'analyse des données de plusieurs satellites, notament celle du satellite WMAP. Cependant, elle pose encore quelques problèmes. Par exemple, le choix de la résolution est assez limité, puisque multiple de 2 de la résolution de base. Ce facteur est assez critique, puisque le nombre de données est multiplié par 4 lorsque l'on passe d'une résolution à la suivante.

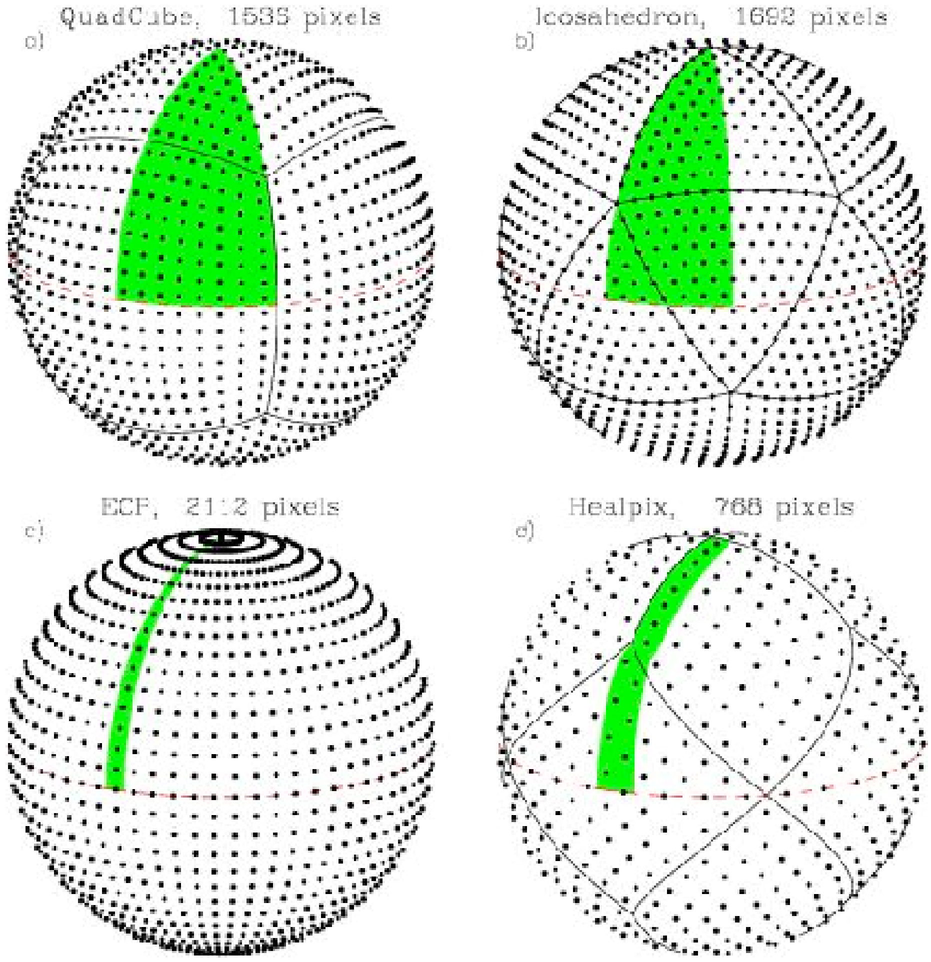

FIG. 2.9 – Comparaison de la séparation en θ, ϕ pour différentes pixélisations. .

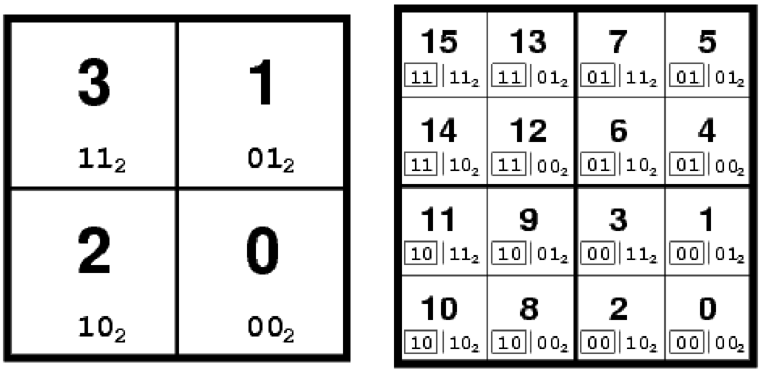

FIG. 2.10 – Passage d'une pixélisation à une autre par la numérotation en chaînes.

2.7 HTM - Hierarchical Triangular Mesh

La pixélisation HTM a pour but principal l'archivage des données sur la sphère, ainsi que des requêtes de régions rapides (Kunszt et al. 2001). Elle repose sur une division simple de la sphère en huit triangles le long de 4 méridiens et de l'équateur. Chaque pixel est ensuite divisé en quatre à chaque étape de la hiérarchie. Les premières résolutions sont montrées par la figure 2.11.

Le principal problème de cette pixélisation est l'absence de pixels de même taille, ainsi que des formes très différentes selon les pixels. En effet, les pixels très près du centre et les pixels de base seront presque équilatéraux. Par contre, ceux situés près des coins deviennent de plus en plus plats à chaque nouvelle itération et à très grande résolution, la déformation devient considérable. Cependant, la requête de régions partielles est très efficace. Ceci est lié à l'utilisation de triangles comme unité de base, plutôt que des quadrilatères comme dans la plupart des autres pixélisations. Cette propriété n'est toutefois pas la plus intéressante pour l'étude du CMB.

2.8 Gauss-Legendre Sky Pixelisation (GLESP)

En pratique, il y existe deux approches pour la pixelisation sur la sphère. Soit on cherche à obtenir un pavage uniforme et compact sur la sphère, qui permet de faire des sommes approchées, soit on cherche à obtenir des sommes exactes et une pixélisation qui permet de s'en approcher (Tegmark 1996). Les pixélisations IGLOO, icosaèdre, HEALPix et HTM permettent d'obtenir des pixélisations homogènes, plus ou moins compactes et effectuent leur somme de manière approchée. Elles se servent aussi de différentes techniques itératives afin d'améliorer la précision de leur résultat. L'approche GLESP (Doroshkevich et al. 2005) considère que le seul moyen de calculer les har-

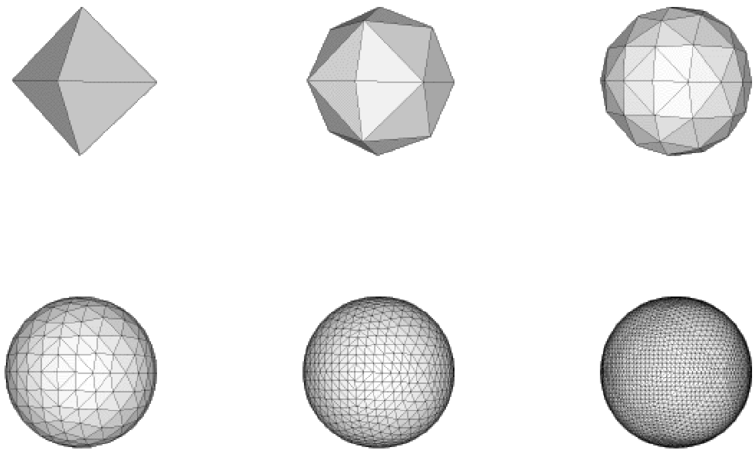

FIG. 2.11 – Différentes résolutions de la pixélisation HTM.

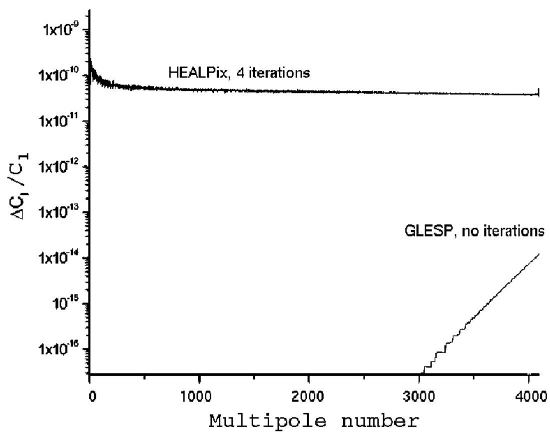

FIG. 2.12 – Comparaison de l'erreur de calcul des C_l avec HEALPix et GLESP.

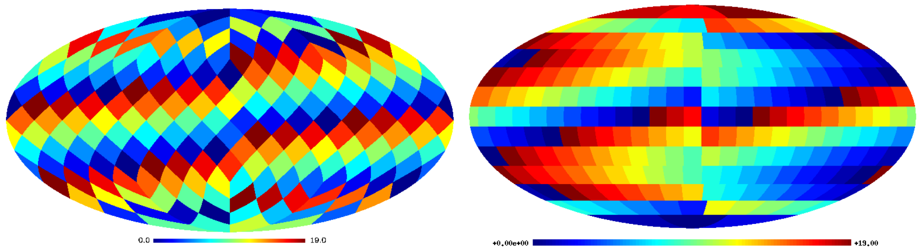

FIG. 2.13 – Comparaison des shémas de pixélisation pour une carte de très faible résolution (880 minutes d'arc) avec la pixélisation Healpix (à gauche) et la pixélisation Glesp (à droite).

moniques sphériques de manière exacte sur la sphère à un ordre donné est de prendre pour centre de pixel des points dont la lattitude se situe sur les zéros du polynôme de Legendre associé.

Pour ce qui est de la détermination en longitude, deux approches ont été proposées. La première consiste à avoir des pixels de surface plus ou moins égale. Pour cela, l'équateur est découpé en $2 * l + 1$ pixels, et les bandes de pixels inférieurs et supérieurs sont découpées de telle façon que les pixels aient une surface la plus proche possible des pixels de la bande équatoriale. L'autre approche a été de découper les pixels à longitude constante. Les pixels n'ont plus alors une suface égale, mais le calcul des harmoniques sphériques est encore plus précis.

Du fait que la quadrature en harmoniques sphériques soit exacte, le calcul des a_{lm} ainsi que l'estimation du spectre de puissance sont bien meilleurs. La figure 2.12 montre la différence entre les erreurs introduites par la pixelisation HEALPix et celles introduites par GLESP. Pour le calcul des a_{lm}, l'erreur moyenne est de l'ordre de $\sim 10^{-5}$ pour HEALPix, tandis qu'elle est de $\sim 10^{-7}$ en simple précision pour GLESP.

La figure 2.13 nous montre pour une carte à très faible résolution (880 minutes d'arc), la forme et la distribution des pixels pour les deux pixélisations Healpix et GLESP. On voit bien la différence de forme des pixels. Il y a 16 anneaux et 192 pixels pour la pixélisations Healpix et 13 anneaux et 220 pixels pour Glesp.

Pour les figures suivantes, il s'agit de cartes avec un nombre de pixels grand mais encore modéré, de résolution de 13.74 minutes d'arc et de 786432 pixels pour Healpix et de 785094 pour Glesp. Rappelons, que des cartes dites grandes ont plus de 150 millions de pixels et une résolution de l'ordre de la minute d'arc. Les deux cartes sont donc assez similaires en résolution.

La figure 2.14 nous montre le nombre de pixels par anneau pour les deux pixélisations. Pour la pixélisation Healpix, on remarque que le nombre de pixels par anneau est croissant en partant des pôles, puis devient constant dans la zone équatoriale correspondant à la moitier des anneaux. Notons

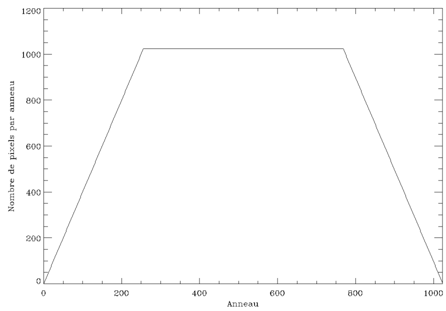

FIG. 2.14 – Comparaison du nombre de pixels par anneau avec la pixélisation Healpix (à gauche) et la pixélisation Glesp (à droite).

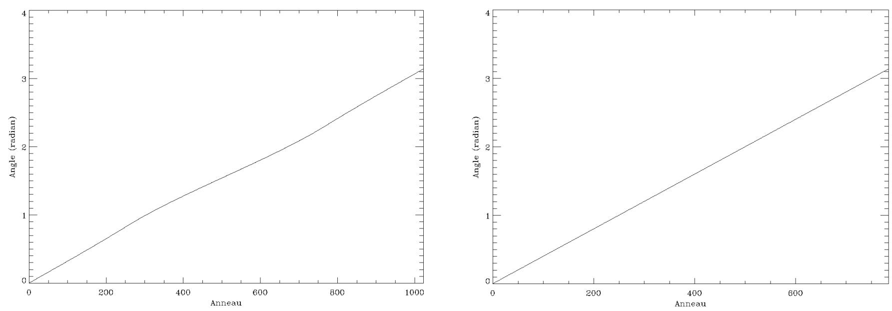

FIG. 2.15 – Comparaison de la position angulaire des anneaux avec la pixélisation Healpix (à gauche) et la pixélisation Glesp (à droite).

aussi que la croissance au niveau des pôles est linéaire (4 pixels de plus par anneau). Pour la pixélisation Glesp, la variation du nombre de pixels par annneau est beaucoup plus régulière.

La figure 2.15 présente la position angulaire des anneaux. Pour un meilleur échantillonnage et une résolution angulaire constante, la linéarité de cette courbe est importante. On remarque ici, que l'échantillonage des zéros du polynôme de Legendre est bien régulier pour Glesp, ce qui n'est pas la cas pour Healpix.

2.9 Conclusions

Parmi les pixélisations que nous venons de décrire, les pixélisations Healpix et Glesp sont certainement les plus intéressantes :

– Healpix présente l'avantage de sa grande diffusion dans la communauté. Les cartes WMAP notamment sont fournies dans ce format. De plus, un logiciel bien documenté existe, et de nouvelles versions sont régulièrement faites. D'autre part, les routines de calcul des harmoniques sphériques sont parallèlisées, et donc très rapides.

– Glesp a pour avantage de permettre une estimation exacte des coefficients en harmoniques sphériques. Cependant, le standard en CMB est le format HEALPix. Il est donc nécessaire pour les données WMAP ou Planck, de re-pixéliser les cartes de CMB pour pouvoir les utiliser avec GLESP. Dans GLESP deux re-pixélisations sont possibles : on peut faire une moyenne en utilisant une fonction de pondération ou bien faire une interpolation en utilisant une fonction spline cubique. Les effets de cette re-pixélisation en termes de conservation de flux et des fréquences ne sont pas bien compris.

Nous verrons dans la suite que la pixélisation Healpix va nous permettre de construire très simplement certaines transformées multi-échelles géométriques comme les curvelets sur la sphère. C'est donc cette pixélisation que nous avons essentiellement utilisées, même si la plupart de nos codes permettent de manipuler les deux types de pixélisation, Healpix et Glesp.

Chapitre 3
Transformée en ondelettes sur la sphère

Sommaire

3.1 Introduction **33**

3.2 Transformées en ondelettes sur la sphère existantes . **34**

3.2.1 Transformée en ondelettes de Haar sur la sphère . 34

3.2.2 Les ondelettes axisymétriques par projection stereoscopique inverse. 35

3.2.3 Ondelette directionelle sur la sphère. 39

3.2.4 Conclusions 43

3.3 Nouvelles transformées/reconstructions en ondelettes sur la sphère **43**

3.3.1 Transformée en ondelettes isotrope non décimées sur la sphère 43

3.3.2 Transformée en ondelettes isotrope pyramidale sur la sphère 49

3.4 Conclusion **52**

3.1 Introduction

Historiquement, la première transformée en ondelettes utilisée sur la sphère a été la transformée de Haar (Schröder and Sweldens 1995). Ellle est à la fois très facile à mettre en oeuvre et très rapide numériquement. Elle est directement dérivée de l'algorithme de la transformée de Haar sur la plan grâce aux pixelisations hiérarchiques, comme la pixelisation Quad-Cube ou Healpix. L'ondelette de Haar manque de régularité et sa forme n'est pas adaptée à une analyse fine du contenu des données. Pour cette raison, de nombreuses autres méthodes ont été proposées ces dernières années, comme la transformée en ondelettes continues isotropes (Antoine 1999; Tenorio et al. 1999; Cayón et al. 2001a; Holschneider 1996) et directionnelles (Wiaux et al. 2006a, 2008). Dans ce chapitre, nous présentons tout d'abord l'état de l'art dans ce domaine. Dans une deuxième partie, nous présentons une nouvelle transformée en ondelettes basée sur les harmoniques sphériques qui dispose d'un algorithme de reconstruction extrêmement simple.

3.2 Transformées en ondelettes sur la sphère existantes

3.2.1 Transformée en ondelettes de Haar sur la sphère

La transformation ondelette de Haar sur la sphère (SHW) est basée pour chaque résolution j et chaque pixel k de la sphère sur une fonction échelle $\phi_{j,k}$ et trois fonctions ondelettes $\psi_{m,j,k}$ avec $m \in \{1,2,3\}$. Pour la pixelisation Healpix, la résolution est donnée par le paramètre N_{side} qui correspond aux nombres de pixel sur le côté de chaque face, avec $N_{side} = 2^{j-1}$. A une résolution j on a donc $n_j = 12 \times 4^{j-1}$ pixels de surfaces μ_j. On peut ainsi définir une fonction d'échelle ϕ et trois fonctions ondelettes ψ avec :

$$
\begin{aligned}
\phi_{j,k}(x) &= \begin{cases} 1 & \text{si } x \in S_{j,k} \\ 0 & \text{sinon} \end{cases} \\
\psi_{1,j,k} &= \frac{\phi_{j+1,k_0} + \phi_{j+1,k_2} - \phi_{j+1,k_1} - \phi_{j+1,k_3}}{4\mu_{j+1}} \\
\psi_{2,j,k} &= \frac{\phi_{j+1,k_0} + \phi_{j+1,k_1} - \phi_{j+1,k_2} - \phi_{j+1,k_3}}{4\mu_{j+1}} \\
\psi_{3,j,k} &= \frac{\phi_{j+1,k_0} + \phi_{j+1,k_3} - \phi_{j+1,k_1} - \phi_{j+1,k_2}}{4\mu_{j+1}}
\end{aligned}
\tag{3.1}
$$

k_0, k_1, k_2, k_3 sont les quatre pixels à l'échelle $j+1$ issus du pixel k à l'échelle j. On calcule les coefficients d'échelle $\lambda_{j,k}$ à l'échelle j à partir de ceux de l'échelle $j+1$ par :

$$
\lambda_{j,k} = \frac{1}{4} \sum_{m=0}^{3} \lambda_{j+1,k_m}, \tag{3.2}
$$

ainsi que les coefficients ondelette de l'échelle j à partir des coefficients d'échelle $j+1$ avec :

$$
\begin{aligned}
\gamma_{1,j,k} &= \mu_{j+1}(\lambda_{j+1,k_0} + \lambda_{j+1,k_2} - \lambda_{j+1,k_1} - \lambda_{j+1,k_3}) \\
\gamma_{2,j,k} &= \mu_{j+1}(\lambda_{j+1,k_0} + \lambda_{j+1,k_1} - \lambda_{j+1,k_2} - \lambda_{j+1,k_3}) \\
\gamma_{3,j,k} &= \mu_{j+1}(\lambda_{j+1,k_0} + \lambda_{j+1,k_3} - \lambda_{j+1,k_1} - \lambda_{j+1,k_2})
\end{aligned}
\tag{3.3}
$$

La figure 3.1 schématise le calcul des coefficients ondelette et d'échelle d'une résolution à la suivante. Les coefficients échelle $\lambda_{j,k}$ des données WMAP sont présentés sur la figure 3.2.

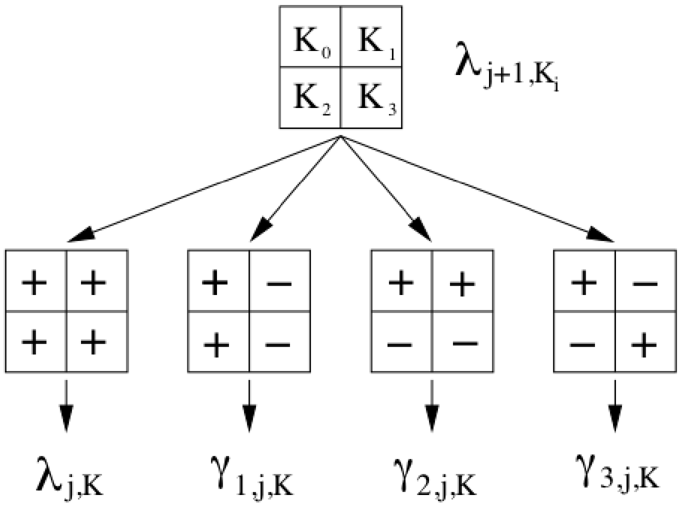

FIG. 3.1 – Passage d'une échelle à la suivante par une transformation de Haar.

La transformation de Haar est orthogonale, et sa reconstruction est exacte. On obtient la transformation inverse par :

$$f(x_i) = \sum_{l=0}^{n_{j_0-1}} \lambda_{j_0,l}\phi_{j_0,j}(x_i) + \sum_{j=j_0}^{J-1}\sum_{m=0}^{3}\sum_{l=0}^{n_j-1} \gamma_{m,j,l}\psi_{m,j,l}(x_i) \qquad (3.4)$$

Cette transformation, bien que rapide, est relativement limitée. En effet, elle n'est pas invariante pas rotation et dépend fortement du schéma de pixelisation, obligatoirement hiérarchique. La forme de l'ondelette de Haar n'est pas non plus adaptée pour les applications relatives à la cosmologie. A l'exception d'une étude sur la Gaussianité des données CMB de COBE (Barreiro et al. 2000), cette transformée n'a jamais été utilisée.

3.2.2 Les ondelettes axisymétriques par projection stereoscopique inverse.

La transformée en ondelettes continue offrre plus de flexibilité sur le choix de la fonction ondelettes. Afin de transposer simplement sur la sphère des fonctions définies sur le plan, il est possible d'utiliser la projection stéréographique inverse d'ondelettes radiales, comme par exemple l'ondelette chapeau mexicain (Cayón et al. 2001a). Cette transformation est très intéressante car elle permet de conserver toutes les propriétés des ondelettes du plan vers la sphère. On définit l'opérateur $\Pi^{-1} : x \to \omega = \Pi^{-1}x = (\theta(r), \phi)$, avec $\theta(r) = 2\arctan(r/2)$ et l'ondelettes radiale peut être déplacée sur la sphère par une rotation unique $\omega_0 = (\theta_0, \phi_0)$, respectivement autour

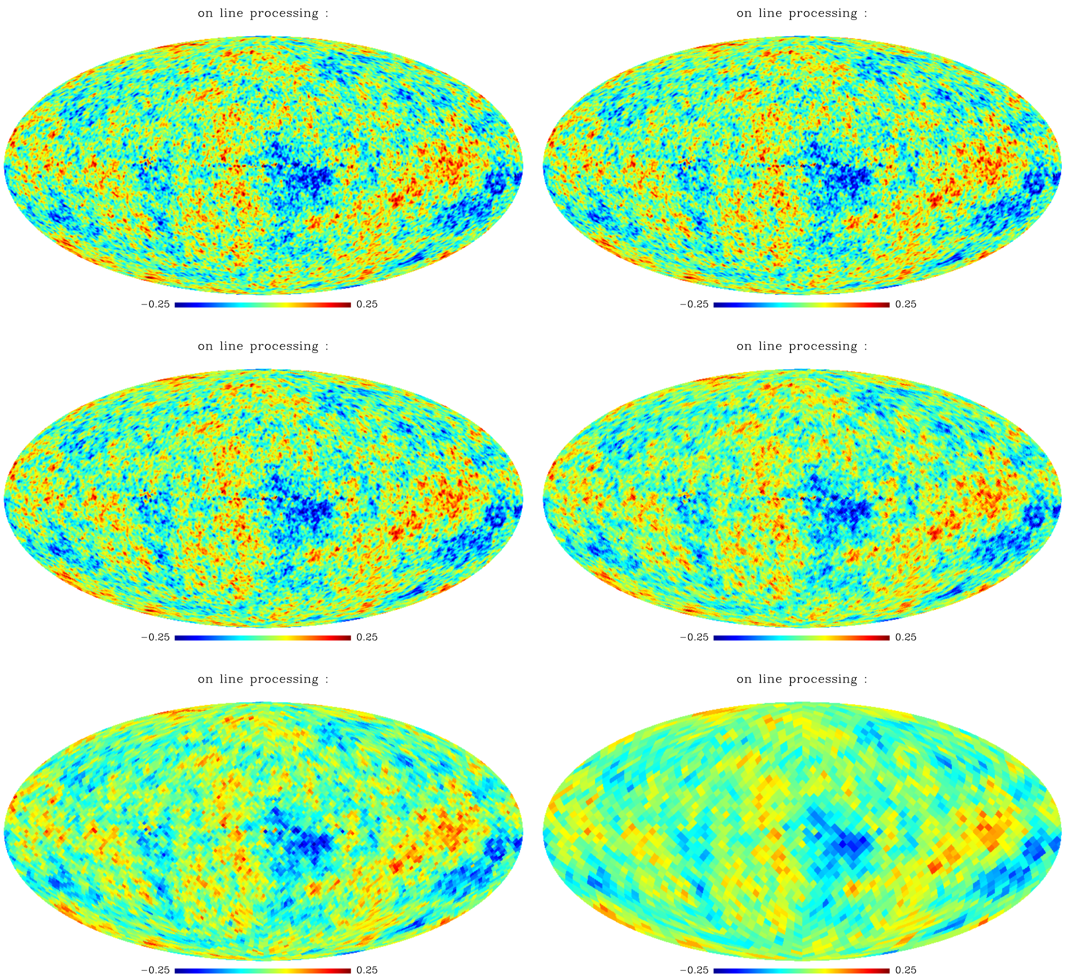

FIG. 3.2 – Coefficients d'échelle d'une transformation de Haar.

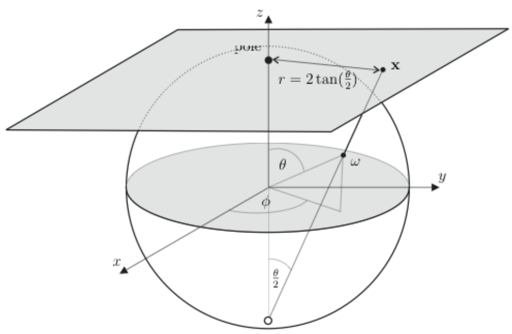

FIG. 3.3 – Projection stéréographique inverse d'une fonction radiale du plan sur la sphère.

des deux axes O_y et O_z. La figure 3.3 montre la projection de fonctions radiales du plan sur la sphère.

Ainsi il est possible de définir une convolution $*$ sur la sphère entre une fonction $f(\omega)$ et un filtre ou une ondelette radiale $\Psi(\theta)$. Cette convolution s'écrit de la façon suivante :

$$\phi(\theta) * f(\theta, \phi) = \int_{S^2} d\Omega \; \psi^*(R_\rho^{-1}\omega)F(\omega) \tag{3.5}$$

Ces ondelettes sont par leur construction axisymétriques. De ce fait, ceci a pour avantage une convolution avec une autre carte très aisée. En effet, les ondelettes étant axisymétriques, elles se décomposent uniquement sur les harmoniques sphériques eux même axisymétriques, c'est-à-dire les $Y_{l,0}$. Ainsi leurs coefficients en harmoniques sphériques $\hat{\Psi}_{lm}$ sont nuls pour $m \neq 0$. La convolution d'une fonction quelconque $f(\omega) = \sum_{l=0,m<|l|}^{L} \hat{a}_{lm} Y_{lm}$ par un filtre axisymétrique se fait donc dans le domaine des harmoniques sphériques suivant la formule suivante :

$$\phi(\theta) * f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \hat{a}_{lm} \hat{\varphi}_{l,0} Y_{lm}(\theta, \phi) \tag{3.6}$$

La dilatation classique des ondelettes du plan s'étend également sur la sphère par l'opérateur $D(a)$ qui, à une sphère associe une sphère dont les points subissent une dilatation D_a :

$$[D(a)G(\omega)] = \lambda^{1/2}(a, \theta)G(D_a^{-1}\omega) \tag{3.7}$$

où $\lambda^{1/2}(a, \theta)$ un facteur de conservation de la norme de S^2 :

$$\lambda^{1/2}(a, \theta) = a^{-}1[1 + \tan^2(\theta/2)]/[1 + a^{-2}\tan^2(\theta/2)] \tag{3.8}$$

La fonction de dilatation des points D_a est telle que :

$$D_a(\theta, \phi) = (\theta_a(\theta), \phi) \tag{3.9}$$

La fonction $\theta_a(\theta)$ est définie de $\theta \in [0, \pi[\to \theta_a \in [0, \pi[$ par la relation :

$$\tan(\theta_a(\theta)/2) = a \tan(\theta/2) \tag{3.10}$$

Le pôle sud est défini par la projection en lui-même.

L'ondelette du chapeau mexicain

L'ondelette radiale du chapeau mexicain sur le plan, s'exprime comme la dérivée seconde d'une gaussienne (Cayón et al. 2001a) :

$$\psi(r) = \frac{1}{\sqrt{2\pi}}\left(2 - \left(\frac{r}{R}\right)^2\right)e^{-\frac{r^2}{2R^2}} \tag{3.11}$$

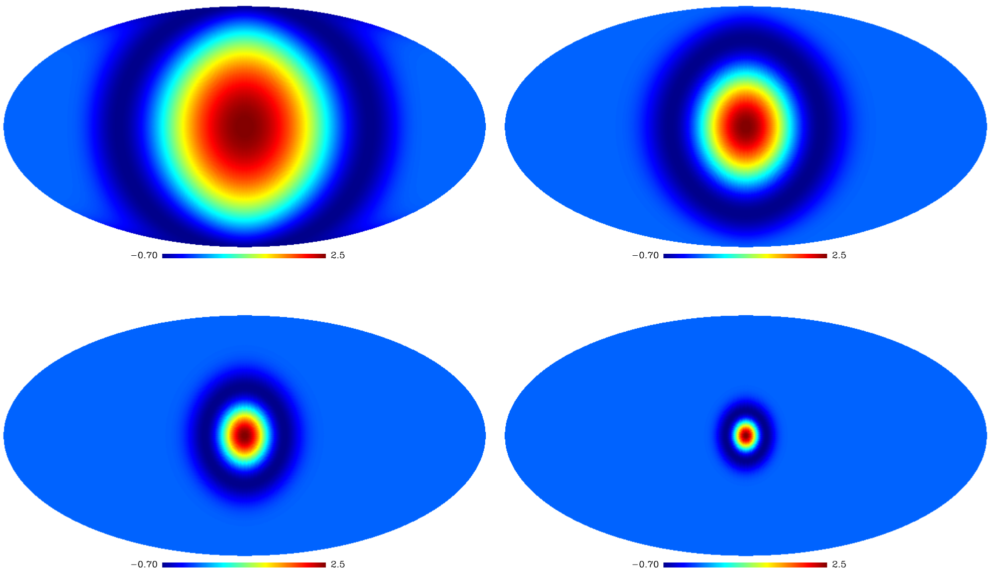

FIG. 3.4 – Ondelette chapeau mexicain pour des coefficients de dilatation $a = \{1, 2, 4, 8\}$.

où R est un facteur d'échelle et r une mesure de distance au centre de l'ondelette. Cette fonction a été utilisée pour le développement d'une transformation ondelette *continue* et a permis l'utilisation de celle-ci dans l'analyse d'images et la détection de motifs isotropes cachés dans du bruit.
En utilisant la projection stéréographique inverse, il est possible de définir une extension de l'ondelette chapeau mexicain sur la sphère, comme le montre Antoine (1999); Tenorio et al. (1999); Cayón et al. (2001a); Holschneider (1996). Cela se résume sous la forme suivante :

$$\psi_s(\delta) = \frac{1}{\sqrt{2\pi} N_R}\left(1 + \left(\frac{\delta}{2}\right)^2\right)^2 \left(2 - \left(\frac{\delta}{R}\right)^2\right) e^{-\frac{\delta^2}{2R^2}}, \qquad (3.12)$$

où R est un facteur d'échelle et N_R un facteur de normalisation :

$$N_R = R\left(1 + \frac{R^2}{2} + \frac{R^4}{4}\right)^{\frac{1}{2}}, \qquad (3.13)$$

et δ la distance entre le point tangent auquel est reliée l'angle polaire θ de la projection stéréographique inverse :

$$\delta = 2\tan\frac{\theta}{2} \qquad (3.14)$$

L'ondelette résultante sur la sphère est clairement zonale, ce qui permet de la transposer dans le domaine des harmoniques sphèriques afin de calculer

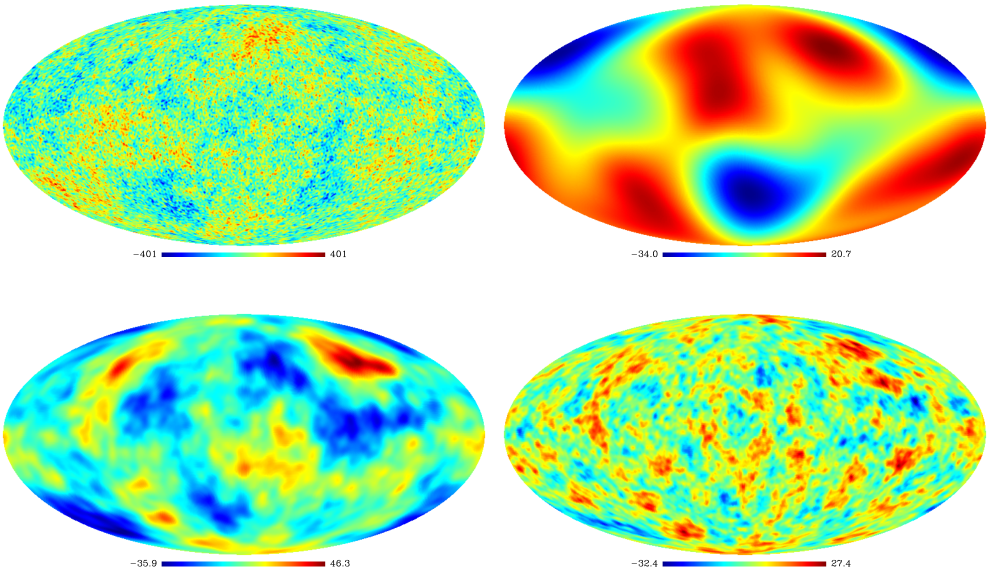

FIG. 3.5 – Simulation d'une carte de CMB et sa transformation ondelette chapeau mexicain pour des coefficients de dilatation $a = \{1, 8, 20\}$.

des coefficients en ondelettes. Cette transformation a été utilisée avec succès pour la détection de points sources dans le CMB et la détection de non gaussianité dans le CMB (Vielva et al. 2004). Cependant, bien que utile pour l'analyse de données, cette transformation ne permet pas la reconstruction. La figure 3.4 nous montre l'ondelette du chapeau mexicain à 4 échelles différentes et la figure 3.5 montre quatre échelles de la transformée en ondelettes d'une carte simulée de CMB.

3.2.3 Ondelette directionelle sur la sphère.

Pour étudier des structures anisotropes, la transformée continue décrite précédemment peut être étendue à des ondelettes directionnelles (Antoine et al. 2002; Vielva et al. 2006; McEwen et al. 2007). L'idée du développement de ces ondelettes est l'étude de fonctions non axisymmétriques sur la sphère. Bien que cette étude est habituelle sur le plan, son extension sur la sphère est bien plus difficile. La première difficulté est le fait que l'espace des projections de l'ensemble des rotations sur la sphère, qui serait l'équivalent sur le plan de l'ensemble des déplacements, n'est pas défini par deux angles, comme la sphère S^2, mais par les trois angles d'Euler $\rho = (\chi, \theta, \phi)$, avec $\chi, \phi \in [0, 2\pi[$ et $\theta \in [0, \pi[$. Ces rotations sont définies par une rotation autour de l'axe $\hat{z}$ d'un angle χ, suivi d'une rotation d'angle θ autour de $\hat{y}$ et enfin une dernière rotation d'angle ϕ autour de nouveau l'axe $\hat{z}$. L'espace $SO(3)$ convient bien pour la définition de ces angles.

Dans un premier temps, plusieurs ondelettes directionnelles vont être d'abord être présentées, dans le paragraphe suivant. Les différentes façons de calculer la corrélation entre ces ondelettes et une fonction de la sphère seront developpées dans la section suivante.

Les ondelettes directionelles par projection

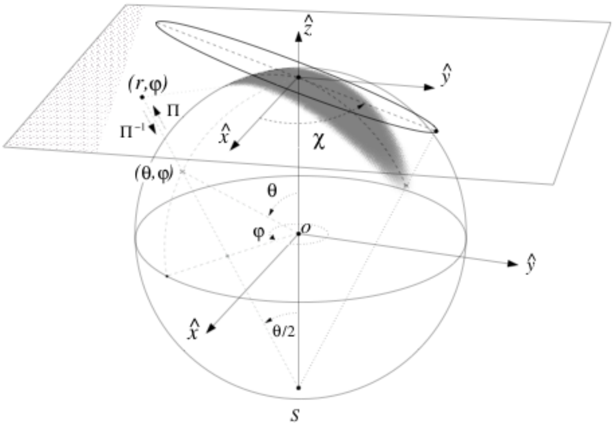

FIG. 3.6 – Projection stéréographique inverse d'une ondelette directionnelle sur la sphère.

La méthode de projection des ondelettes du plan sur la sphère est toujours la projection stéréographique inverse. Cependant, les ondelettes utilisées ne sont plus radiales, mais directionnelles. La figure 3.6 montre la projection d'une ondelette de forme elliptique du plan vers la sphère. Une ondelette directionnelle est donc caractérisée par son échelle a mais aussi son angle par rapport à l'axe $\hat{x}$: χ.

Ondelette chapeau mexicain étiré

Dans le cas radial, l'ondelette du chapeau mexicain utilise la correspondance $r = \sqrt{x^2 + y^2}$. Il est possible de priviligier une direction particulière, et l'équation finale de l'ondelette du chapeau mexican étiré, s'écrit de la façon suivante :

$$\begin{aligned}\Psi^{(mex)}(\omega) &= \sqrt{\frac{2}{\pi}} N(\sigma_x, \sigma_y)(1 + \tan^2 \frac{\theta}{2})[1 - \frac{4 \tan^2 \theta/2}{\sigma_x^2 + \sigma_y^2}(\frac{\sigma_y^2}{\sigma_x^2} \cos^2 \phi \\ &\quad + \frac{\sigma_x^2}{\sigma_y^2} \sin^2 \phi)] e^{-2 \tan \frac{\theta}{2} (\cos^2 \phi / \sigma_x^2 + \sin^2 \phi / \sigma_y^2)} \qquad (3.15)\end{aligned}$$

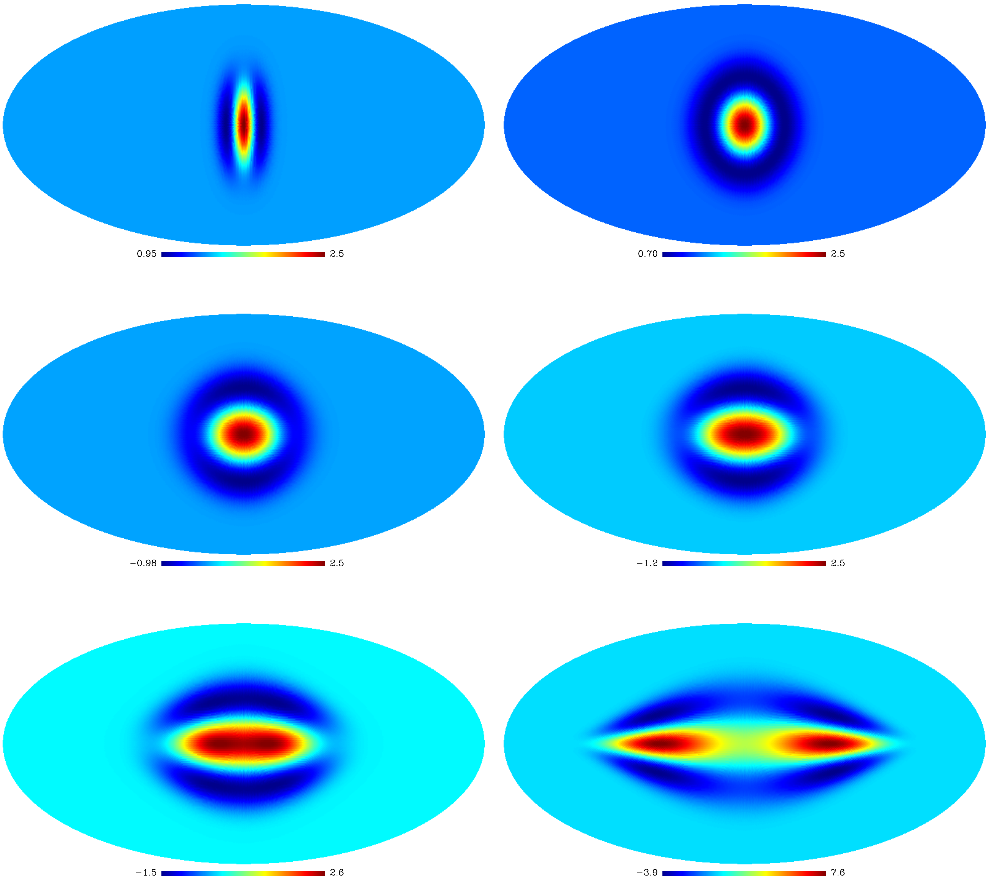

FIG. 3.7 – Ondelette chapeau mexicain pour un coefficient de dilatation $a = 4$, $\sigma_x = 1$ et $\sigma_y = \{0.5, 1., 1.25, 1.5, 2, 4\}$.

La constante $N(\sigma_x, \sigma_y)$ permet de normaliser l'ondelette.

$$N(\sigma_x, \sigma_y) = (\sigma_x^2 + \sigma_y^2)[\sigma_x \sigma_y (3\sigma_x^4 + 3\sigma_y^4 + 2\sigma_x \sigma_y)]^{-1/2} \tag{3.16}$$

La figure 3.7 nous montre pour un facteur de dilatation donnée, les différentes ondelettes que l'on peut obtenir avec différents paramètres de σ_x et σ_y.

Ondelette de Morlet sur la sphère

L'ondelette de Morlet est une ondelette directionnelle classique du plan. Elle permet de choisir assez finement le nombre d'oscillations de l'ondelette ainsi que leur direction. Sa projection stéreographique sur la sphère s'ex-

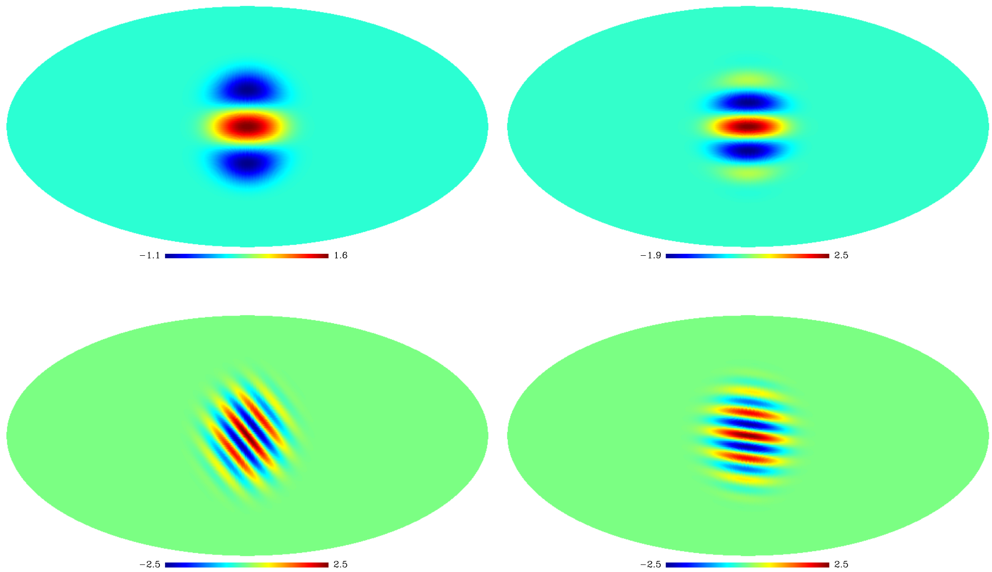

FIG. 3.8 – Ondelette de Morlet sur la sphère. Les vecteurs $\vec{k}$ ont respectivement pour valeur (2,0), (4,0), (6,6) et (9,1).

prime de la façon suivante :

$$\Psi^{(mor)}(\omega) = \sqrt{\frac{2}{\pi}}N(\sigma_x,\sigma_y)(1+\tan^2\frac{\theta}{2})[\cos(\frac{\vec{k}\cdot(\pi^{-1})\vec{x}}{\sqrt{2}} - e^{-\vec{k}^2/4}]e^{-2\tan^2(\theta/2)} \tag{3.17}$$

avec $(\pi^{-1})\vec{x}$ qui s'écrit :

$$(\pi^{-1})\vec{x} = (2\tan(\theta/2)\cos\phi, 2\tan(\theta/2)\sin\phi) \tag{3.18}$$

Le vecteur $\vec{k} = (k_x, k_y)$ permet le contrôle du nombre d'oscillations dans les directions $\hat{x}$ et $\hat{y}$. Il permet ainsi de choisir une forme et une direction général de l'ondelette. Cette ondelette est normalisée par $N(k) = (1+3e^{-\vec{k}^2/2} - 4e^{-3\vec{k}^2/8})^{-1/2}$. La figure 3.8 présente quelques formes particulières de cette ondelette, pour un vecteur $\vec{k}$ valant respectivement (2,0), (4,0), (6,6) et (9,1).

Ondelette directionnelle steerable

Un filtre Ψ sur la sphère est dit 'steerable' si n'importe quelle rotation d'angle χ sur l'axe O_z de ce filtre s'exprime comme combinaison linéaire d'un nombre fini M de filtres χ_m (Wiaux et al. 2006a; McEwen et al. 2007).

Ceci s'exprime par la relation :

$$[R^{\hat{z}}\Psi](\omega) = \sum_{m=1}^{M} k_m(\chi)\Psi_m(\omega) \tag{3.19}$$

Ces filtres ont une propriété très intéressante pour le calcul d'une corrélation entre un filtre et une fonction de la sphère. En effet, les harmoniques sphériques de tels filtres sont nuls pour certaines valeurs de m.

Les dérivées n-ième dans la direction $\hat{x}$ de fonctions radiales sur le plan, sont des ondelettes "steerable". Du fait que la projection stéreographique inverse soit radiale, la propriété de steerabilité est conservée lors de la projection. Ainsi la projection stéreographique de dérivée de Gaussiennes nous permet d'obtenir des filtre "steerable" sur la sphère. Nous allons maintenant présenter les formes explicites de ces dérivées pour les deux premiers ordres.

3.2.4 Conclusions

Les transformées en ondelettes continues isotropes et directionnelles ont été utilisées ces dernières années pour tester la Gaussianité du CMB (Vielva et al. 2004). Le fameux *Cold Spot* a été détecté avec le chapeau mexicain (Cruz et al. 2005). Les ondelettes directionnelles "steerable" sont relativement coûteuse en temps de calcul, mais ont tout de même été employées pour étudier l'isotropie du CMB (Vielva et al. 2006; Wiaux et al. 2006b).

3.3 Nouvelles transformées/reconstructions en ondelettes sur la sphère

A l'exception de la transformée de Haar, les transformées en ondelettes continues sont relativement coûteuse en temps de calcul et n'ont pas d'algorithme de reconstruction, ce qui limite très fortement le domaine d'application. Nous montrons dans cette partie qu'il est possible de construire une transformée en ondelettes basée sur les harmoniques sphériques qui présente un algorithme de reconstruction extrêmement simple. En effet cette transformée est très similaire à l'algorithme à trous car la reconstruction se fait par une simple sommation des échelles et du plan lissé. De plus, en utilisant la pixélisation Healpix, il devient aussi possible de décimer les échelles tout en gardant une reconstruction exacte (Starck et al. 2006).

3.3.1 Transformée en ondelettes isotrope non décimées sur la sphère

Notre transformation isotropique est obtenue en utilisant une fonction d'échelle isoptrope, $\phi_{l_c}(\vartheta, \varphi)$ avec une fréquence de coupure l_c. Ceci a pour effet que les coefficients en harmoniques sphériques de cette fonction d'échelle

$\hat{\phi}_{l_c}(l,m)$ de ϕ_{l_c} sont nulles quand $m \neq 0$, ce qui implique que :

$$\phi_{l_c}(\vartheta,\varphi) = \sum_{l=0}^{l=l_c} \hat{\phi}_{l_c}(l,0) Y_{l,0}(\vartheta,\varphi) \tag{3.20}$$

où les $Y_{l,m}$ sont les harmoniques sphériques.

Ainsi, convoluer une carte $f(\vartheta,\varphi)$ avec ϕ_{l_c} est grandement simplifié et les coefficients en harmoniques sphériques $\hat{c}_0(l,m)$ de la carte convoluée sont obtenus par (Bogdanova et al. 2005) :

$$\hat{c}_0(l,m) = \widehat{\phi_{l_c} * f}(l,m) = \sqrt{\frac{4\pi}{2l+1}} \hat{\phi}_{l_c}(l,0) \hat{f}(l,m) \tag{3.21}$$

où $*$ représente la convolution.

La transformation

D'une échelle à la suivante

Une suite d'approximations d'une fonction f à des échelles dyadiques est obtenue en utilisant les fonctions d'échelle suivantes :

$$\begin{aligned} c_0 &= \phi_{l_c} * f \\ c_1 &= \phi_{2^{-1}l_c} * f \\ &\dots \\ c_j &= \phi_{2^{-j}l_c} * f \end{aligned}$$

où $\phi_{2^{-j}l_c}$ est une version dilatée de ϕ_{l_c} avec une fréquence de coupure $2^{-j}l_c$: $\hat{\phi}_{2^{-j}l_c} = \frac{\hat{\phi}_{l_c}(l,/m)}{2^j}$

La séquence multirésolution précédente peut être obtenue de manière récursive. On définit un filtre passe bas h_j pour chaque échelle j par

$$\hat{H}_j(l,m) = \sqrt{\frac{4\pi}{2l+1}} \hat{h}_j(l,m) = \begin{cases} \frac{\hat{\phi}_{\frac{l_c}{2^{j+1}}}(l,m)}{\hat{\phi}_{\frac{l_c}{2^j}}(l,m)} & \text{si } l < \frac{l_c}{2^{j+1}} \quad \text{et} \quad m = 0 \\ 0 & \text{sinon} \end{cases} \tag{3.22}$$

On remarque aisément que c_{j+1} dérive de c_j par convolution avec h_j : $c_{j+1} = c_j * h_j$.

Les coefficients ondelette

À partir d'une fonction ondelette symétrique ψ_{l_c}, on peut également obtenir de la même façon un filtre passe haut g_j pour chaque échelle j :

$$\hat{G}_j(l,m) = \sqrt{\frac{4\pi}{2l+1}} \hat{g}_j(l,m) = \begin{cases} \frac{\hat{\psi}_{\frac{l_c}{2^{j+1}}}(l,m)}{\hat{\phi}_{\frac{l_c}{2^j}}(l,m)} & \text{si } l < \frac{l_c}{2^{j+1}} \quad \text{et} \quad m = 0 \\ 1 & \text{si } l \geq \frac{l_c}{2^{j+1}} \quad \text{et} \quad m = 0 \\ 0 & \text{sinon} \end{cases} \tag{3.23}$$

Ainsi, les coefficients ondelette w_{j+1} à l'échelle $j+1$ sont calculés à partir de la résolution précédente par une simple convolution : $w_{j+1} = c_j * g_j$.

Comme avec l'algorithme *à trous* , les coefficients ondelette peuvent être calculés comme la différence de deux résolutions successives, $w_{j+1}(\vartheta, \varphi) = c_j(\vartheta, \varphi) - c_{j+1}(\vartheta, \varphi)$, lesquelles correspondent en fait au choix pour la fonction ondelette ψ_{l_c} :

$$\hat{\psi}_{\frac{l_c}{2^j}}(l,m) = \hat{\phi}_{\frac{l_c}{2^{j-1}}}(l,m) - \hat{\phi}_{\frac{l_c}{2^j}}(l,m) \tag{3.24}$$

Les filtres passe haut g_j définis au dessus sont, dans ce cas particulier, exprimés de la façon suivante :

$$\hat{G}_j(l,m) = \sqrt{\frac{4\pi}{2l+1}}\hat{g}_j(l,m) = 1 - \sqrt{\frac{4\pi}{2l+1}}\hat{h}_j(l,m) = 1 - \hat{H}_j(l,m) \tag{3.25}$$

Cependant, d'autres fonctions ondelettes pourraient aussi être choisies.

Le choix de la fonction d'échelle

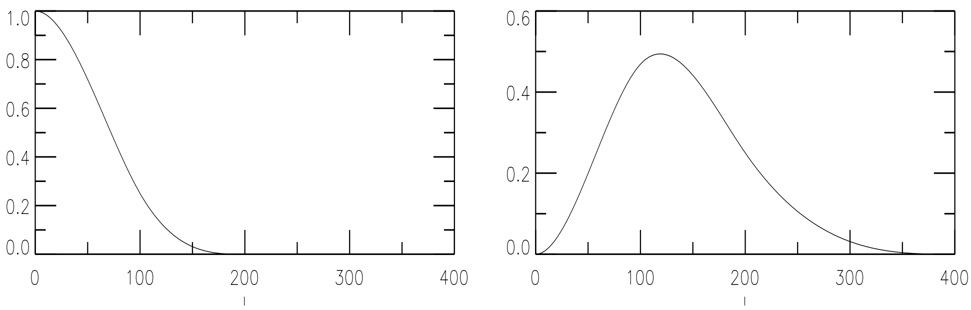

FIG. 3.9 – À gauche, la fonction d'échelle $\hat{\phi}$ et à droite la fonction ondelette $\hat{\psi}$.

N'importe quelle fonction avec une fréquence de coupure est utilisable. Nous avons retenu une fonction B-spline d'ordre 3. Elle est similaire à une gaussienne mais converge assez rapidement vers 0 :

$$\hat{\phi}_{l_c}(l, m=0) = \frac{3}{2}B_3(\frac{2l}{l_c}) \tag{3.26}$$

où $B(x) = \frac{1}{12}(\mid x-2\mid^3 - 4\mid x-1\mid^3 + 6\mid x\mid^3 - 4\mid x+1\mid^3 + \mid x+2\mid^3)$.

Sur la Fig. 3.9 la fonction d'échelle choisie est dérivée d'une B-spline de degré 3. La fonction ondelette résultante est affichée dans l'espace fréquentielle. La Fig. 3.10 montre en haut à gauche la carte WMAP du CMB, et sa transformée en ondelettes avec quatre échelles ondelettes et le plan lissé.

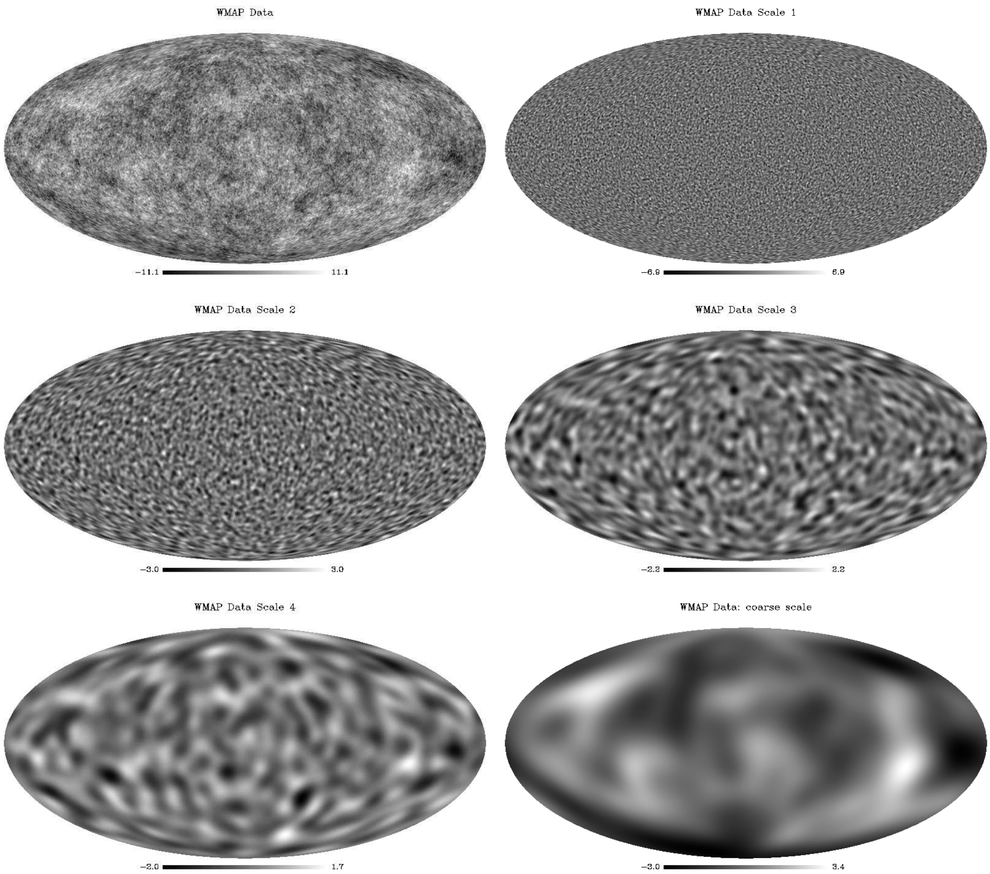

FIG. 3.10 – Carte du CMB obtenue par WMAP (en haut à gauche) et sa transformée ondelette sur la sphère sur 5 échelles (4 échelles ondelettes et le plan lissé). La somme de ces 5 cartes reproduit exactement les données d'origine. Dans l'ordre, de gauche à droite et de haut en bas : données originales, 4 échelles ondelette, 1 plan lisse.

1. Calculer la fonction d'échelle B_3-spline puis en deduire numériquement ψ, h et g.
2. Calculer les coefficients $\hat{c}_0$. en harmoniques sphériques de l'image c_0.
3. Initialiser l'échelle j à 0. Itérer :
   4. Effectuer le produit de $\hat{c}_j$ par $\widehat{H}_j$. On obtient $\hat{c}_{j+1}$.
   À partir de la transformation en harmonique sphérique inverse, on en déduit l'image a l'échelle $j+1$.
   5. Effectuer le produit de $\hat{c}_j$ by $\widehat{G}_j$. On obtient $\hat{w}_{j+1}$.
   De la même façon, la transformation inverse en harmonique sphérique de $\hat{w}_{j+1}$ nous permet d'obtenir les coefficients ondelette w_{j+1} à l'échelle $j+1$.
   6. j=j+1 et si $j \leq J$, retourner à l'étape 4 .
7. L'ensemble $\{w_1, w_2, \dots, w_J, c_J\}$ est la transformation en ondelette de c_0.

Algorithme de la transformée en ondelettes non décimée sur la sphère.

La reconstruction

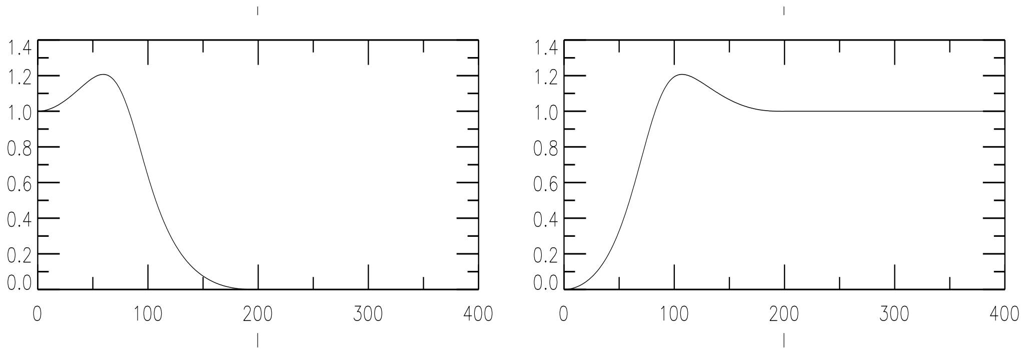

FIG. 3.11 – Sur la gauche, le filtre passe haut associé $\hat{\tilde{h}}$, sur la droite le filtre passe bas associé $\hat{\tilde{g}}$.

Si l'ondelette est la différence entre deux résolutions, l'étape 5 de l'algorithme décrit ci-dessus peut être remplacée par la simple soustraction $w_{j+1} = c_j - c_{j+1}$. Dans ce cas, la transformation ondelette inverse à partir des coefficients $\mathcal{W} = \{w_1, \ldots, w_J, c_J\}$ est triviale. Il suffit de sommer les différentes échelles et le plan lissé :

$$c_0(\theta, \phi) = c_J(\theta, \phi) + \sum_{j=1}^{J} w_j(\theta, \phi) \tag{3.27}$$

Il s'agit de la même formule de reconstruction que l'algorithme *à trous* : la somme de toutes les échelles reproduit les données de départ. Cependant, comme cette transformation est redondante, la transformation inverse n'est pas unique et peut donc se faire d'autres manières, et ainsi, on peut ajouter d'autres contraintes à la fonction de synthèse pendant la reconstruction (Starck et al. 2007; Starck and Fadili 2007). Par exemple, en utilisant les relations suivantes :

$$\begin{aligned} \hat{c}_{j+1}(l,m) &= \widehat{H}_j(l,m)\hat{c}_j(l,m) \\ \hat{w}_{j+1}(l,m) &= \widehat{G}_j(l,m)\hat{c}_j(l,m) \end{aligned} \tag{3.28}$$

une estimation au moindre carré de c_j à partir de c_{j+1} et w_{j+1} donne :

$$\hat{c}_j = \hat{c}_{j+1}\widehat{\tilde{H}}_j + \hat{w}_{j+1}\widehat{\tilde{G}}_j \tag{3.29}$$

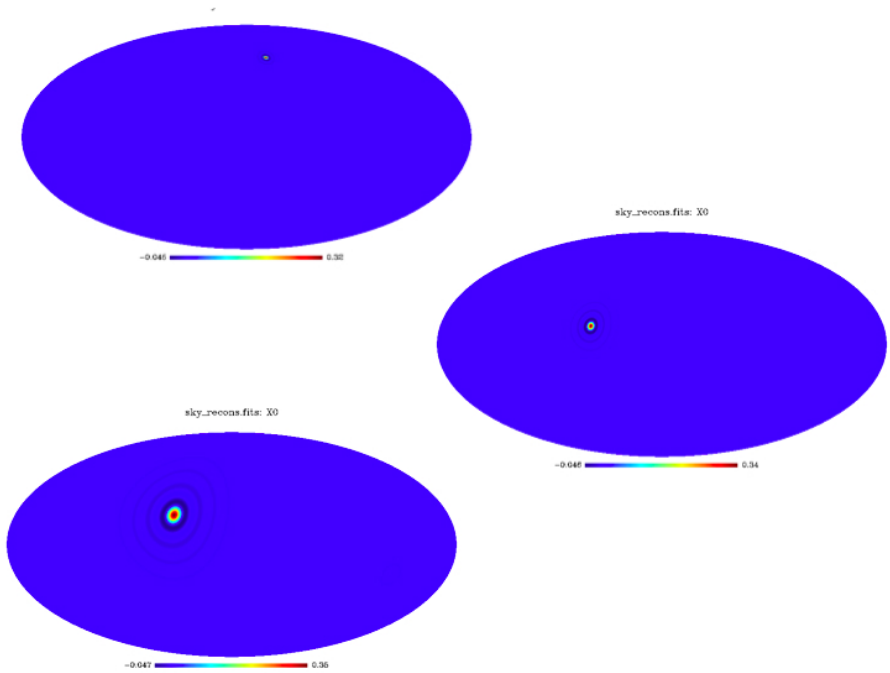

FIG. 3.12 – Projection inverse d'un coefficient ondelette à différentes echelles. Chaque carte est obtenue en forçant tous les coefficients à zéro sauf un, et en appliquant la projection ondelette inverse. En fonction de la position du coefficient et de son échelle, la carte reconstruite nous montre des motifs isotropiques de différentes tailles et à différentes positions.

où les filtres conjugués $\widehat{\tilde{H}}_j$ et $\widehat{\tilde{G}}_j$ ont pour formes :

$$\widehat{\tilde{H}}_j = \sqrt{\frac{4\pi}{2l+1}}\hat{\tilde{h}}_j \quad = \widehat{H}_j^* / (\mid \widehat{H}_j \mid^2 + \mid \widehat{G}_j \mid^2) \tag{3.30}$$

$$\widehat{\tilde{G}}_j = \sqrt{\frac{4\pi}{2l+1}}\hat{\tilde{g}}_j \quad = \widehat{G}_j^* / (\mid \widehat{H}_j \mid^2 + \mid \widehat{G}_j \mid^2) \tag{3.31}$$

Ainsi l'algorithme de reconstruction s'exprime de la façon suivante :

1. Calculer la fonction d'échelle B_3-spline puis en déduire $\hat{\psi}$, $\hat{h}$, $\hat{g}$, $\hat{\tilde{h}}$, $\hat{\tilde{g}}$ numériquement.
2. Calculer les harmoniques sphériques du plan lissé c_J. On obtient $\hat{c}_J$.
3. Initialiser j à $J-1$. Itérer :
   4. Calculer les harmoniques sphériques des coefficients ondelette w_{j+1} à l'échelle $j+1$. On obtient $\hat{w}_{j+1}$.
   5. Multiplier $\hat{c}_{j+1}$ par $\widehat{\tilde{H}}_j$.
   6. Multiplier $\hat{w}_{j+1}$ par $\widehat{\tilde{G}}_j$.
   7. Ajouter les résultats des étapes 6 and 7. On obtient $\hat{c}_j$.
   8. j=j-1 et si $j \geq 0$, retourner à l'étape 4.
9. Calculer la transformée en harmoniques sphériques inverse de $\hat{c}_0$

Les filtres de synthèse passe-haut et passe-bas $\hat{\tilde{h}}$ et $\hat{\tilde{g}}$ sont représentés sur la figure Fig. 3.11. La figure 3.12 montre des projections inverses d'un coefficient ondelette à différentes échelles. Chaque carte est obtenue en forçant tous les coefficients à zéro sauf un, et en appliquant la projection ondelette inverse. En fonction de l'échelle du coefficient, la carte reconstruite nous montre des motifs isotropiques de différentes tailles qui correspondent à la forme de l'ondelettes dans l'espace direct.

3.3.2 Transformée en ondelettes isotrope pyramidale sur la sphère

Dans l'algorithme décrit précédemment, aucune décimation n'est effectuée : chaque échelle ondelette a la même taille que les données originales. Ainsi, le nombre de coefficients d'ondelette final est égal au nombre de pixels de l'image originale multiplié par le nombre d'échelles. Pour des applications futures aux données fournies par le satellite PLANCK, il est souhaitable de pouvoir introduire une décimation dans la décomposition pour réduire la consommation mémoire et le temps d'exécution de la transformation ondelette. Ceci peut être effectué en utilisant une propriété spécifique de la fonction échelle choisie. En effet, comme nous choisissons une fonction échelle avec une fréquence de coupure l_c dans les harmoniques sphèriques, le nombre des coefficients d'ondelette significatifs est divisé par quatre après chaque changement d'échelle, la fréquence de coupure étant divisée par deux à chaque étape. Ainsi, moins de pixels sont nécessaires dans l'espace direct lors du calcul de la transformation en harmoniques sphériques inverse. En utilisant la pixelisation HEALPix (Górski et al. 2005), ceci peut être fait facilement en divisant par deux le paramètre *nside* à chaque échelle. Pour la pixelisation Glesp, on utilise le nombre de pixels nécessaire à la répresentation pour une fréquence l_{max} donnée.

L'algorithme

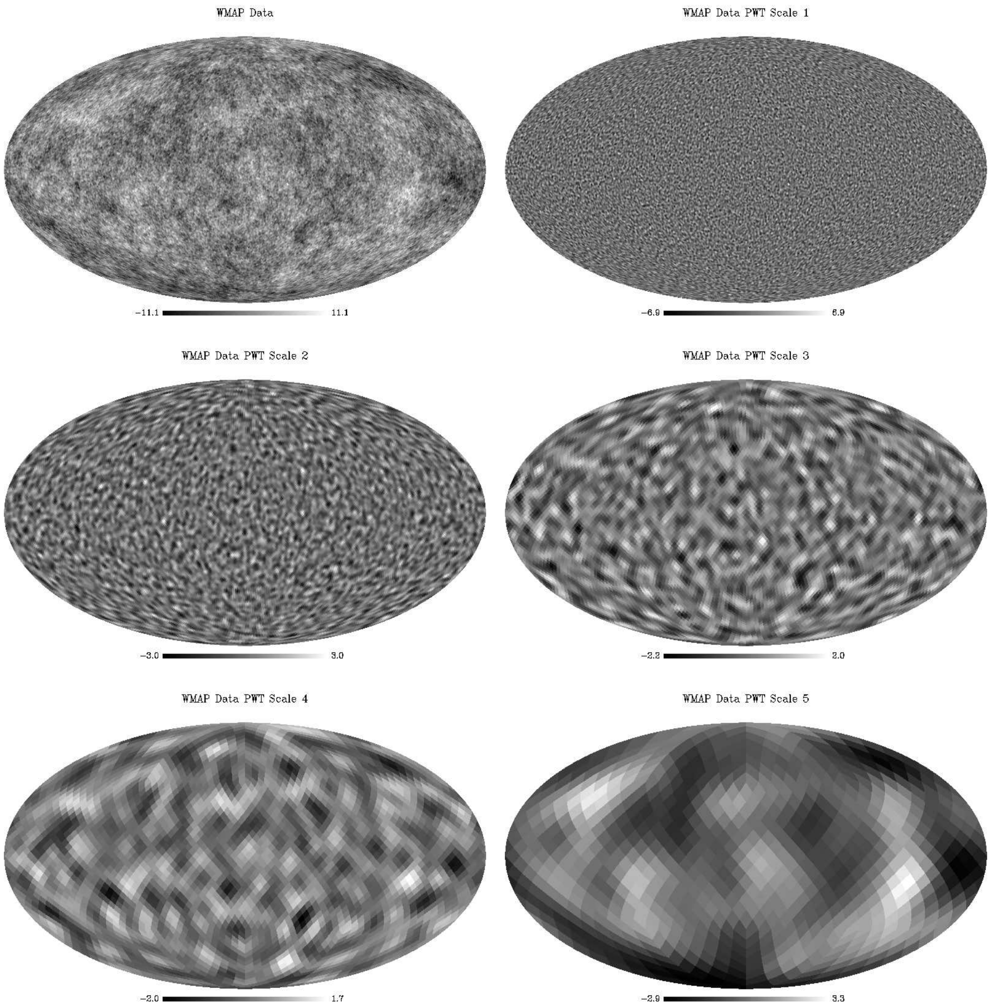

FIG. 3.13 – Carte de CMB obtenue par WMAP (en haut à gauche) et sa transformée ondelette pyramidale sur la sphère sur 5 échelles (4 échelles ondelettes et le plan lissé). Dans l'ordre, de gauche à droite et de haut en bas : données originales, 4 échelles ondelette, 1 plan lisse.

Bien que l'algorithme de la transformation en ondelettes pyramidale soit assez proche de sa version non décimée, il est nécessaire d'apporter plusieurs modifications. On aboutit alors à l'algorithme suivant :

1. Calculer la fonction d'échelle B_3-spline puis en deduire numériquement ψ, h et g .
2. Calculer les coefficients $\hat{c}_0$ en harmoniques sphériques de l'image c_0.
3. Initialiser j à 0. Itérer :
   4. Effectuer le produit de $\hat{c}_j$ par $\widehat{g}_j$. On obtient $\hat{c}_{j+1}$.
   5. Effectuer la transformation inverse de $\hat{c}_{j+1}$ à la résolution 2^j
   6. Sur échantillonner c_{j+1} à la résolution 2^{j+1}
   Calculer, par différence de plan ondelette, $w_{j+1} = c_j - c_{j-1}$.
   7. Effectuer le produit de $\hat{c}_j$ by $\widehat{G}_j$. On obtient $\hat{w}_{j+1}$.
   Diviser la taille de l'échantillonage par 2, les données c_{j+1} n'ayant plus d'information au dela de la fréquence 2^j.
   8. j=j+1 et si $j \leq J$, retourner à l'étape 4 .
9. L'ensemble $\{w_1, w_2, \ldots, w_J, c_J\}$ est la transformation en ondelette de c_0.

Algorithme de la transformée en ondelettes pyramidale sur la sphère.

Cet algorithme permet également une reconstruction exacte.

La reconstruction

La reconstruction est un peu plus difficile que dans le cas non décimée, car maintenant, chaque échelle ondelette à une résolution différente. Il est donc nécessaire de remettre ces échelles à la même résolution avant de les sommer. Cet algorithme s'écrit alors de la façon suivante :

1. Calculer la fonction d'échelle B_3-spline et en déduire $\hat{\psi}$, $\hat{h}$, $\hat{g}$, $\hat{\tilde{h}}$, $\hat{\tilde{g}}$ numériquement.
2. Calculer les harmoniques sphériques du plan lissé c_J. On obtient $\hat{c}_J$.
3. Initialiser j à $J-1$. Itérer :
   4. Sur échantillonner c_{j+1} à la résolution 2^{j+1}.
   5. $c_j = w_j + c_{j+1}$.
   6. j=j-1 et si $j \geq 0$, retour à l'étape 4.

On obtient alors exactement c_0, l'image originale.

Algorithme de reconstruction de la transformée pyramidale sur la sphére.

Cette reconstruction est strictement exacte, puisque les différences entre les différentes échelles se fait dans l'espace pixel et non l'espace des harmoniques sphériques.

La figure 3.13 montre les données CMB de WMAP et sa transformée ondelette pyramidale en cinq échelles. Quand l'échelle augmente(la résulotion diminue), la taille des pixels augmente. Ainsi, pour les deux dernières échelles, la pixélisation commence à devenir visible.

3.4 Conclusion

Nous venons d'introduire deux nouvelles transformées en ondelettes isotropes sur la sphère qui possèdent la propriété très importante d'être inversibles :

- la première non-décimée qui est préférable pour certaines applications telles que le débruitage,
- la deuxième décimée qui sera bien utile pour le traitement de données colossales telles que les données Planck.

Avec la propriété d'inversibilité, de nombreuses applications deviennent possibles dans un cadre multi-échelles telles que : le débruitage, la déconvolution, la séparation de composantes...

L'analyse en ondelettes a connu de nombreux succès. Pour certaines applications, cependant, elle est loin d'offrir une solution optimale. Les ondelettes sont optimales pour représenter des structures isotropiques, elles rencontrent, par contre, des difficultés à représenter des images contenant des structures fortement anisotropiques. Dans le chapitre suivant nous allons introduire deux nouvelles transformées sur la sphère qui pourront être utilisées pour certaines applications où les ondelettes présentent quelques faiblesses.

Chapitre 4
Ridgelets et Curvelets sur la sphère

Sommaire

4.1 Introduction **53**
4.2 Transformée en ridgelets sur la sphère **54**
4.2.1 Transformée en ridgelets sur le plan 54
4.2.2 Transformée en ridgelets sur la sphère. 58
4.3 Transformée en curvelets sur la sphère **59**
4.3.1 Transformée en curvelets sur le plan 59
4.3.2 Transformée en curvelets sur la sphère 63
4.3.3 Transformée en curvelets pyramidale sur la sphère 64
4.4 Conclusions **66**

4.1 Introduction

Si les ondelettes sont particulièrement efficaces pour la détection de structures isotropes de différentes échelles, elles ne sont par contre pas optimales pour la recherche d'objets anisotropes. De nouvelles transformées multi-échelles ont récemment été développées, les ridgelets et les curvelets, qui permettent de rechercher des objets de manière optimale, quand ces objets présentent de fortes anisotropies (Candès and Donoho 1999b,c). Les curvelets ont été utilisées dans des applications de restauration de données et de rehaussement de contraste (Starck et al. 2002, 2003b, 2004a, 2003c). Il a aussi été prouvé que la transformée en curvelets peut être utile pour la détection et la discrimination de non gaussianité dans le CMB (Starck et al. 2004a; Jin et al. 2005).

Nous présentons dans ce chapitre comment les transformées en ridgelets et en curvelets peuvent être implémentées sur la sphère (Starck et al. 2006;

Abrial et al. 2007).

4.2 Transformée en ridgelets sur la sphère

4.2.1 Transformée en ridgelets sur le plan

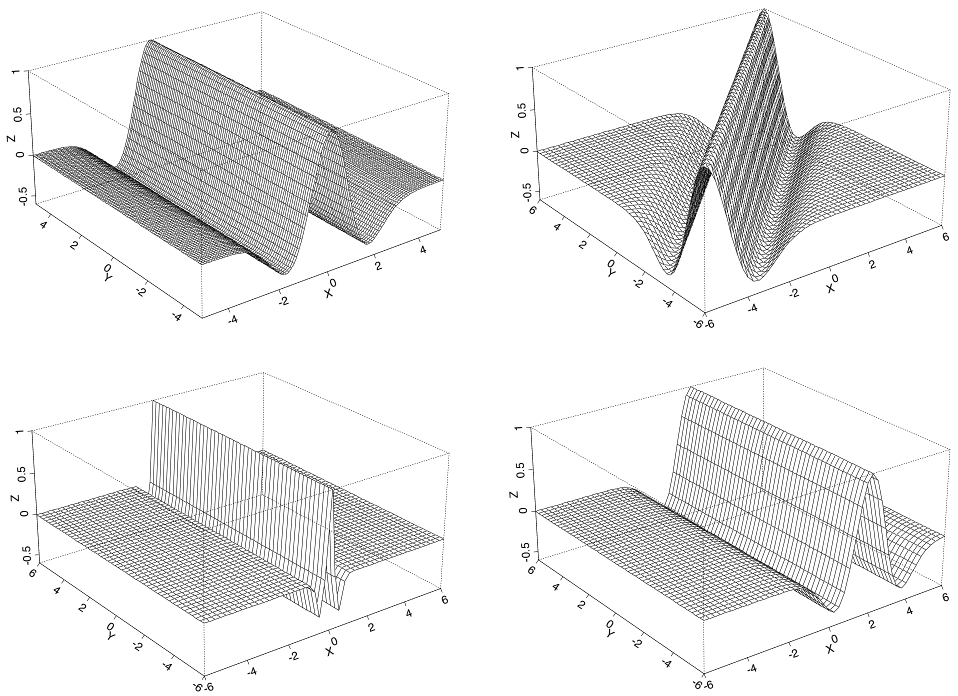

FIG. 4.1 – Exemples de fonctions ridgelets - Les figures en haut à droire, en bas à gauche et à droite sont obtenues par de simples manipulations géométriques (rotation, changement d'échelle et déplacement) de la figure en haut à gauche.

La transformée en ridgelets à deux dimensions peut être définie de la manière suivante (Candès and Donoho 1999b) : prenons une fonction lisse φ $\psi : \mathbf{R} \to \mathbf{R}$ à décroissance rapide et satisfaisant la condition d'admissibilité suivante

$$\int |\hat{\varphi}(\nu)^2|/|\nu|^2 d\nu < \infty, \tag{4.1}$$

qui est vérifiée si ψ a une moyenne nulle : $\int \psi(t)dt = 0$. Nous supposons que la normalisation de ψ est telle que , $\int_0^\infty |\hat{\psi}(\nu)|^2 \nu^{-2} d\nu = 1$

Pour chaque échelle $a > 0$, chaque position $b \in \mathbf{R}$ et chaque orientation $\theta \in [0, 2\pi]$, la fonction ridgelet est définie $\psi_{a,b,\theta} : \mathbf{R}^2 \to \mathbf{R}$ par :

$$\psi_{a,b,\theta}(x) = \psi_{a,b,\theta}(x_1, x_2) = a^{1/2} \cdot \psi((x_1 \cos\theta + x_2 \sin\theta - b)/a). \tag{4.2}$$

Ainsi, une fonction ridgelet est constante le long des lignes d'équation : $x_1 \cos\theta + x_2 \sin\theta = cst$. Perpendiculairement à ces lignes, on a une fonction ondelette. La figure 4.1 montre quelques exemples de ridgelets. Les figures en haut à droire, en bas à gauche et à droite sont obtenues par de simples manipulations géométriques (rotation, changement d'échelle et déplacement) de la figure 4.1 en haut à gauche.

Et les coefficients de la transformée en ridgelets d'une fonction f sont définis par :

$$\mathcal{R}_f(a,b,\theta) = \int \overline{\psi}_{a,b,\theta}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}. \tag{4.3}$$

Nous avons alors la formule de reconstruction exacte :

$$f(\mathbf{x}) = \int_0^{2\pi} \int_{-\infty}^{\infty} \int_0^{\infty} \mathcal{R}_f(a,b,\theta) \psi_{a,b,\theta}(\mathbf{x}) \frac{da}{a^3} db \frac{d\theta}{4\pi} \tag{4.4}$$

Celle-ci est valable partout pour des fonctions à la fois intégrable et de carré intégrable.

Implémentation de la transformée en ridgelets

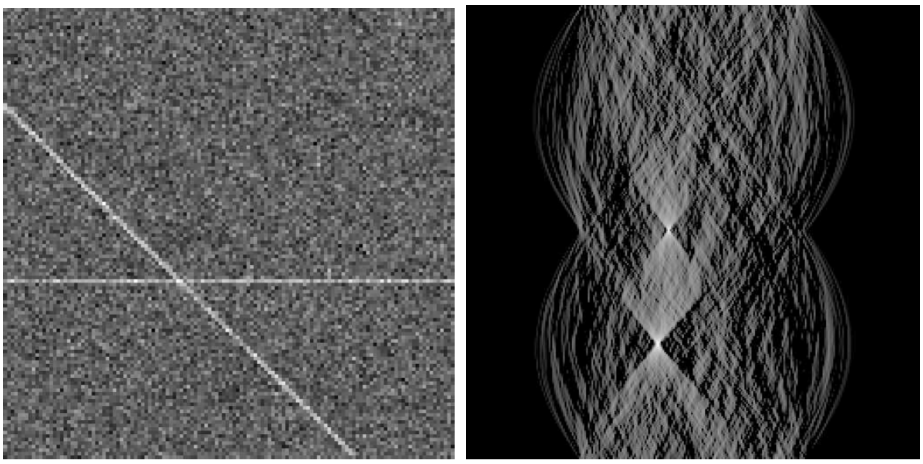

Fig. 4.2 – A gauche, image contenant deux lignes et du bruit Gaussien. A droite, sa transformée de Radon.

La tranformée en ridgelets peut être construite en faisant une analyse en ondelettes dans le domaine de Radon (Candès and Donoho 1999b). Rappelons que la transformée de Radon d'une fonction f est l'ensemble des intégrales sur les lignes indexées par $(\theta, t) \in \mathbf{R}$ formulée par :

$$Rf(\theta,t) = \int f(x_1,x_2) \delta(x_1 \cos\theta + x_2 \sin\theta - t)\, dx_1 dx_2, \tag{4.5}$$

où δ est l'opérateur de Dirac. La figure 4.2 montre une image contenant deux lignes et du bruit Gaussien (gauche), et sa transformée de Radon (droite).

La transformée en ridgelets est alors exactement une transformée en ondelettes 1D le long des lignes de la transformée de Radon où la variable angulaire θ est constante et t varie. Ainsi, une stratégie élémentaire pour calculer la transformée en ridgelets est de calculer d'abord la transformation de Radon $Rf(t, \theta)$, puis d'appliquer une transformée en ondelettes unidimensionelles sur les lignes $Rf(\cdot, \theta)$.

Un aspect fondamental lié à la transformée de Radon est le théorème *projection-slice formula* (Deans 1983) :

$$\hat{f}(\lambda \cos \theta, \lambda \sin \theta) = \int Rf(t, \theta) e^{-i\lambda t} dt.$$

La transformée de Radon peut donc être obtenue en appliquant tout d'abord une transformée de Fourier 2D sur l'image, puis en extrayant des lignes dans l'espace de Fourier qui passe par l'origine, et finalement en faisant une transformée de Fourier 1D de chacune de ces lignes (Toft 1996; Averbuch et al. 2001).

Ridgelets locales

La transformée en ridgelets est optimale pour trouver des lignes globales dont la taille a pour ordre de grandeur celle de l'image. Pour détecter des segments de taille plus petite, une partition de l'image doit être introduite. L'image peut être décomposée en blocs se chevauchant de b pixels par côté, de telle façon que le chevauchement entre deux blocs verticaux adjacents soit un rectangle de taille $b \times b/2$. Le chevauchement est utilisé de manière à prévenir l'apparition d'artefacts liés aux blocs. Ainsi pour une image de taille $n \times n$, nous avons donc $2n/b$ blocs dans chaque direction, et ainsi le facteur de redondance est égal à 4.

La partion introduit de la redondance, car un pixel appartient à 4 blocs voisins. Nous présentons deux stratégies intéressantes pour effectuer l'analyse et la synthèse :

1. Les valeurs des blocs sont pondérées par une fenêtre spatiale w (analyse) de telle façon que la somme de tous les blocs reproduise exactement la valeur originale du pixel (synthèse).
2. Les valeurs des blocs sont celles des pixels de l'image (analyse), mais elles sont pondérées lorsque que l'image est reconstruite (synthèse).

L'expérience a montré que la seconde approche apporte de meilleurs résultats, surtout pour les problèmes de restauration. La valeur d'un pixel, $f[i_1, i_2]$ à partir de ces quatre blocs correspondants de demi-taille $m = b/2$, nommés, $B_1[k_1, l_1], B_2[k_2, l_1], B_3[k_1, l_2]$ et $B_4[k_2, l_2]$, avec $k_1, l_1 > b/2$ et $k_2 =$

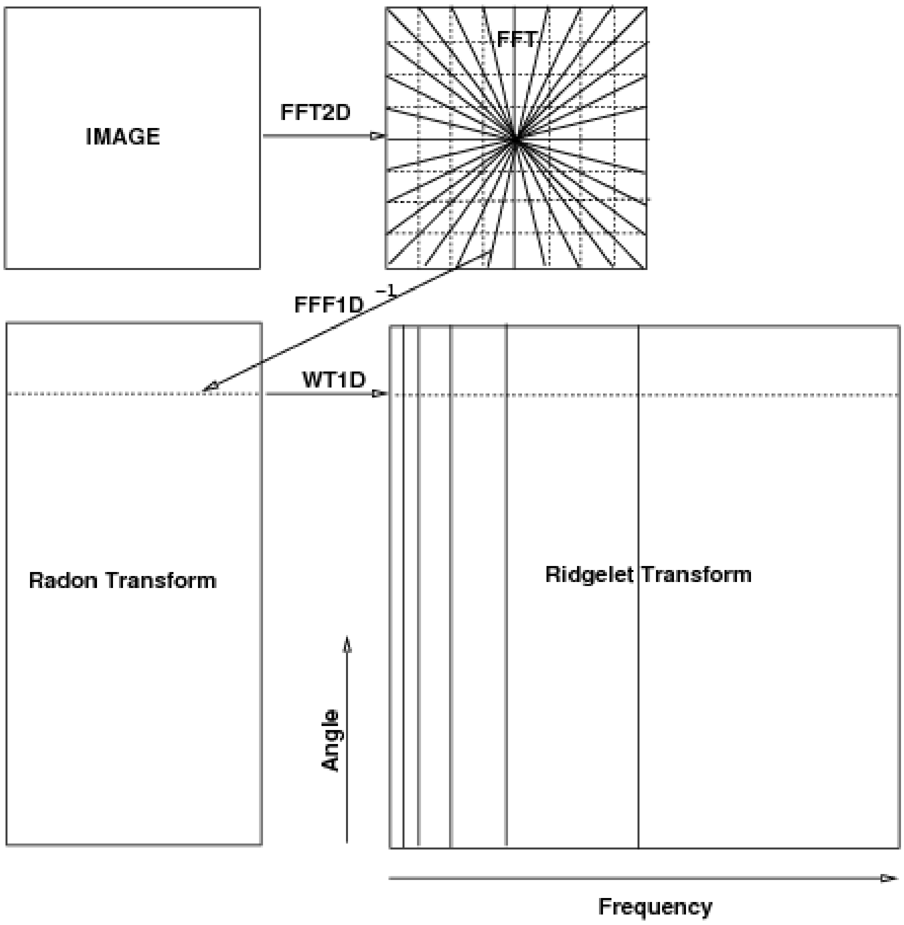

Fig. 4.3 – Diagramme de l'algorithme de la transformée en ridgelets. La transformée de Radon est obtenue en appliquant tout d'abord une transformée de Fourier 2D sur l'image, puis en extrayant des lignes dans l'espace de Fourier qui passe par l'origine, et finalement en faisant une transformée de Fourier 1D de chacune de ces lignes. La transformée en ondelettes 1D le long ces lignes dans l'espace de Radon nous donne les coefficients ridgelets.

$k_1 - m$, $l_2 = l_1 - m$, est calculée de la manière suivante :

$$\begin{aligned} f_1 &= w(k_2/m)B_1[k_1, l_1] + w(1 - k_2/m)B_2[k_2, l_1] \\ f_2 &= w(k_2/m)B_3[k_1, l_2] + w(1 - k_2/m)B_4[k_2, l_2] \\ f[i_1, i_2] &= w(l_2/m)f_1 + w(1 - l_2/m)f_2 \end{aligned} \tag{4.6}$$

où $w(x) = \cos^2(\pi x/2)$ est la fonction fenêtre choisie. Bien sûr, d'autres fonctions sont possibles, à la condition qu'elles vérifient :

$$\begin{aligned} w(0) &= 1 \\ w(1) &= 0 \\ w'(0) &= 0 \\ w(x) + w(1 - x) &= 1 \end{aligned}$$

4.2.2 Transformée en ridgelets sur la sphère.

Partition de la sphère.

Afin de pourvoir appliquer une transformée en ridgelets locale, un partition préalable de la sphère est nécessaire. Comme cela a été décrit dans le second chapitre, la pixélisation HEALPix permet de diviser la sphère en blocs dont la projection plane est assez proche du quadrilatère. Cette division se fait de plus de manière hiérarchique, si bien que chaque face est donc composée de $nside^2$ pixels. Cette particularité nous permet donc de découper des données sphériques en blocs de taille de la forme 2^n. Ainsi, certaines transformations 2D usuelles peuvent être appliquées sur la sphère. Un effet de bloc peut apparaître lors d'opérations de reconstruction nonlinéaire sur la transition entre blocs. Pour supprimer cet effet, il est possible de travailler sur différentes rotations de la sphère, afin de déplacer ces transitions en des endroits différents des données, puis de fusionner les différents résultats (*cycle spinning*). Ceci est rendu possible par le fait qu'il existe plusieurs rotations qui ne modifient pas le schéma de pixélisation avec HEALPix.

Algorithme

Une fois que la partition de la sphère est effectuée, la transformée en ridgelets 2D décrite dans Starck et al. (2003a) peut être appliquée sur chaque bloc. Voici la description de la transformée en ridgelets sur le plan (voir figure 4.3) :

1. Calcul de la transformée de Fourier 2D.
2. Extraction des lignes passant par l'origine des fréquences.
3. Calcul de la transformée de Fourier inverse 1D sur chaque ligne. Ceci nous permet de calculer la transformée de Radon.

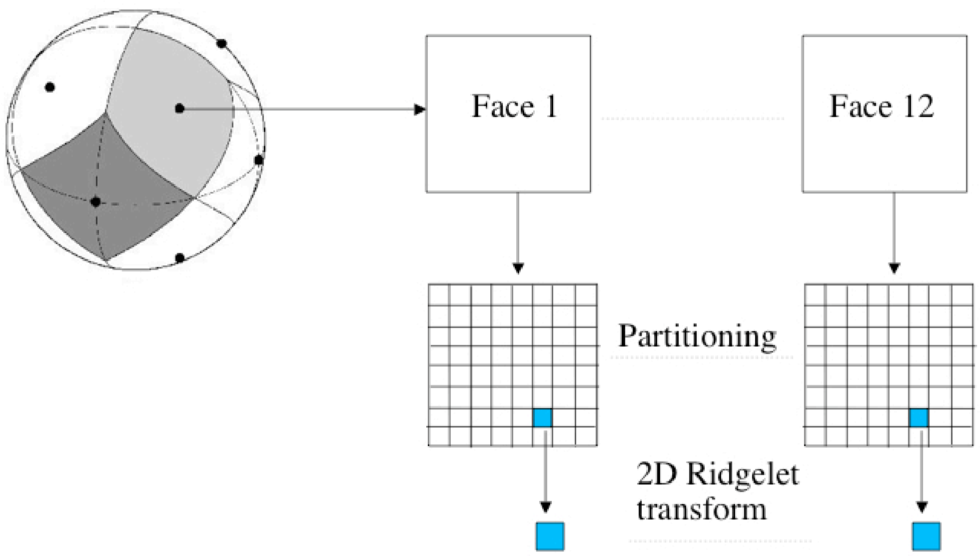

FIG. 4.4 – Diagramme de la transformée en ridgelets sur la sphère.

4. Calcul de la transformée en ondelettes 1D des lignes de la transformée de Radon.

Les trois premières étapes correspondent à la transformée de Radon appelée *linogramme*. D'autres méthodes pour la transformée de Radon peuvent aussi être utilisées, comme la *Slant Stack Radon Transform* (Donoho and Flesia 2002).

La figure 4.4 montre le diagramme de la transformée en ridgelets sur la sphère. La figure 4.5 présente la projection inverse de coefficients ridgelets à différentes échelles et différentes orientations.

4.3 Transformée en curvelets sur la sphère

Dans le traitement d'images, les contours contenus dans une image ne sont jamais vraiment droits, mais plutôt courbés, et les ridgelets ne sont pas capables de représenter correctement de telles images. Cependant, il est toujours possible d'utiliser l'algorithme ridgelet de manière locale de telle façon qu'une courbe peut être approximée par des segments de droite. La principale difficulté est alors de fixer la taille des segments de droite, qui correspond à la taille des blocs dans l'analyse en ridgelets.

4.3.1 Transformée en curvelets sur le plan

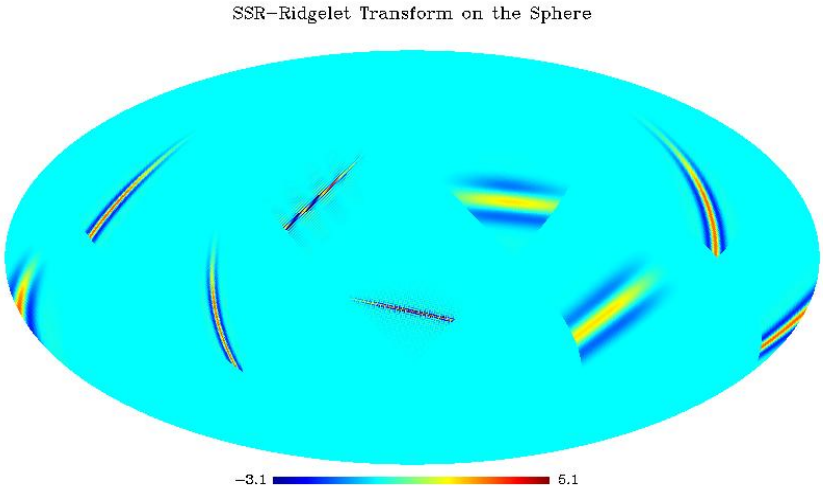

Fig. 4.5 – Projection inverse de coefficients ridgelets à différentes échelles et à différentes orientations. Tous les coefficients sont mis à zéro sauf un, et une transformation inverse est appliquée à ces coefficients. En fonction de l'échelle et de la position du coefficient différent de zéro, les données reconstruites ont des formes de longueur définie et d'orientation définie.

La transformée en curvelets (Candès and Donoho 1999a; Donoho and Duncan 2000; Starck et al. 2002) a ouvert la possibilité d'analyser une image avec des tailles de bloc différentes, mais avec une seule transformation. L'idée est de commencer par décomposer l'image en un ensemble de plans d'ondelettes, et ensuite d'analyser chaque bande par une transformée en ridgelets locale. La taille des blocs peut ainsi être changée à chaque échelle. Les différents niveaux de la pyramide en ridgelets multi-échelles sont utilisés pour représenter les différentes sous-bandes de la transformée en ondelettes. Cette décomposition en sous-bandes impose une relation entre la longueur et la largeur des élements, de telle façon qu'ils soient anisotropes et obéissent approximativement à une loi d'échelle parabolique $largeur^2 \approx longueur$.

La transformée en curvelets discrète d'une fonction continue $f(x)$ utilise une séquence dyadique d'échelles, et un banc de filtres, qui a pour propriété que sa bande passante δ_j est concentrée sur des fréquences $[2^{2j}, 2^{2j+2}]$, c'est-à-dire :

$$\delta_j(f) = \Psi_{2j} * f, \hat{\Psi}_{2j}(\nu) = \hat{\Psi}(2^{-2j}\nu) \tag{4.7}$$

Dans la théorie des ondelettes, on utilise habituellement une décomposition en bandes dyadiques $[2^j, 2^{j+1}]$. Par opposition, les sous-bandes utilisées dans la transformée en curvelets discrète ont la forme non standard $[2^{2j}, 2^{2j+2}]$. Ceci est une particularité non-standard de la transformée en curvelets qui est importante, car c'est ceci qui fixe la loi d'échelle parabolique.

Les différentes étapes de la décomposition sont les suivantes :

- *Décomposition en sous-bandes.* L'image f est décomposée en sous-bandes.
- *Partionnement doux.* Chaque sous-bande est fenêtrée avec un filtre lisseur à une échelle appropriée (de taille $\sim 2^{-j}$).
- *Transformation en ridgelets.* Chaque bloc est analysé par une transformée en ridgelets.

Dans cette définition, les deux sous-bandes $[2^{2j}, 2^{2j+1}]$ et $[2^{2j+1}, 2^{2j+2}]$ sont fusionnées avant d'appliquer la transformée en ridgelets.

Implémentation numérique

La transformée en ondelettes isotrope "à trous" est spécialement bien adaptée au besoin de la transformée en curvelets discrète (Starck et al. 2002). Cet algorithme décompose une image de taille $n \times n$ en une superposition de plans :

$$f[i_1, i_2] = c_J[i_1, i_2] + \sum_{j=1}^{J} w_j[i_1, i_2] \tag{4.8}$$

où c_J est le plan lissé de l'image originale et les w_j représentent les coefficients en ondelettes de f à l'échelle j. Ainsi, l'algorithme fournit $J+1$ sous-bandes de taille $n \times n$.

L'algorithme de la transformée en curvelets 2D peut être résumé par le pseudo-code suivant :

1. Appliquer la transformation isotrope "à trous" à J échelles.
2. Initialiser $B_1 = B_{min}$, où B_{min} est la taille du bloc de la première échelle.
3. Pour $j = 1, \cdots, J$ faire
4. Partionner la sous-bande w_j avec une taille de bloc B_j et appliquer la transformée en ridgelets sur chaque bloc.
5. si j modulo $2 = 1$ alors
6. $B_{j+1} = 2B_j$
7. sinon $B_{j+1} = B_j$
8. fin Si
9. fin Pour

La taille des fenêtres de localisation est doublée toutes les deux sous-bandes dyadiques, permettant de garantir la propriété fondamentale de la transformation en curvelets, qui est que les éléments de longueur proche de $2^{-j/2}$ sont utilisés pour l'analyse et la synthèse de la $j$ème sous-bande $[2^j, 2^{j+1}]$. On peut également noter que le plan lissé c_J de l'image est laissé intact. La valeur $B_{min} = 16$ est couramment utilisée dans la plupart des implémentations. La figure 4.6 montre le diagramme de l'algorithme.

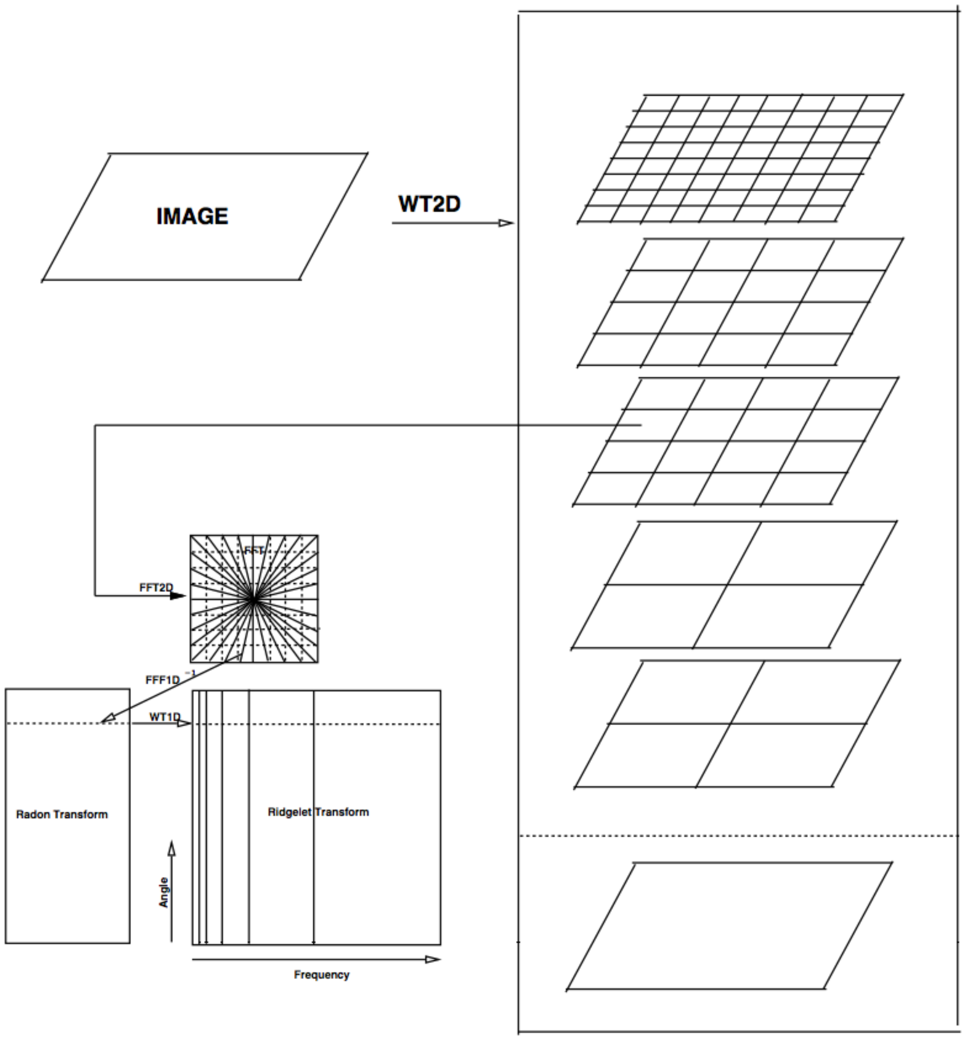

Fig. 4.6 – Diagramme de la transformée en curvelets discrète. Cette figure montre la décomposition en sous-bandes de l'image originale, suivie de la partition spatial de chaque sous-bande. Une transformée en ridgelets est appliquée ensuite sur chaque bloc.

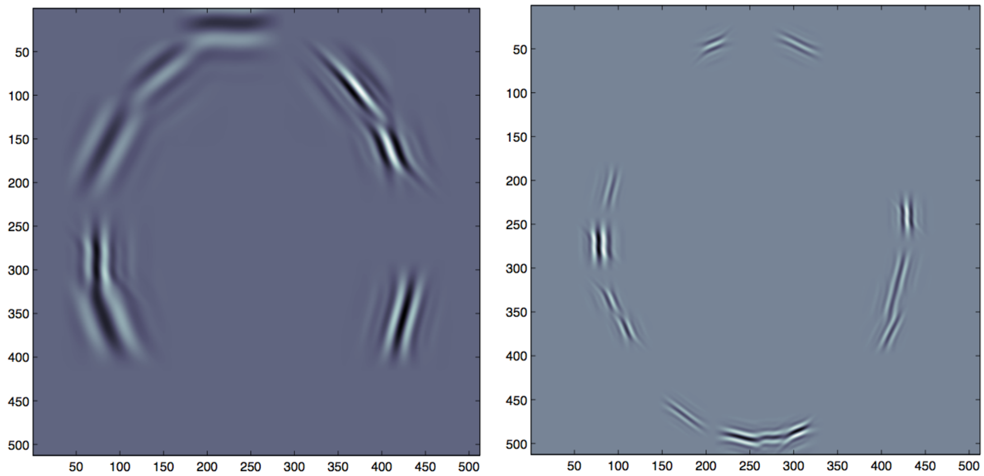

FIG. 4.7 – Quelques exemples de curvelets.

Cette implémentation est redondante. Le facteur de redondance est égal à $16J+1$, lorsque J échelles sont utilisées. Cet algorithme permet une reconstruction exacte, du fait que chaque étape de la décomposition est elle-même inversible.

On peut ainsi montrer que la compléxité numérique de la transformée en curvelets est en $O(n^2(\log n)^2)$ pour une image de taille $n \times n$ (Fadili and Starck 2008). La figure 4.7 montre quelques curvelets à différentes échelles, orientations et positions. Plus de détails sur l'implémentation de la transformée en curvelets 2D sont diponibles dans Starck et al. (2002, 2003a, 2004b); Fadili and Starck (2008).

4.3.2 Transformée en curvelets sur la sphère

A partir de la transformée en ondelettes isotrope sphérique décrite au chapitre précedent et de la transformée en ridgelets sur la sphère présentée dans ce chapitre, il est maintenant possible d'imaginer une extension de la transformée en curvelets pour la sphère.

Algorithme de la transformée en curvelets

La transformée en curvelets sur la sphère se décompose de la manière suivante :

1. Réaliser une transformée en ondelettes isotrope sur la sphère avec J échelles,
2. Initialiser la taille des blocs initiaux à $B_1 = B_{min}$,
3. pour $j = 1, \ldots, J$:
   – faire une partition des échelles w_j, avec une taille de bloc B_j.

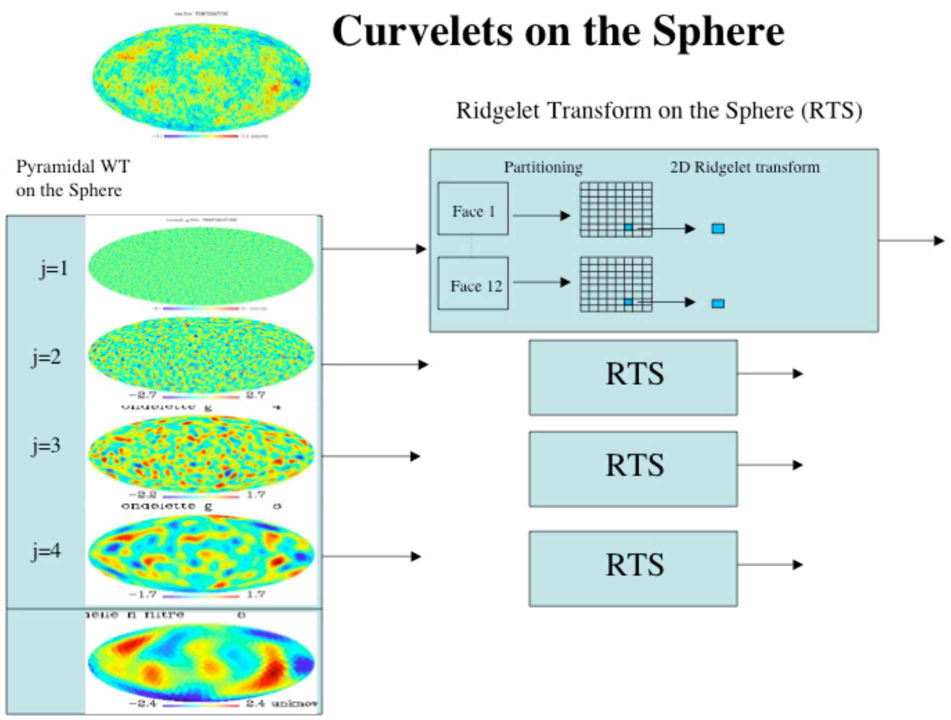

FIG. 4.8 – Diagramme de l'algorithme de la transformée en curvelets sur la sphère.

- réaliser une transformée en ridgelets sur chaque bloc.
- si j modulo $2 = 1$ alors $B_{j+1} = 2B_j$,
- sinon $B_{j+1} = B_j$.

La longueur de la fenêtre est doublée pour chaque sous-bande dyadique, afin de conserver la propriété fondamentale de la transformée en curvelets qui consiste à dire que des éléments de taille $2^{-j/2}$ sont utilisés pour l'analyse et la synthèse de la $j$ème sous-bande $[2^j, 2^{j+1}]$. On utilise dans notre implémentation une taille de bloc minimale $B_{min} = 16$ pixels. La figure 4.8 présente un diagramme de l'algorithme.

La figure 4.9 montre la transformation inverse de quelques coefficients curvelets à différentes échelles et orientations.

4.3.3 Transformée en curvelets pyramidale sur la sphère

La transformée en curvelets est redondante, avec un facteur de redondance de $16J+1$ où J échelles sont utilisées. Cette forte redondance peut être un problème pour manipuler de grands ensembles de données comme ceux de

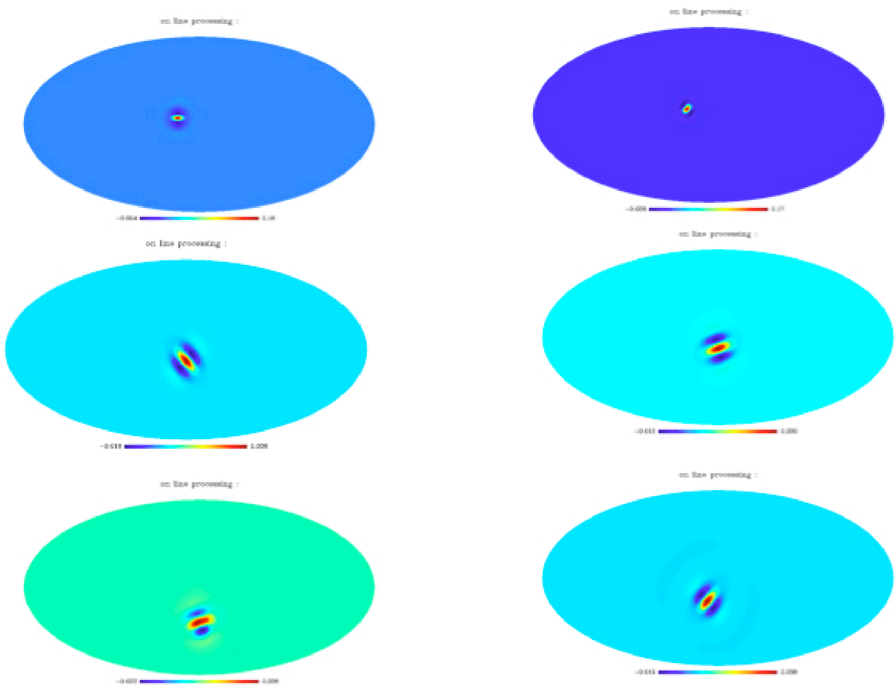

FIG. 4.9 – Projection inverse de coefficients curvelets à différentes échelles et différentes orientations. Tous les coefficients sont mis à zéro sauf un, et une transformation inverse est appliquée à ces coefficients. En fonction de l'échelle et de la position du coefficient, choisi différent de zéro, les données reconstruites ont des formes de longueur définie et d'orientation définie.

PLANCK. Cette redondance peut être aisément réduite en subsituant dans la transformée en curvelets sphérique la transformée en ondelettes pyramidale sur la sphère à la transformée en ondelettes non décimée. L'algorithme de la transformée en curvelets pyramidale est le suivant :

1. Réaliser une transformée en ondelettes pyramidale isotrope sur la sphère avec J échelles,
2. Initialiser la taille des blocs initiaux à $B_1 = B_{min}$,
3. pour $j = 1, \ldots, J$:
   – faire une partition des échelles w_j, avec une taille de bloc B_j.
   – réaliser une transformée en ridgelets sur chaque bloc.
   – si j modulo $2 = 1$ alors $B_{j+1} = MAX(8, B_j/2)$,
   – sinon $B_{j+1} = B_j$.

4.4 Conclusions

La transformée en curvelets 2D sur la sphère que nous avons présentée dans ce chapitre permet l'analyse directionnelle d'images à des échelles différentes. La transformée en curvelets sur la sphère est assez semblable à la transformée en curvelets sur la plan. Cependant, l'algorithme à trous utilisé sur le plan est, dans ce cas précis, remplacé par la transformée en ondelettes isotrope, décrite au chapitre précédent. Ainsi, les données sont d'abord transformées par une transformée ondelettes isotrope non décimée ou pyramidale. Ensuite, chaque échelle j est décomposée en blocs de taille B_j pixels qui se superposent et une transformée en ridgelets est appliquée sur chacun des blocs. La transformée en ridgelets se décompose elle même finalement, en l'application d'une transformée en ondelettes 1D sur les lignes de la transformée de Radon.

Dans les chapitres suivants, nous verrons comment la transformée en curvelets peut être utilisée dans des applications de restauration et de test de Gaussianité.

Chapitre 5

Débruitage multi-échelles sur la sphère

Sommaire

5.1 Introduction au débruitage 68

5.2 Coefficients significatifs 68

5.2.1 Définition 68

5.2.2 Modélisation du bruit 69

5.2.3 Estimation des caractéristiques d'un bruit gaussien 70

5.3 Méthodes de seuillage 71

5.3.1 Le seuillage *dur* et le seuillage *doux* 71

5.3.2 La méthode de seuillage FDR 72

5.4 Méthode de filtrage combinée sur la sphère 75

5.4.1 Principe 76

5.4.2 Algorithme 77

5.4.3 Résultats 77

5.5 Conclusion 79

Les différentes transformations multi-échelles présentées dans les chapitres précédents permettent l'analyse et la reconstruction effective de données sur la sphère. À l'instar de ce qui se fait dorénavant de manière usuelle en traitement de signaux 1D ou d'images 2D, où les décompositions multi-échelles interviennent largement dans des applications allant de la restauration à la compression, la détection etc., nous avons étudié la possibilité d'utiliser les transformations multi-échelles sur la sphère à des fins analogues s'agissant de traiter des données structurées dans une topologie sphérique.

Dans ce chapitre, nous présentons plusieurs méthodes de débruitage et de restauration sur la sphère utilisant les représentations multi-échelles décrites précédemment.

5.1 Introduction au débruitage

Les ondelettes et les curvelets ont déjà été utilisé avec succès pour le débruitage d'images 2D usuelles *via* un filtrage non linéaire ou des méthodes de seuillage, comme cela est montré dans Starck and Murtagh (2006). Le seuillage dur, par exemple, consiste à mettre à zéro tous les coefficients non significatifs dans une représentation des données initiales, c'est à dire tous les coefficients dont la valeur absolue est plus faible qu'un certain seuil. En pratique, le seuil retenu est une fonction de l'écart-type du bruit dans la représentation utilisée, σ_j qu'il convient d'estimer éventuellement dans chaque bande j voire chaque direction, etc. Un coefficient w_j dans la représentation des données est alors jugé significatif et donc intéressant au sens de témoin d'une structure ou d'un signal dans les données, dès lors que $\mid w_j \mid > k\sigma_j$. La valeur du paramètre k est fixée par l'utilisateur typiquement entre 3 et 5, en fonction de son conservatisme. L'écart-type σ_j dans chaque bande j peut, selon la complexité de la transformation retenue, être obtenu soit analytiquement, soit à partir de simulations numériques comme indiqué dans Starck and Murtagh (2006). En notant D les données bruitées et δ l'opérateur de seuilage, les données filtrées $\tilde{D}$ sont obtenues par :

$$\tilde{D} = \mathcal{R}\ \delta(\ \mathcal{T}\ D) \tag{5.1}$$

où $\mathcal{T}$ et $\mathcal{R}$ désignent respectivement l'opérateur de transformation (*e.g.* en ondelettes, en curvelets) et son inverse ou une reconstruction associée lorsque la transformation est redondante.

5.2 Coefficients significatifs

5.2.1 Définition

Dans la plupart des applications, il est nécessaire de déterminer si la valeur d'un coeffcient dans une représentation des données en ondelettes ou autre, traduit l'existence d'une structure intéressante et comporte donc une contribution *signal* ou au contraire si cette valeur est complètement anodine et relève simplement du bruit sur les observations. Les premiers coefficients sont qualifiés de significatifs tandis que les seconds pourront être rejetés.

La transformation en ondelettes ou plus généralement multi-échelle decompose une image d'entrée en un ensemble d'images, de bandes à des résolutions différentes. Nous notons $w_{j,k}$ les coeffcients produits par cette transformation à l'échelle j. L'indice k est généralement un indice de position mais peut également repérer une direction dans le cas des curvelets par exemple. Connaissant *a priori* les propriétés du bruit entachant les données, il est possible de déterminer les propriétes statistiques du bruit dans la nouvelle représentation, soit analytiquement soit numériquement. En s'appuyant sur la loi du bruit dans la représentation choisie, il est simple d'établir un test d'hypothèse pour juger de la pertinence d'un coefficient au regard de l'objectif de débruitage.

Notons $\mathcal{H}_0$ l'hypothèse que l'image est localement constante à l'échelle j. Le réjet de l'hypothèse $\mathcal{H}_0$ dépend de la probabilité suivante :

$$P = Prob(\mid w_{j,k} \mid < \ \tau \mid \mathcal{H}_0) \tag{5.2}$$

où le seuil de détection, τ, est défini pour chaque échelle. Pour un seuil d'estimation donné, ϵ, si $P = P(\tau) > \epsilon$ l'hypothèse nulle n'est pas exclue. Bien que non nulle, la valeur du coeffcient pourrait bien n'être due qu'au seul bruit. En revanche, si $P < \epsilon$, la valeur du coeffcient ne peut vraisemblablement pas être due au bruit seul, et donc l'hypothèse 0 est rejeté. Dans ce cas, un coeffcient significatif a été détecté.

5.2.2 Modélisation du bruit

Supposons, sous l'hypothèse $\mathcal{H}_0$ que les données ne contiennent que du bruit, que la distribution des coeffcients $w_{j,k}$ est gaussienne centrée et d'écart-type σ_j :

$$p(w_{j,l}) = \frac{1}{\sqrt{2\pi}\sigma_j} e^{-w_{j,l}^2/2\sigma_j^2} \tag{5.3}$$

Le rejet de l'hypothèse $\mathcal{H}_0$ dépend donc de :

$$P = Prob(w_{j,l} > W) = \frac{1}{\sqrt{2\pi}\sigma_j} \int_{w_{j,l}}^{+\infty} e^{-W^2/2\sigma_j^2} dW \tag{5.4}$$

Pour un bruit gaussien donné, il suffit donc de comprarer $w_{j,l}$ à $k\sigma_j$. Souvent k est choisi avec pour valeur 3, ce qui correspond à $\epsilon = 0.002$ approximativement :

$$\begin{array}{lll} \text{si } \mid w_{j,l} \mid \geq \ k\sigma_j & \text{alors } w_{j,l} \text{ est significatif} \\ \text{si } \mid w_{j,l} \mid < \ k\sigma_j & \text{alors } w_{j,l} \text{ n'est pas significatif} \end{array} \tag{5.5}$$

Nous avons donc besoin d'estimer, dans le cadre d'un modèle de bruit Gaussien, l'écart-type du bruit à chaque échelle. Cet écart-type peut être déterminé de façon analytique, mais ces calculs peuvent devenir assez vite très compliqués. La bonne valeur de σ_j dans la succession des plans ondelettes est définie par l'écart-type du bruit σ_N dans les données originales D, et à partir de l'étude de ce bruit dans l'espace ondelette. Cette étude consiste en la simulation d'un jeu de données contenant un bruit gaussien avec un écart-type égal à 1, et à calculer la tranformation en ondelettes de ce jeu de données. Ensuite, l'écart-type σ_j^e est calculé dans chaque bande. Nous obtenons une courbe σ_j^e en fonction de j, donnée par le comportement du bruit dans les échelles ondelettes. Dans le cas d'une transformation en ondelettes orthonormée, cette courbe est triviale. Grâce aux propriétés de linéarité de la transformation en ondelettes (mais aussi de la transformation en curvelets), nous avons la propriété $\sigma_j = \sigma_N \sigma_j^e$. Ainsi, l'écart-type du bruit sur les données à l'échelle j est égal à l'écart-type du bruit σ_N multiplié par l'écart-type à l'échelle j des données simulées.

5.2.3 Estimation des caractéristiques d'un bruit gaussien

Le bruit gaussien d'écart-type σ_N peut être estimé automatiquement dans un jeu de données D. Cette estimation est particulièrement importante, car c'est à partir de l'écart-type σ_N que vont être déduits tous les écart-type σ_j aux échelles j. Ainsi, une erreur associée à l'estimation de σ_N va introduire une erreur sur tous les σ_j. On obtient une estimation plus fine du niveau de bruit dans les hautes fréquences des données, où en général le bruit domine le signal.

Cas d'un bruit blanc Gaussien

La méthode résultante consiste d'abord en un filtrage des données D avec un filtre moyenneur ou encore un filtre médian, puis à soustraire à D sa version filtré F : $S = D - F$. Dans notre cas, nous remplacons S par la première échelle d'une transformation ondelette ($S = w_1$), ce qui est plus rapide d'un point de vue numérique que d'appliquer un des filtres précédement cités. Cela est équivalent à soustraire à nos données D un filtre moyenneur.

L'histogramme de S montre un pic gaussien autour de zéro. Un écrétage à k-sigma est ensuite utilisé pour rejeter les pixels où le signal est fort de façon significative. Nous appelons $S^{(1)}$ le sous ensemble de S qui contient uniquement les pixels tels que $| S_l | < k\sigma_S$, où σ_S est l'écart-type de S, et k une constante en général prise égale à 3. En itérant, nous obtenons un sous ensemble $S^{(n+1)}$ vérifiant l'inégalité $| S_l^{(n)} | < k\sigma_{S^{(n)}}$, avec toujours $\sigma_{S^{(n)}}$ l'écart-type de $S^{(n)}$. Ainsi, une estimation robuste du bruit σ_1 dans w_1 (car $S = w_1$) est effectué par le calcul de l'écart-type de $S^{(n)}$ ($\sigma_1 = \sigma_{S^{(n)}}$). En pratique, trois itérations seulement sont nécessaires, et la précision est en général meilleure que 5%. σ_N est finalement calculé par :

$$\sigma_N = \frac{\sigma_1}{\sigma_1^e} = \frac{\sigma_{S^{(n)}}}{\sigma_1^e} \tag{5.6}$$

Cas d'un bruit coloré

Dans le cas d'un bruit gaussien corrèlé, les données peuvent être traitées comme dans le cas gaussien blanc avec comme différence que l'écart-type du bruit σ_j doit être estimé indépendement à chaque échelle j. Deux méthodes peuvent être utilisées :

1. σ_j peut être estimé par la méthode d'écrétage à k-sigma appliquée à chaque échelle j
2. La déviation médiane absolue, (*median absolute deviation–MAD*), peut être utilisé comme un estimateur de l'écart-type du bruit :

$$\sigma_j = \text{median}(\mid w_j \mid)/0.6745 \tag{5.7}$$

5.3 Méthodes de seuillage

De nombreuses techniques de seuillage ont été proposées ces dernières années.

5.3.1 Le seuillage dur et le seuillage doux

Le *seuillage dur* consiste à mettre à zéro tous les coefficients d'ondelette qui ont une valeur absolue inférieure à un certain seuil T_j (coefficients d'ondelette non significatifs) :

$$\tilde{w}_{j,k} = \begin{cases} w_{j,k} & \text{si } \mid w_{j,k} \mid \geq T_j \\ 0 & \text{sinon} \end{cases}$$

où $w_{j,k}$ est un coefficient ondelette à l'échelle j et à la position k.

Le *seuillage doux* consiste à remplacer chaque coefficient $w_{j,k}$ par la valeur $\tilde{w}_{j,k}$ où :

$$\tilde{w}_{j,k} = \begin{cases} sgn(w_{j,k})(\mid w_{j,k} \mid - T_j) & \text{si } \mid w_{j,k} \mid \geq T_j \\ 0 & \text{sinon} \end{cases}$$

Cette opération s'écrit en général :

$$\tilde{w}_{j,k} = \text{soft}(w_{j,k}) = sgn(w_{j,k})(\mid w_{j,k} \mid - T_j)_+ \tag{5.8}$$

où $(x)_+ = MAX(0, x)$.

Quand une transformation en ondelettes orthogonale discrète est utilisée, il est intérressant de noter que les estimateurs de seuillage dur et de seuillage doux sont les solutions respectives des problèmes de minimisations suivants :

$$\tilde{w} = \arg_w \min \frac{1}{2} \parallel D - \mathcal{W}^{-1}w \parallel^2_{l^2} + \lambda \parallel w \parallel^2_{l^0} \qquad \textbf{seuillage dur}$$

$$\tilde{w} = \arg_w \min \frac{1}{2} \parallel D - \mathcal{W}^{-1}w \parallel^2_{l^2} + \lambda \parallel w \parallel^2_{l^1} \qquad \textbf{seuillage doux}$$

où D sont les données d'entrée, $\mathcal{W}$ l'opérateur de transformation ondelette, et l^0 indique la limite de l^δ quand $\delta \rightarrow 0$. Ceci compte en fait le nombre d'éléments non nuls dans la séquence.

Comme décrit auparavant, dans le cas d'un bruit gaussien, $T_j = K\sigma_j$, où j est l'échelle du coefficient ondelette, σ_j est l'écart-type du bruit à l'échelle j, et K est une constant généralement choisi à 3. D'autres méthodes de seuillage ont également été proposées, commme le *le seuillage universel* (Donoho and Johnstone 1994; Donoho 1993), ou la méthode SURE (Stein Unbiased Risk Estimate) (Coifman and Donoho 1995), mais en général, elles ne donnent pas des résultats aussi bons que les méthodes basées sur le seuillage dur des coefficients significatifs. Pour les données astronomiques, le seuillage doux ne devrait jamais être utilisé car il cause une perte de la photométrie associée à chaque objet, ce qui peut être facilement vérifié en regardant une carte de résidu (c'est-à-dire une carte de la différence entre les données de départ et les données filtrées). En ce qui concernce le niveau de seuillage, le seuillage universel correspond au risque minimum. Plus le nombre de pixels est grand, plus le risque est grand, et il est normal que le seuil T dépende du nombre de pixels ($T = \sqrt{2 \log n}\sigma_j$, n étant le nombre de pixel). Le seuillage à $K\sigma$ correspond à un probabilité de fausse détection, c'est-à-dire la probabilité de détecter un coefficient comme significatif alors qu'il est dû au bruit. La valeur 3σ correspond à un taux de fausses détections de 0.27%.

La figure 5.1 montre les différentes possibilités de débruitage sur la sphère d'une carte de synchrotron simulée. En haut à gauche une simulation d'emission synchrotron renormalisée entre 0 et 255, en haut à droite, la même image à laquelle à été ajouté un bruit gaussien d'écart-type $\sigma = 5$. Sur ces données simulées, nous avons appliqué la méthode non-linéaire de débruitage par seuillage dans deux représentations différentes à savoir en ondelettes et en curvelets. Dans le premier cas, nous avons utilisé la transformation en ondelettes sphériques isotropes pyramidales décrites dans le troisième chapitre. Dans le deuxième cas, nous avons utilisé la transformation en curvelets pyramidale sur la sphère. La figure 5.1 montre les résultats obtenus dans les deux cas. Un seuil de détection à $5\sigma_j$ a été utilisé dans les deux expériences. Remarquons que les structures très anisotropes présentes dans les données initiales sont plus efficacement représentées par une forme d'onde elle même anisotrope de sorte que le filtrage dans la représentation en curvelets fournit de biens meilleurs résultats qu'un filtrage dans la representation en ondelettes isotropes utilisée.

5.3.2 La méthode de seuillage FDR

Le False Discovery Rate (FDR) est une méthode statistique (Benjamini and Hochberg 1995) qui permet de choisir un seuil de manière adaptative

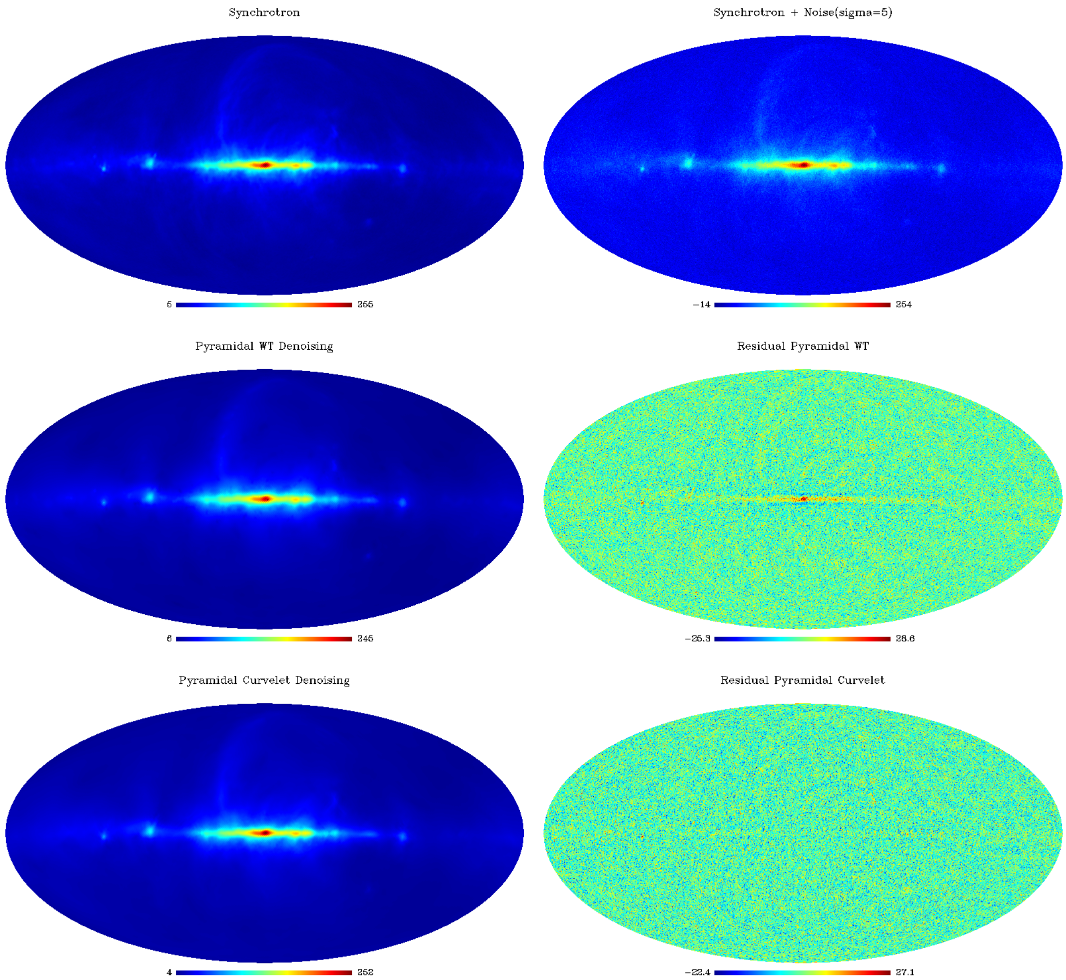

FIG. 5.1 – Débruitage, En haut à gauche et droite : simulation d'une image d'émission de synchroton galactique et la même image avec un bruit additif gaussien (i.e. données simulées). Au milieu : image filtrée par une transformation en ondelettes pyramidale et le résidu associé. En bas : image filtrée par une transformation en curvelets pyramidale et le résidu associé. Sur de telles données, présentant de nombreuses structures très anisotropes, le résidu de filtrage par transformation en curvelets est beaucoup plus faible que celui obtenu par débruitage dans une représentation en ondelettes.

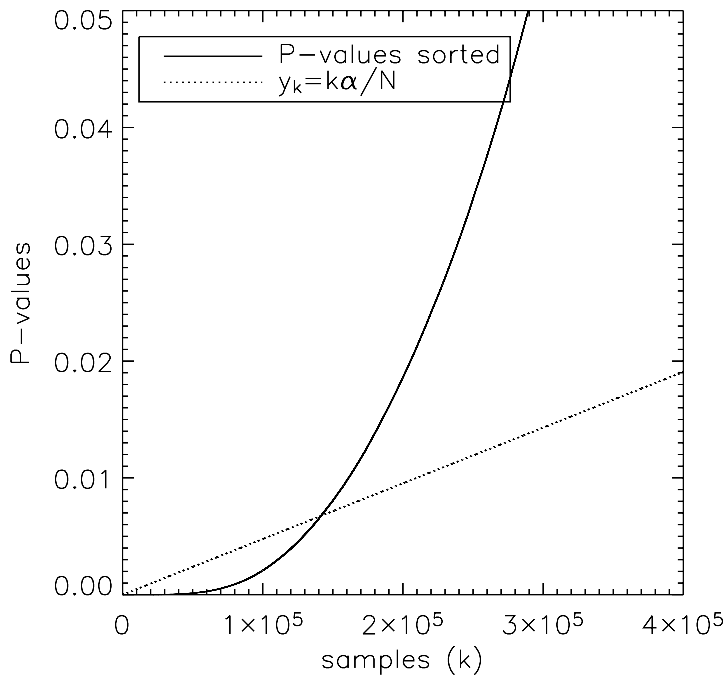

FIG. 5.2 – Détermination graphique du seuil FDR

et ainsi de déterminer le support multi-résolution. Cette technique a été utilisée récemment dans différents problèmes d'analyse de données en astrophysique (Miller et al. 2001; Hopkins et al. 2002; Starck et al. 2006; Pires et al. 2006). La procedure FDR fournit le moyen de contrôler de manière adaptative la part des fausses détections dans l'ensemble des objets détectés. L'intérêt d'une telle procédure s'agissant de débruitage d'images réside dans le rapprochement que l'on peut faire entre ce problème et celui du test d'hypothèses multiples : il faut pour chaque coefficient décider s'il doit ou non être conservé. Pour un seuil fixé τ, fixé *a priori* sur la base d'un test individuel, il est possible de montrer, que la probabilité d'avoir au moins une fausse détection, tend vers 1 lorsque le nombre de coefficients devient grand. La tendance naturelle serait à relever le seuil τ au risque d'être finalement trop conservatif. La méthode FDR propose de rechercher un compromis et permet de trouver un seuil de manière adaptative qui garantisse en moyenne un taux de fausses détections inférieur ou égal à la valeur α fixé par l'uilisateur.

Le FDR est donné par le rapport :

$$FDR = \frac{V_{ia}}{D_a} \tag{5.9}$$

où V_{ia} est le nombre de fausses détections et D_a est le nombre total de détections. Le formalisme FDR assure qu'en moyenne le taux de fausses détections n'est pas plus grand qu'un certain α pris entre 0 et 1 par l'utilisateur. Avec cette procédure, le paramètre ϵ discuté plus haut dans le paragraphe 5.2 est tel que :

$$\epsilon_{FDR} \leq \frac{T_i}{V}\alpha \leq \alpha \tag{5.10}$$

où le facteur inconnu $\frac{T_i}{V}$ est le rapport du nombre T_i de pixels vraiment inactifs au nombre total de pixels V.

La procédure FDR suit le schéma suivant :

- Soient $p_1, ..., p_N$ les p-values créées à partir de N tests, classées dans l'ordre croissant.

- On définit alors $d = max(k : p_k < \frac{k \cdot \alpha}{c_N \cdot N})$ où $c_N = 1$ si les p-values sont statistiquement indépendantes.

- Tous les pixels associés à des p-values inférieures ou égales à p_d sont considérés comme significatifs.

Graphiquement, ceci correspond à tracer la courbe p_k en fonction de $\frac{k}{N}$ ainsi que la droite passant par l'origine et de pente $\frac{\alpha}{c_N}$ (voir figure 5.2). On recherche alors le dernier point dont la p-value p_d est sous la droite. À partir de cette valeur p_d, on peut en déduire un seuil T. Tous les pixels ou coefficients plus grands que T ont une p-value plus petite que P_d et sont considérés comme significatifs.

Cette méthode a fait ses preuves en comparaison à d'autres méthodes traditionnelles pour fixer le seuil de détection, comme cela a été montré dans Starck et al. (2006). Dans le contexte de l'analyse de données sur la sphère, nous avons utilisé cette technique pour déterminer les seuils de manière adaptative pour les transformations multi-échelle que nous avons développées précédemment. Pour chaque échelle j, nous utilisons le seuil de détection T_j obtenu par la méthode FDR.

5.4 Méthode de filtrage combinée sur la sphère

Bien que les résultats obtenus par un simple filtrage des coefficients en curvelets soient très encourageants, il est possible d'améliorer encore ces derniers. Une étude rapide du résidu des filtrages, aussi bien en ondelettes que curvelets, sur la figure 5.1, montre l'éxistence de structures très différentes. Par exemple, le filtrage basé sur la transformation en ondelettes ne restitue pas correctement les structures longues et allongées, tandis que le filtrage à base de curvelets est mis à mal par des structures locales et isotropes. Chaque transformation a son propre domaine de prédilection où il est le plus efficace et la complémentarité des différentes transformations peut être mise à profit. La méthode de filtrage combinée (Combined Filtering Method–CFM) proposée dans Starck et al. (2001) permet de bénéficier des avantages des deux représentations simultanément. Cette méthode itérative détecte les coefficient significatifs alternativement dans le domaine ondelette et dans le domaine curvelet et garantit que la carte reconstruite prendra en compte tous les motifs detectés comme significatifs par chacune des transformations.

5.4.1 Principe

Supposons que nous ayons K transformations linéaires $T_1, \ldots, T_K$ et appelons α_k les coefficients d'un objet x après application de la transformation T_k, c'est à dire $\alpha_k = T_k x$. Nous allons considérer que pour chaque transformation T_k, nous avons une transformation inverse disponible que nous noterons T_k^{-1}, même s'il s'agit d'un abus de langage. Ainsi, on peut appeler T une matrice diagonale par bloc, construite à partir de T_k ; et α un vecteur construit à partir de l'ensemble des α_k.

La règle du seuillage dur associé à la transformation T_k fournit un estimateur $\tilde{s}_k$ à partir de la formule :

$$\tilde{s}_k = T_k^{-1} \delta(\alpha_k) \tag{5.11}$$

où δ est un opérateur qui met à zéro tous les coefficients α_k dont la valeur absolue est plus faible qu'un certain seuil predéfini, c'est-à-dire les coefficients non significatifs.

Pour des données y sous la forme $y = s + \sigma z$, où s est l'image que nous souhaitons restaurer et z un bruit blanc standard, nous allons résoudre le problème de minimisation suivant (Starck et al. 2001) :

$$\min \|T\tilde{s}\|_{\ell_1}, \quad \text{sous contrainte de} \quad s \in C, \tag{5.12}$$

où C est l'ensemble des vecteurs $\tilde{s}$ obéissant aux contraintes linéaires :

$$\begin{cases} \tilde{s} \geq 0, \\ |T\tilde{s} - Ty| \leq e; \end{cases} \tag{5.13}$$

Ici, la deuxième contrainte, l'inégalité, concerne seulement l'ensemble des coefficients significatifs, c'est-à-dire les indices μ tel que $\alpha_\mu = (Ty)_\mu$ soit plus grand (en valeur absolue) qu'un seuil t_μ. Un vecteur de tolérance (e_μ) étant donné, nous cherchons une solution $(T\tilde{s})_\mu$ à une distance au plus e_μ des coefficients bruités empiriques α_μ donnés par :

$$y = \langle y, \varphi_\mu \rangle,$$

de telle façon que α_μ suis une distribution normale de moyenne $\langle f, \varphi_\mu \rangle$ et de variance $\sigma_\mu^2 = \sigma^2 \|\varphi_\mu\|_2^2$. En pratique, les valeurs des seuils sont typiquement entre trois et quatre fois le niveau de bruit σ_μ. Dans nos expériences nous avons pris $e_\mu = \sigma_\mu/2$. En clair, nos contraintes garantissent que la reconstruction va prendre en compte tout motif détecté comme étant significatif par l'une au moins des K transformations.

Nous proposons comme méthode pour résoudre le problème de minimisation donnée par l'équation 5.12 d'utiliser l'algorithme *hybrid steepest descent* décrit par Yamada (2001). Cette technique est basée sur :

$$s^{n+1} = P(s^n) - \lambda_{n+1} \nabla_J (P(s^n)); \tag{5.14}$$

Ici, P est la l'opérateur de projection ℓ_2 sur l'ensemble C, ∇_J représente le gradient et $(\lambda_n)_{n\geq 1}$ est une suite satifaisant $(\lambda_n)_{n\geq 1} \in [0,1]$ et $\lim_{n\to+\infty} \lambda_n = 0$.

5.4.2 Algorithme

L'algorithme de filtrage combiné est le suivant :

1. Initialiser $L_{\max} = 1$, le nombre d'itérations N_i, et $\delta_\lambda = \frac{L_{\max}}{N_i}$.
2. Estimer l'écart-type du bruit σ et prendre $e_k = \frac{\sigma}{2}$.
3. Pour k = 1, .., K calculer la transformée : $\alpha_k^{(s)} = T_k s$.
4. Mettre $\lambda = L_{\max}$, $n = 0$, et $\tilde{s}^n$ à 0.
5. Tant que $\lambda >= 0$,
   - $u = \tilde{s}^n$.
   - Pour k = 1, .., K,
     - Calculer la transformée $\alpha_k = T_k u$.
     - Pour tous les coefficients $\alpha_{k,l}$ faire
       - Calculer le résidu $r_{k,l} = \alpha_{k,l}^{(s)} - \alpha_{k,l}$
       - si $\alpha_{k,l}^{(s)}$ est significatif et $\mid r_{k,l} \mid > e_{k,l}$ alors $\alpha_{k,l} = \alpha_{k,l}^{(s)}$
       - $\alpha_{k,l} = sgn(\alpha_{k,l})(\mid \alpha_{k,l} \mid - \lambda)_+$.
     - $u = T_k^{-1} \alpha_k$
   - Seuiller les valeurs négatives dans u et $\tilde{s}^{n+1} = u$.
   - $n = n + 1$, $\lambda = \lambda - \delta_\lambda$, et retourner à l'étape 5.

5.4.3 Résultats

Figure 5.3 montre le résultat du débruitage d'une carte de synchrotron par la méthode de filtrage combinée et la carte résiduelle associée.

La Figure 5.4 fait un zoom sur une face Healpix (la face 6) des données et résultats de cette même expérience permettant une comparaison plus fine des trois techniques de filtrage proposées. Du fait de l'existence de formes de nature très différentes dans les données, les résidus après filtrage en ondelettes ou en curvelets exhibent encore des structures. Le résultat est bien meilleur quand le filtrage combiné est appliqué, et aucune structure ne peut plus être detecté à l'oeil nu dans le résidu.

Les résultats en termes de *rapport signal-sur-bruit*, synthétisés dans le tableau 5.1, corroborent l'analyse visuelle des résidus.

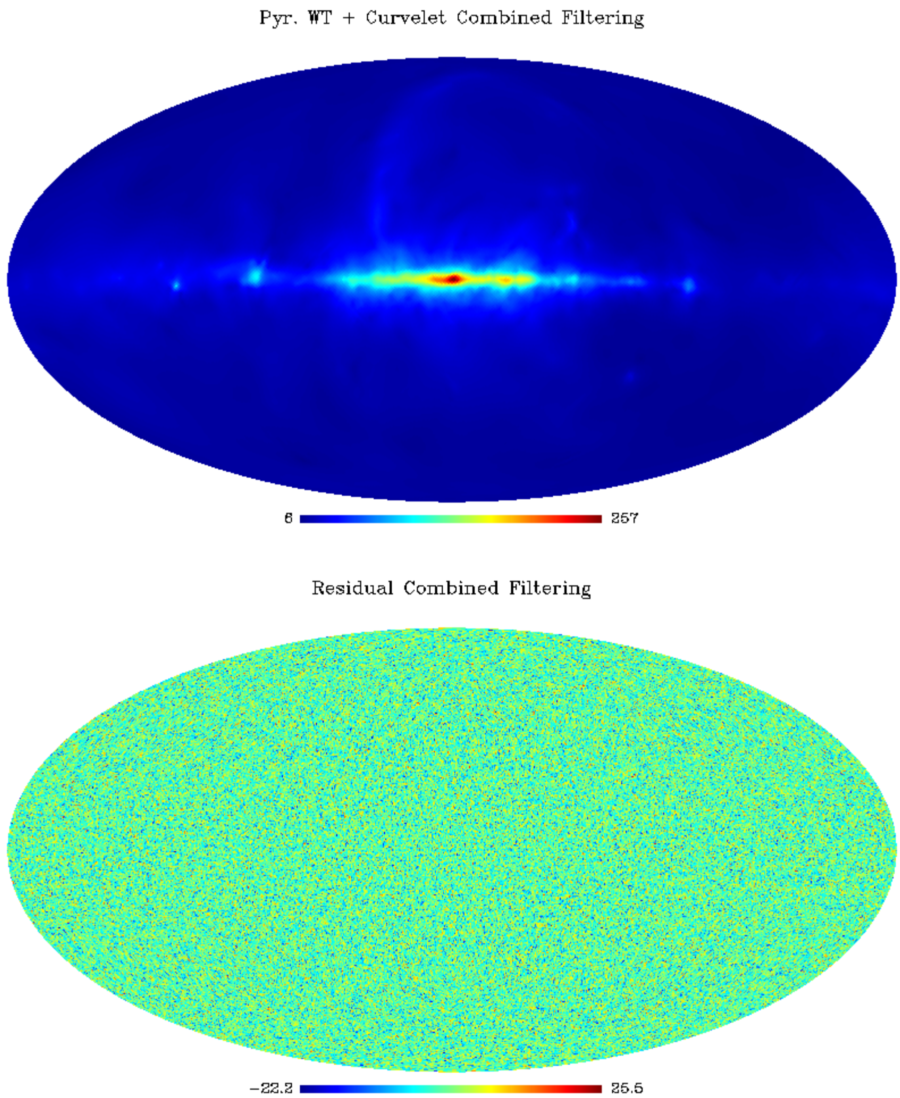

FIG. 5.3 – **Débruitage.** Carte de synchrotron débruitée obtenue par application de la méthode de filtrage combiné utilisant les transformations pyramidales en ondelettes et en curvelets, et carte du résidu associé.

Méthode	Ecart-type	SNR (dB)
Carte bruitée	5.	13.65
Filtrage ondelettes	1.30	25.29
Filtrage curvelets	1.01	27.60
Filtrage combiné (CFM)	0.86	28.99

TAB. 5.1 – Table des écart-type et des rapports signal-sur-bruit (SNR) après filtrage de la carte de synchrotron, initialement bruitée par un bruit blanc gaussien d'écart-type $\sigma = 5$, par la méthode de filtrage non-linéaire décrite plus haut, appliquée dans la représentation en ondelettes, en curvelets ainsi que par une méthode de filtrage non-linéaire itérative dans une représentation combinée en ondelettes et curvelets.

5.5 Conclusion

Cette application montre l'apport direct de ces nouvelles transformations sur la sphère. Il est désormais possible de filtrer et de débruiter des images sur la sphère sans devoir passer par des projection sur le plan. Dans le chapitre suivante, nous allons voir une autre application de ces nouvelles transformées multi-échelles : l'interpolation des données manquantes sur la sphère.

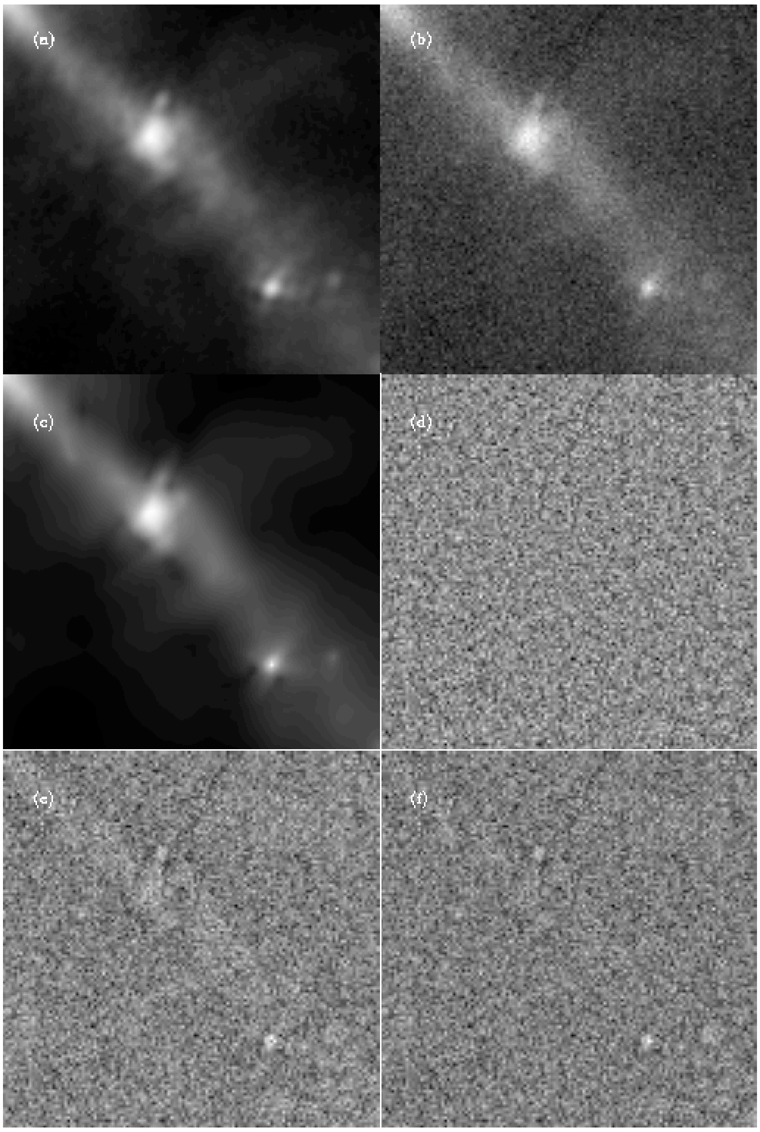

FIG. 5.4 – **Méthode de filtrage combinée** : Représentation de la sixième face dans le schéma de pixelisation Healpix de l'image présentée à la figure 5.3. De haut en bas et de gauche à droite : a) l'image originale, b) l'image bruitée, c) l'image filtrée par la méthode combinée , d) résidu de l'image filtrée par le méthode combinée, e) résidu du filtrage ondelette et f) résidu du filtrage curvelet.

Chapitre 6

Analyse en Composantes Morphologiques et Inpainting sur la Sphère

Sommaire

6.1 MCA sur la sphère **82**
6.1.1 Représentations parcimonieuses 82
6.1.2 MCA : Principe et Algorithme 85
6.1.3 Expériences numériques et applications 88
6.2 Inpainting sur la sphère **93**
6.2.1 Principe et Algorithme 94
6.2.2 Experiences Numériques - Application en Astrophysique 97
6.3 Conclusion **104**

Nous proposons dans ce chapitre d'étendre au cas de données sur la sphère les algorithmes d'Analyse en Composantes Morphologiques (MCA) et d'*Inpainting* introduits initialement dans Elad et al. (2006); Starck et al. (2004b) pour l'analyse des signaux et des images usuels. Ces extensions ont été rendues possibles par la variété des transformations et représentations de données sur la sphère développées récemment et en particulier les ondelettes et curvelets sur la sphère construites aux chapitres et .
L'algorithme MCA a été conçu pour séparer les différents constituants d'une image, tels que les contours d'une part et les régions texturées d'autre part, en exploitant leur diversité morphologique. Si on considère la concision d'une représentation des données comme quelque chose d'important, la diversité morphologique est simplement le constat que différents choix de base conduisent à des représentations plus ou moins concises de différentes familles de signaux. La théorie de l'approximation établit des résultats théori-

ques s'agissant de l'optimalité en termes de parcimonie d'une représentation pour une classe donnée de signaux (Candès and Donoho 1999c, 2000).
Lorsque le signal étudié présente des caractéristiques morphologiques diverses il est tentant de chercher à construire une représentation parcimonieuse dans un dictionnaire plus vaste et redondant, formé de l'union de plusieurs bases, chacune étant *a priori* bien adaptée à l'une des composantes. On en arrive alors à rechercher la solution la plus parcimonieuse d'un problème inverse linéaire sous-déterminé. L'algorithme MCA permet dans ce cadre d'atteindre rapidement une solution approchée, largement satisfaisante dans de nombreuses applications. Il en est ainsi de l'interpolation de données manquantes ou *inpainting* qui peut également être vu comme un problème inverse linéaire sous déterminé. Une variante de l'algorithme MCA proposée dans Starck et al. (2004b); Elad et al. (2006) s'appuie sur l'existence d'une représentation parcimonieuse des données complètes dans un dictionnaire spécifié *a priori* pour reconstruire les structures déteriorées.

Les paragraphes suivants décrivent plus en détails l'extension de l'algorithme MCA et de sa variante pour l'interpolation de données manquantes dans le cas où les données à analyser ou à interpoler sont définies sur la sphère. Différents résultats expérimentaux sont rapportés illustrant la pertinence des méthodes proposées, en particulier dans le contexte de la préparation à l'analyse des données du fond diffus cosmologique (CMB) que fournira le satellite Planck.

6.1 MCA sur la sphère

6.1.1 Représentations parcimonieuses

Une pratique courante du traitement du signal, des images voire de cartes sur la sphère, est de décomposer les données en constituants élémentaires. Cette *analyse* s'exprime comme un problème inverse dans lequel les données sont supposées avoir été générées selon le modèle suivant :

$$y = \sum_i \alpha_i \phi_i + \eta \tag{6.1}$$

c'est à dire une combinaison linéaire de fonctions élémentaires $\phi_i \in \mathbb{R}^n$ avec les poids α_i. Dans ce modèle, η représente une possible contamination par un bruit additif, typiquement un bruit blanc gaussien. Le problème est de retrouver les structures qui participent à la formation des données $y \in \mathbb{R}^n$, c'est à dire de trouver un ensemble de fonctions ϕ_i et les poids associés $\tilde{\alpha}_i$ qui permettent de former les données. La solution de ce problème d'estimation dépend à l'evidence fortement de l'information disponible *a priori*. Nous supposons ici que les *bonnes* fonctions élementaires ϕ_i doivent être recherchées au sein d'un ensemble de fonctions défini *a priori*. Cet ensemble

peut-être une base, une trame ou encore plusieurs bases ou trames réunies en un dictionnaire volumineux et redondant.
En pratique, s'agissant d'analyser des données sur la sphère, les transformations *numériques* disponibles incluent la transformation en harmoniques sphériques et différentes transformations en ondelettes. Ainsi, des logiciels tels que Healpix[1] (Górski et al. 2005) ou Glesp (Doroshkevich et al. 2005) fournissent des routines pour le calcul des harmoniques sphériques d'une fonction échantillonnée sur la sphère selon leurs schémas de pixelisation respectifs. Mentionnons les travaux de Schröder and Sweldens (1995) qui ont développé une transformation orthogonale en ondelettes sur la sphère inspirée de la transformation de Haar, ainsi que les travaux de Freeden and Maier (2002); Freeden et al. (2003) qui ont proposé une transformation en ondelettes inversible fondée sur la transformée en harmoniques sphériques. Sur la base de ces travaux, nous avons décrit aux chapitres 3 et 4 plusieurs nouvelles transformations multi-échelles inversibles sur la sphère à savoir une transformation en ondelettes isotrope non décimée sur la sphère, une transformation en ondelettes pyramidale, une transformation en ridgelets ainsi qu'une transformation en curvelets sur la sphère (Starck et al. 2006). Notre intérêt ici est que chacune de ces transformations est à même de représenter avec parcimonie différents types de structures, de morphologies, dans les données. Les ondelettes détecteront facilement des objets localisés et isotropes tandis que les curvelets sont davantage adaptées à la détection de structures anisotropes.
Il est clair que tout $y \in \mathbb{R}^n$ admet une représentation exacte dans n'importe quelle base de $\mathbb{R}^n$, ou même plusieurs telles représentations exactes dans le cas d'un dictionnaire redondant. Néanmoins, toutes ces représentations ne sont pas également intéressantes s'agissant de modèlisation, de détection de structures, etc. En pratique, il y a une tendance forte à préférer *a priori* les représentations des données y qui n'utilisent qu'un petit nombre de fonctions élémentaires aboutissant à une représentation plus concise et peut-être plus interprétable des données. La construction de représentations ou d'approximations parcimonieuses est en fait l'art et le cœur de l'analyse de données structurées. La conception d'algorithmes performants de détection, de débruitage, de restauration, de compression s'appuie sur l'existence de dictionnaires adaptées et sur l'existence d'algorithmes efficaces pour la sélection d'éléments dans ces dictionnaires.
Sur ce dernier point, soulignons en effet que trouver le plus petit ensemble d'éléments dans un dictionnaire étendu, tels que leurs combinaisons permettent de reconstruire les structures les plus saillantes d'un signal ou d'une carte donnée, ce qu'on peut exprimer formellement par :

$$\min_{\alpha} \|\alpha\|_{\ell_0} \text{ sous contrainte de } y = \Phi\alpha \tag{6.2}$$

[1]http ://www.eso.org/science/healpix

où la pseudo-norme ℓ_0 comptabilise le nombre de coefficients non-nuls, est un problème combinatoire NP-complexe, le plus souvent insoluble. De manière surprenante, des résultats récents (Bruckstein and Elad 2002; Donoho and Elad 2003a; Donoho and Huo 2001; Donoho 2006a) montrent que lorsque y a une representation suffisamment parcimonieuse dans le dictionnaire Φ, la solution du problème ci-dessus est unique et coincide avec celle du problème suivant :

$$\min_{\alpha} \ \|\alpha\|_{\ell_1} \text{ sous contrainte de } y = \Phi\alpha. \tag{6.3}$$

qui à l'avantage d'être convexe du fait que la parcimonie d'un vecteur de $\mathbb{R}^n$ y est mesurée avec une norme ℓ_1. Un grand nombre d'algorithmes ont été proposés pour essayer de resoudre (6.2) directement ou (6.3) ou encore une version relaxée de ce dernier problème, éventuellement sous forme lagrangienne :

$$\min_{\alpha} \frac{1}{2}\|y - \Phi\alpha\|_{\ell_2}^2 + \lambda\|\alpha\|_{\ell_1} \quad \lambda > 0 \tag{6.4}$$

parmi lesquels l'algorithme *Matching Pursuit* (Mallat and Zhang 1993), l'algorithme de résolution d'un programme linéaire *Basis Pursuit* (Chen et al. 1998) basée sur la Méthode du point intérieur, ou encore les algorithmes LARS (Efron et al. 2004), Stomp (Donoho et al. 2006b) et *Polytope Faces Pursuit* (Plumbley 2006). Un algorithme de descente pour rédoudre (6.4) est décrit dans Alliney (1992); Alliney and Ruzinsky (1994). Les conditions pour lesquelles chacune de ces méthodes fournit une solution parcimonieuse unique et les conditions pour que ces solutions coincident avec la solution optimale au problème (6.2) ont été étudiées récemment par de nombreux auteurs (e.g. Donoho and Elad 2003a; Elad and Bruckstein 2002; Fuchs 2005; Gribonval and Nielsen 2003; Donoho et al. 2006a). Ils ont montré que les méthodes proposées sont effectivement capables de recouvrer la solution la plus parcimonieuse à condition que celle-ci soit effectivement suffisamment parcimonieuse et que le dictionaire Φ soit suffisamment décohérent. Rappelons que la cohérence d'un dictionnaire est une mesure de la ressemblance entre ses éléments au sens de la valeur absolue de leur produit scalaire. Dans une analyse *pire cas*, la cohérence est définie comme le plus grand terme hors-diagonale de la matrice de Gram du dictionnaire Φ, en valeur absolue. Des analyses plus fines utilisant des notions de cohérence plus subtiles ont été menées par exemple dans Tropp (2006).
S'agissant d'analyser des grands jeux de données tels que des images ou des cartes sphériques, qui plus est à la résolution du satellite Planck, le coût calculatoire des algorithmes précédents peut rapidement devenir rédhibitoire. L'Analyse en Composantes Morphologiques (MCA) est une alternative récente et bien plus rapide décrite dans Starck et al. (2004b). Le prochain paragraphe détaille cet algorithme en l'étendant à l'analyse de données sur la sphère.

FIG. 6.1 – Photo de Barbara séparée en contours et textures en utilisant MCA avec un dictionnaire de cosinus locaux et de curvelets.

6.1.2 MCA : Principe et Algorithme

L'Analyse en Composantes Morphologiques (MCA) construit une representation parcimonieuse d'un signal ou d'une image, ou d'une carte sur la sphère en considérant que ce signal ou cette image est une combinaison linéaire de composantes morphologiquement différentes, chacune admettant une représentation parcimonieuse dans différents dictionnaires pour lesquel il existe des algorithmes de transformation et de reconstruction rapides. Il en est ainsi par exemple de l'image de la figure 6.1 qui combine des régions texturées séparées par des contours : les premières sont bien représentées par les éléments d'une transformation en cosinus local tandis que les seconds sont bien détectés par des curvelets.
L'exemple de la figure 6.2 est plus évident encore : l'image initiale y peut-être vue comme la somme d'une image ne contenant que des segments de droite s_1 et d'une image ne contenant que des structures *ponctuelles* isotropes s_2. Avec, l'algorithme MCA, il est possible de construire une représentation parcimonieuse de x dans un dictionnaire redondant formé de ridgelets et d'ondelettes et par là même de séparer s_1 et s_2. La diversité morphologique est ce qui permet cette séparation.
Plus générallement, étant donnée une carte sur la sphère y modélisée comme une combinaison linéaire de K composantes s_k à estimer,

$$y = \sum_{k=1}^{K} \alpha_k s_k \tag{6.5}$$

nous faisons pour MCA l'hypothèse qu'il existe pour chaque k un dictionnaire éventuellement redondant $\mathbf{\Phi}_k$ tel que s_k admette une représentation parcimonieuse α_k dans $\mathbf{\Phi}_k$ avec $s_k = \mathbf{\Phi}_k \alpha_k$. On suppose de surcroît que la décomposition la plus éparse de s_k dans un dictionnaire $\mathbf{\Phi}_{k' \neq k}$ est en revanche très peu parcimonieuse. Formellement, cette diversité morphologique

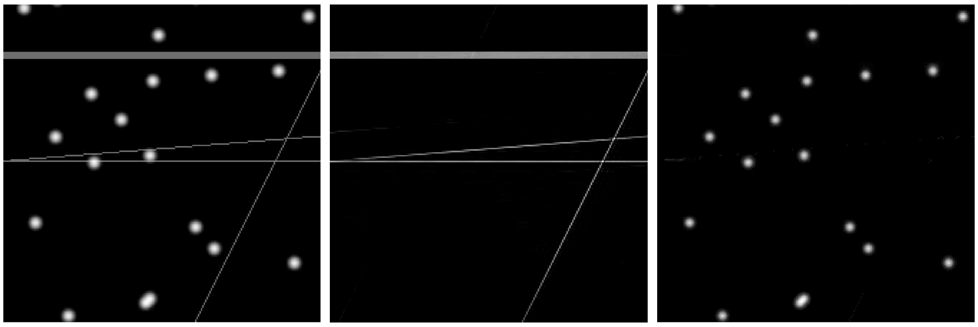

FIG. 6.2 – Lignes et sources ponctuelles séparées en utilisant MCA avec un dictionnaire d'ondelettes et de ridgelets.

des s_k requise par MCA, s'exprime simplement de la manière suivante :

$$\forall k, \quad \forall k' \neq k, \quad ||\mathbf{\Phi}_k^T s_k||_0 < ||\mathbf{\Phi}_{k'}^T s_k||_0 \tag{6.6}$$

dans le cas où les $\mathbf{\Phi}_k$ sont des bases orthogonales. Ainsi, les différents $\mathbf{\Phi}_k$ peuvent être vus comme amplifiant les différences perçues entre les différentes composantes et MCA s'appuie sur le pouvoir discriminant des différents dictionnaires $\mathbf{\Phi}_k$ pour estimer et donc séparer les composantes $(s_k)_{k=1,\cdots,K}$ à partir du mélange observé y.
Idéalement, les α_k sont les solutions du problème suivant :

$$\min_{\alpha_1,\ldots,\ \alpha_K} \sum_{k=1}^{K} \|\alpha_k\|_0 \quad \text{subject to} \quad y = \sum_{k=1}^{K} \mathbf{\Phi}_k \alpha_k \tag{6.7}$$

Néanmoins, comme déjà indiqué plus haut et bien que sensé du point de vue de la solution recherchée, ce dernier problème est non convexe et combinatoire. Sa complexité augmente exponentiellement avec le nombre d'éléments dans le dictionnaire dans son ensemble. Finalement, compte tenu des résultats d'équivalence obtenus récemment dans Donoho and Elad (2003a,b) par exemple, l'algorithme MCA recherche plutôt une solution au problème de minimisation suivant :

$$\min_{s_1,\ldots,s_K} \lambda \sum_{k=1}^{K} \|\alpha_k\|_1 + \left\| y - \sum_{k=1}^{K} s_k \right\|_2^2 \quad \text{avec} \;\; s_k = \mathbf{\Phi}_k \alpha_k. \tag{6.8}$$

où une norme de parcimonie ℓ_1 est substitué à la norme ℓ_0 suivant une prescription de l'algorithme *Basis Pursuit* décrit dans Chen et al. (1998). Dans cette nouvelle formulation, la contrainte d'égalité a été relaxée faisant apparaitre un terme quadratique permettant la prise en compte éventuelle d'une erreur de modélisation ou d'un bruit additif gaussien sur la carte observée y. On a toujours :

$$\forall k, \quad s_k = \mathbf{\Phi}_k \alpha_k \tag{6.9}$$

Finalement, après ces modifications, le problème de minimisation (6.8) est convexe. Nous proposons de le résoudre par *blocs* (Bruce et al. 1998), composante par composante, en exigeant qu'à la solution le gradient partiel relativement à chaque s_k soit nul. Dans le cas où chaque $\mathbf{\Phi}_k$ est une base orthonormée, cette approche conduit au système d'équations couplées suivants :

$$\forall k, s_k = r_k - \frac{\lambda_k}{2}\mathbf{\Phi}_k \mathrm{sign}(\mathbf{\Phi}_k^T s_k) \text{ with } r_k = s - \sum_{k' \neq k} s_{k'} \tag{6.10}$$

L'algorithme MCA cherche à résoudre numériquement ce système d'équations où en tous les cas à déterminer une solution approchée satisfaisante. MCA consiste en un seuillage itératif, alternant sur les composantes s_k, avec un seuil décroissant.

Algorithme

MCA construit une approximation parcimonieuse des données y en raffinant progressivement l'approximation courante en incorporant des détails plus fins alternativement dans chacune des composantes morphologiques s_k, au fur et à mesure des itérations. En effet, à chaque itération, un opérateur de seuillage ou filtrage non-linéaire est utilisé pour détecter des structures dans le résidu courant. Le seuil étant abaissé progressivement au cours des itérations, les structures détectées sont de plus en plus fines. Finalement, nous proposons d'approcher une solution du système d'équations couplées 6.10 ci-dessus en utilisant la méthode itérative *Block-Coordinate Relaxation Method* (Bruce et al. 1998), avec, pour un k donné, un seuillage doux de la décomposition de r_k sur $\mathbf{\Phi_k}$.
Soulignons que lorsque les transformations utilisées sont non-unitaires ou redondantes, l'approche proposée n'est plus strictement valable. Néanmoins, des études théoriques récentes (M.Elad 2006; Chaux et al. 2007; Elad et al. 2007) permettent d'envisager une extensions au cas de trames ou de dictionnaires redondants. En notant $\mathbf{T}_k$ et $\mathbf{R}_k$ les transformations directes et inverses associées au dictionnaire redondant $\mathbf{\Phi_k}$, MCA recherche en fin de compte une solution au problème (6.8) avec l'algorithme suivant :

1. Fixer le nombre d'itérations $I_{\max}$ et la valeur initiale du seuil $\left(\lambda_k^{(0)}\right)_k$
2. Tant que $\lambda_k^{(t)} > \lambda_{\min}$ ($\lambda_{\min}$ égale à zéro ou dépendant du niveau de bruit),
   – Réaliser les itérations suivantes pour estimer les composantes $(s_k)_{k=1,\dots,K}$ à l'itération t :
   Pour $k = 1, \cdots, K$
   - Calculer le terme résiduel $r_k^{(t)}$ en supposant l'estimation $\tilde{s}_{k' \neq k}^{(t-1)}$ de $s_{k' \neq k}$:
     $r_k^{(t)} = y - \sum_{k' \neq k} \tilde{s}_{k'}^{(t-1)}$
   - Estimer les coefficients à l'itération t de $\tilde{s}_k^{(t)}$ en seuillant avec $\lambda_k^{(t)}$:
     $\tilde{\alpha}_k^{(t)} = \delta_{\lambda_k^{(t)}}\left(\mathbf{T}_k r_k^{(t)}\right)$
   - Reconstruire une estimation de s_k en reconstruisant à partir des coefficients sélectionnées $\tilde{\alpha}_k^{(t)}$:
     $\tilde{s}_k^{(t)} = \mathbf{R}_k \tilde{\alpha}_k^{(t)}$

   – Faire décroître le seuil λ_k en suivant une stratégie donnée.

L'opérateur de seuillage doux apparait dans ce qui précède comme une conséquence de l'utilisation de la norme ℓ_1 comme mesure de parcimonie. Vers la fin du processus itératif, l'utilisation d'un opérateur de seuillage dur, qu'on peut relier à la norme ℓ_0, à la place peut conduire à de meilleurs résultats (*e.g.* en termes d'erreur de reconstruction) (Starck et al. 2004b).
L'adoption d'un seuil λ décroissant au fur et à mesure des itérations est une caractéristique importante de l'algorithme proposé par Starck et al. (2004b). En l'absence de bruit sur les données, le seuil est progressivement abaissé jusqu'à atteindre une valeur finale nulle. Cette approche constructive a quelques ressemblances avec les algorithmes dits *greedy* tels que *Matching Pursuit*, *Orthogonal Matching Pursuit* (Mallat and Zhang 1993) ou encore la variante *stagewise* de ce dernier algorithme *STOMP* (Donoho et al. 2006b). Le seuil a néanmoins autant une valeur algorithmique que physique lorsque les données sont entachées de bruit et que l'équation 6.8 est interprétée dans un cadre probabiliste : la valeur finale du seuil est dans ce dernier cas liée au bruit sur les données ramené dans la représentation des composantes. Arrêter la reconstruction des composantes dès lors que les coefficients détectés ne sont plus significativement au-dessus du niveau de bruit. On prend typiquement une valeur finale du seuil égale à trois fois l'écart-type du bruit.
Le choix d'une stratégie de décroissance du seuil λ n'est pas anodin. En effet, un choix judicieux permet d'accélérer la vitesse de convergence de l'algorithme, et donc de diminuer le temps de calcul nécessaire à l'algorithme MCA, sans pour autant dégrader la qualité de l'estimation des composantes. Différentes stratégies, selon les distributions *a priori* du bruit et des composantes (*e.g.* décroissance linéaire ou exponentielle, adaptative ou pas) sont étudiées dans Bobin et al. (2007b).

6.1.3 Expériences numériques et applications

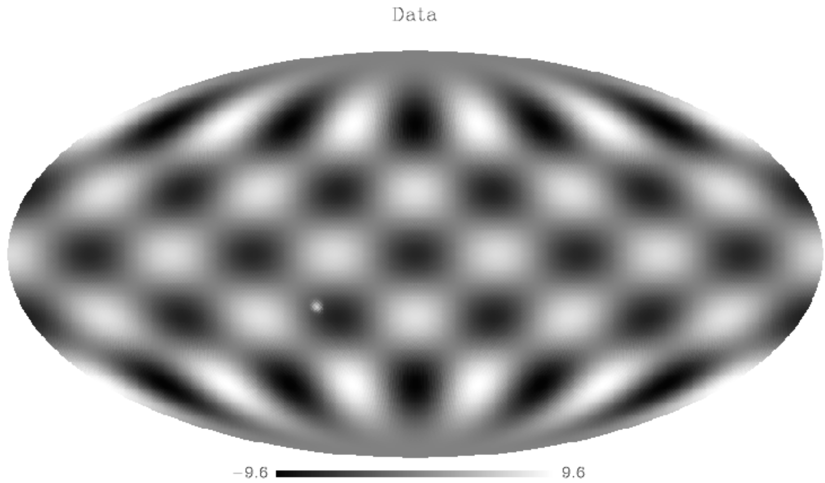

FIG. 6.3 – Séparation d'une harmonique sphérique et d'un profil gaussien avec MCA : données synthétiques sur la sphère, en entrée de l'algorithme.

Expérience numérique de principe

Les figures 6.3 et 6.4 reproduisent les cartes sur la sphère en entrée et en sortie d'une expérience de principe utilisant l'algorithme MCA détaillé plus haut. La figure 6.3 représente la carte d'entrée : il s'agit d'une simulation obtenue en sommant deux composantes admettant visiblement des représentations *parcimonieuses* dans deux dictionnaires très différents. La première composante est la représentation spatiale sur la sphère d'une harmonique sphérique obtenue avec l'inverse de la transformation en harmoniques spheriques de *Healpix*. La deuxième composante est parcimonieuse dans la représentation en ondelettes isotropes sur la sphère. Il s'agit d'un profil gaussien 2D projeté sur la sphère selon l'application point à point entre les pixels d'une des 12 faces de la pixélisation Healpix et ceux d'une image plate de même taille. L'algorithme MCA, utilisé avec la transformation en harmoniques sphériques et la transformation en ondelettes isotropes sur la sphère, est sans surprise capable de retrouver ces deux composantes. Les résultats de cette séparation sont reproduits à la figure 6.4. Cette expérience de principe ouvre la voie à des applications plus conséquentes telle que celle décrite dans le paragraphe suivant.

Application à la détection de défauts de surface

Nous décrivons ici succinctement une application de MCA, détaillée dans Abrial et al. (2007), pour le contrôle de la sphéricité de cibles utilisées dans la fusion inertielle par confinement (ICF). Cette technologie compte sur l'implosion de cibles sphépriques multicouches complexes pour générer

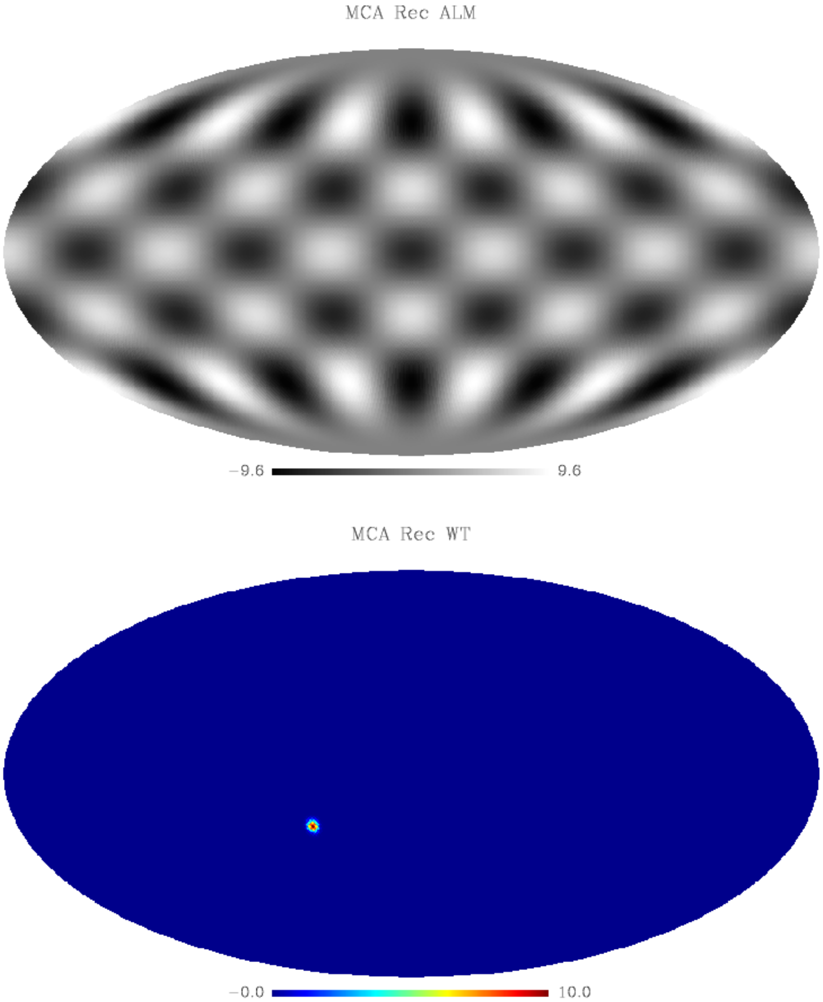

FIG. 6.4 – Séparation d'une harmonique sphérique et d'un profil gaussien avec MCA : cartes sur la sphère des composantes séparées.

en leur centre les conditions de température et de pression nécessaires à l'amorçage de la réaction de fusion nucléaire (Atzeni and ter Vehn 2004). L'existence de défauts sur l'une ou l'autre couche d'une cible idéalement parfaitement sphérique, peut être à l'origine du développement d'instabilités hydrodynamiques oblitérant la formation de points chauds au sein du combustible nucléaire au centre de la cible et donc sa mise à feu (Lindl 1997). Il est donc capital de pouvoir caractériser de l'état de surface d'une cible et juger statistiquement de la qualité d'un processus de fabrication.
La surface d'une cible sphérique peut être imagée à l'échelle nanométrique par différentes techniques : microscopie électronique, interférométrie optique,

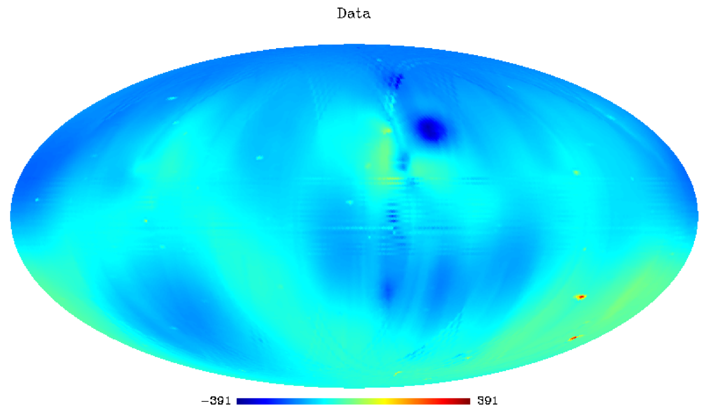

FIG. 6.5 – Structures en surface d'une capsule sphérique utilisée pour la fusion à confinement inertiel, à l'échelle nanométrique. On observe la superposition de déformations à grande échelle, des pics isolés et des *rayures* ainsi que des artefacts qui ressemblent à des interférences sur des échelles spatiales intermédiaires.

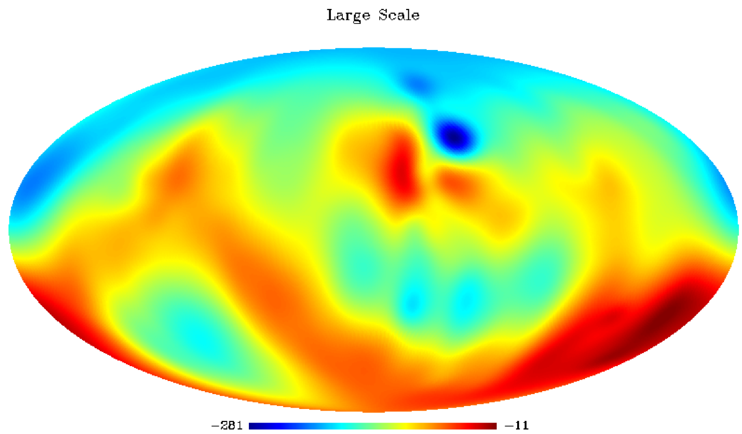

FIG. 6.6 – Plan lissé de la transformation ondelette isotrope non-décimée sur la sphère de la carte des défauts surfaciques d'une cible reproduite à la figure 6.5.

etc. La figure 6.6 reproduit un exemple de carte sphérique résultant de telles mesures. Il apparait des écarts à la sphère idéale avec des structures de morphologies variées : des variations lentes à grande échelle sur la sphère, des pics isolés, des rayures ainsi que des artefacts qui ressemblent à des interférences

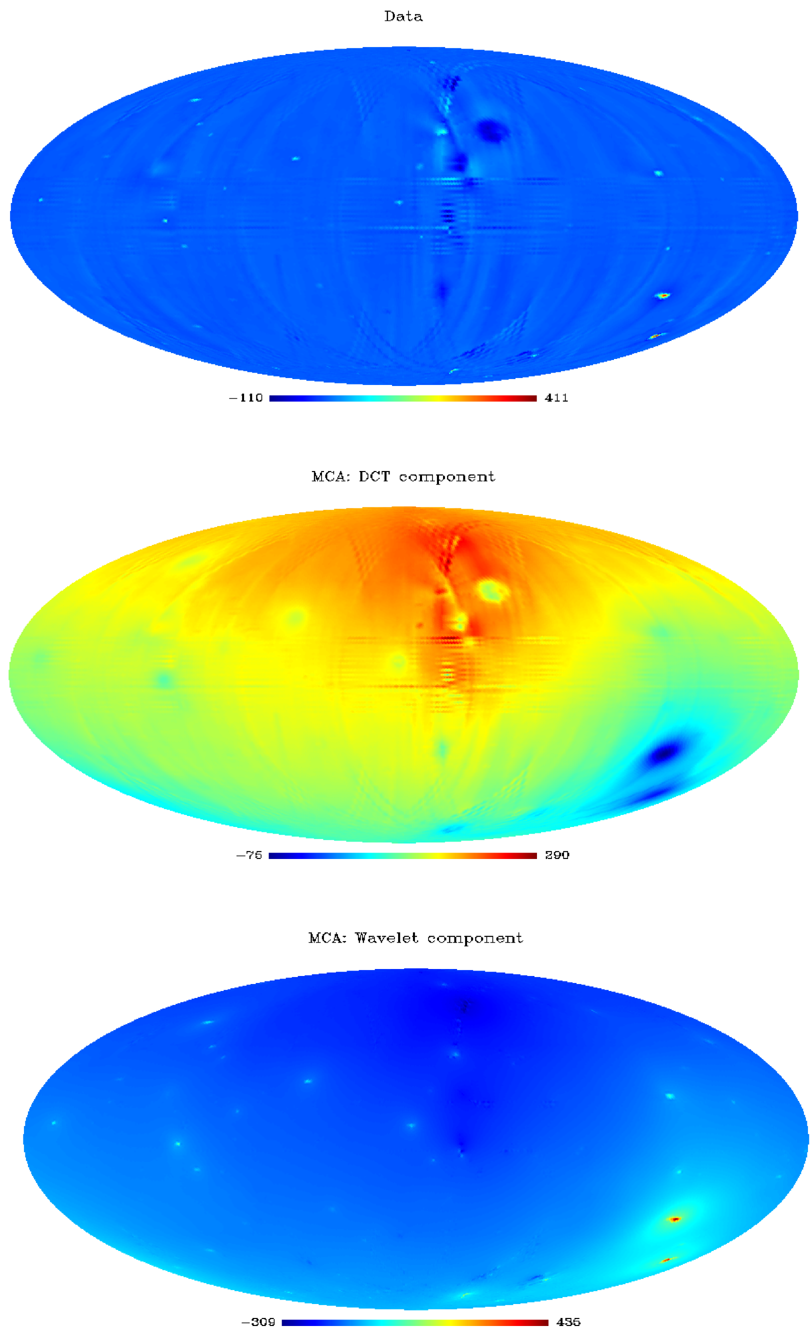

Fig. 6.7 – **En haut :** Carte sphérique obtenue à partir des données initiales de la figure 6.5 après soustraction du plan lissé représenté à la figure 6.6 à laquelle l'algorithme MCA a été appliqué. **Au centre :** Carte des motifs d'interférence et les artefacts de mesure captés par la tranformation en cosinus locale sur la sphère. **En bas :** Carte des pics isolés détectés par la transformation en ondelettes isotropes non-décimée sur la sphère. La somme du plan lissé (figure 6.6) et de cette dernière carte permet d'obtenir des données nettoyées des motifs d'interférence et des artefacts d'observation.

sur des échelles spatiales intermédiaires. Ces artefacts n'entrent évidemment pas en ligne de compte s'agissant de qualifier une capsule pour l'implosion. L'algorithme MCA s'est avéré utile pour analyser les petites échelles des données de la figure 6.5, les grandes échelles dans une représentation en ondelettes isotropes ayant été préalablement soustraites. Les dictionnaires que nous avons utilisés sont la transformation en ondelettes isotropes non-décimée sur la sphère et la transformation locale en cosinus, la première préssentie pour représenter au mieux les pics isolés et la seconde pour representer les artefacts oscillants dus à la technique d'observation. Avec cette méthode d'analyse, il est possible de reconstruire une carte nettoyée des artefacts en recombinant la carte des grandes échelles et la carte des pics isolés, telle reproduite à la figure 6.7. L'étude de l'évolution des cartes nettoyées au moyen d'un code de simulation hydrodynamique permet ensuite de révéler d'éventuels défauts systématiques dans le processus de fabrication des cibles. Davantages de détails sur cette étude sont accessibles dans Abrial et al. (2007); Afeyan et al. (2006).

6.2 Inpainting sur la sphère

Appelé ainsi d'après les procédés utilisés dans la restauration d'œuvres d'art détériorées, l'*inpainting* fait référence à un ensemble de techniques utilisées pour modifier des images de façon indécelable pour des personnes qui n'auraient pas eu connaissance des images originales (voir figure 6.8). Il y a de nombreuses applications telles que la suppression de rayures ou d'objets entiers dans des photographies, la suppression de texte ou graphiques surimprimés, le remplissage de blocs manquants dans des images transmises de manière incomplète, ou encore l'interpolation des structures dans une image en vue d'obtenir de meilleurs taux de compression, etc. Très généralement, l'*inpainting* tente d'interpoler à travers les trous dans une image en s'appuyant sur l'*information* portée par l'arrangement des pixels encore disponibles : il s'agit de prolonger les contours, d'exploiter la periodicité des textures apparentes, etc.

Récemment, le développement d'algorithmes pour la préservation des contours et des textures au travers de trous dans les données a suscité un vif intérêt, et de nombreux chercheurs ont contribué à la résolution de ce problème d'interpolation. Depuis la contribution remarquable de Masnou and Morel (1998, 2002), qui ont proposé des méthodes de désocclusion basées sur des principes variationnels, l'*inpainting* d'images non-texturées à été particulièrement étudié. Plus récemment, les contributions de Ballester et al. (2001); Bertalmio et al. (2001, 2000), puis de Chan and Shen (2001) ont donné une nouvelle impulsion au développement de techniques numériques d'*inpainting*. Dans ces travaux, les auteurs mettent en avant l'importance de la géométrie dans le problème d'interpolation d'image. Cette observation

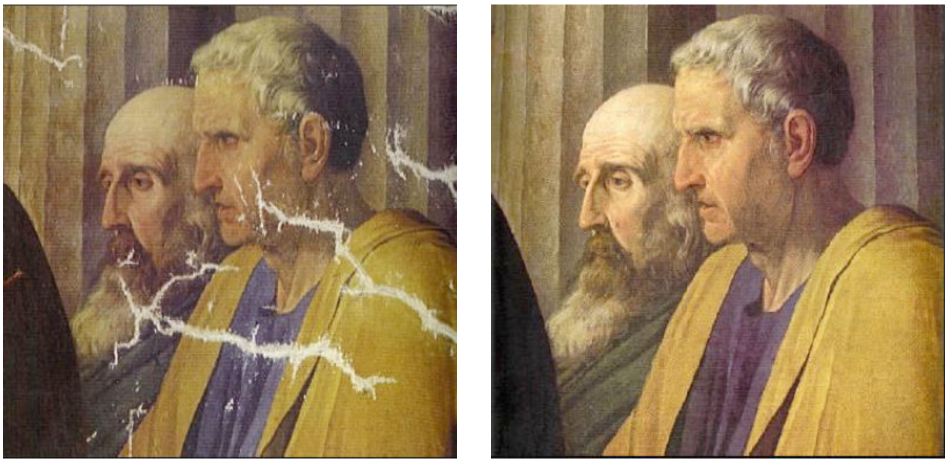

Fig. 6.8 – A gauche, exemple d'oeuvre d'art présentant une éraflure, à droite, l'image après *inpainting* numérique. Détails de "Cornelia Mère de Gracques" par J. Suvee (Louvre).

les conduits à proposer une approche par la résolution d'EDP (Equation Différentielle Partielle) de diffusion anisotrope permettant le remplissage des trous dans une image par prolongement régulier des isophotes. Ces méthodes se sont essentiellement montrées performantes dans le cas d'images lisses par morceaux. Des méthodes plus récentes encore sont décrites dans Verdera et al. (2003); Bornemann and März (2006); Chan et al. (2006).
Une approche très différente de l'*inpainting* est introduite dans Elad et al. (2006) où il est montré qu'une légère modification de l'algorithme MCA permet d'apporter une solution au problème d'interpolation de données manquantes. Cette méthode est étendue ci-après au cas de données représentées sur la sphère et illustrée par différentes expériences numériques, en particulier dans le contexte du projet Planck.

6.2.1 Principe et Algorithme

La méthode d'interpolation de données manquantes décrite dans Elad et al. (2005) suppose que les données complètes admettent une représentation parcimonieuse dans un dictionnaire connu *a priori* qui doit de surcroit être décohérent du dictionnaire dans lequel les données incomplètes sont visualisées, en l'occurrence le dictionnaire *direct* des pixels. En revanche, la représentation des données complètes dans ce dernier dictionnaire est censée ne pas être parcimonieuse. Il est clair en fin de compte que ces hypothèses sont identiques à celles du *compressed sensing* (Donoho 2006b; Candès et al. 2004; Candès and Tao 2005, 2004) qui généralise le mécanisme d'interpolation que nous décrivons ici.
Le sens de ces hypothèses et le lien entre *représentations parcimonieuses* et

inpainting peut être très simplement illustré par l'expérience suivante illustré par la figure 6.9 : considérons l'effet d'un masque rectangulaire sur les coefficients de Fourier d'une onde sinusoidale monochromatique. À cause du caractère non local des éléments de la base de Fourier, il faut un grand nombre de coefficients dans cette base pour représenter le *trou* dans les données, un résultat connu comme le *Phénomène de Gibbs*. En recherchant, dans la base de Fourier, une représentation parcimonieuse de la sinusoide incomplète en dehors de la zone masquée, c'est à dire sans chercher à représenter la région masquée typiquement par des zéros, l'algorithme fini par compléter les données manquantes de manière à remplir l'objectif de parcimonie fixé et permet ainsi de retrouver la sinusoide initiale.

Plus formellement, il est possible d'exprimer notre objectif d'interpolation

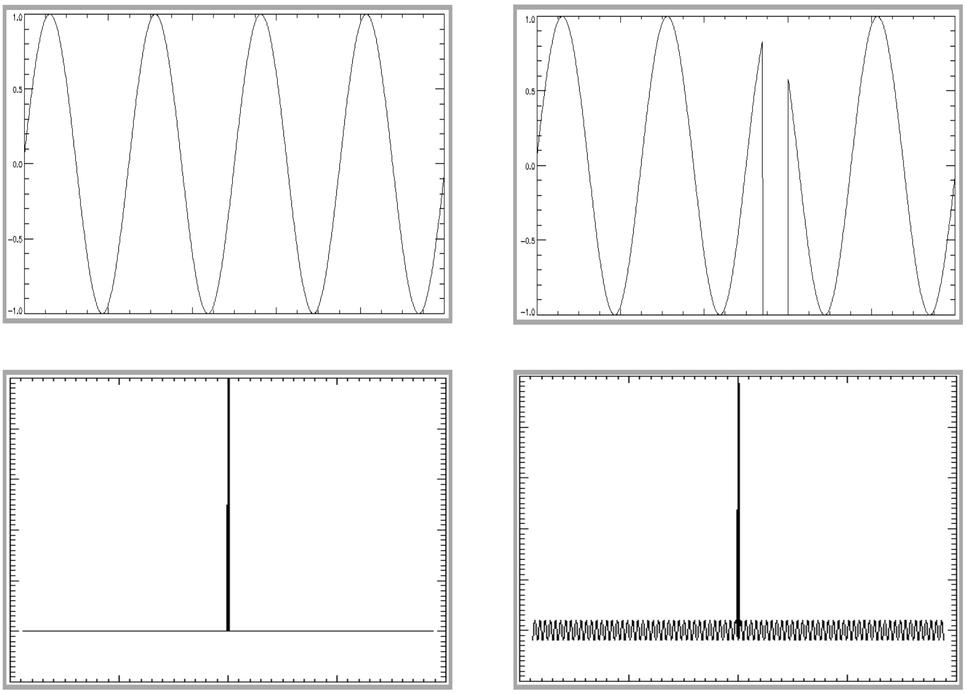

FIG. 6.9 – En haut à gauche, une fonction sinusoide, en haut à droite, la même fonction sinusoide tronquée, en bas à gauche la transformée de Fourier de la fonction sinusoide et en bas à droite, la transformée de Fourier de la fonction sinusoide tronquée.

de données manquantes sur la sphère sous contrainte de parcimonie de la solution dans un dictionnaire spécifié *a priori* en modifiant très légèrement l'equation (6.8). Considérons de nouveau une carte sur la sphère y et supposons que nous connaissons *a priori* les positions des pixels manquants ce que nous formalisons par la donnée d'une carte binaire M sur la sphère. Les données valides dans y sont repérées par des 1 dans M et des 0 dans

M marquent des pixels manquants ou non-valides. Pour l'*inpainting*, nous proposons de modifier la fonction coût de MCA de la manière suivante (Elad et al. 2006) :

$$\min_{s_1,\ldots,s_n} \lambda \sum_{k=1}^{K} \|\alpha_k\|_1 + \left\| M \odot (y - \sum_{k=1}^{K} s_k) \right\|_2^2 \quad \text{with} \quad s_k = \mathbf{\Phi}_k \alpha_k. \tag{6.11}$$

où l'opérateur $\odot$ désigne une multiplication *terme à terme*. De cette manière, la construction d'une représentation parcimonieuse des données y n'est pas contrainte par des données invalides puisque celles-ci ne contribuent pas au terme quadratique d'adéquation aux données. Là où nous n'avons pas de données pour contraindre la construction d'une représentation parcimonieuse de y, l'algorithme est *libre* de compléter la solution à sa guise au regard de l'objectif de parcimonie et ainsi de réaliser l'objectif d'interpolation. De nouveau, la parcimonie de la solution est renforcée par une pénalité sur sa norme ℓ_1 plus aisée à manipuler que la pseudo-norme ℓ_0.
D'autres contraintes peuvent être imposées à la solution. Par exemple, il est montré dans Elad et al. (2006) que l'ajout d'une pénalité sur la *variation totale* dans l'équation (6.11) permet d'améliorer sensiblement la reconstruction des composantes lisses par morceaux. Il est également possible d'imposer une certaine régularité des variations spatiales de quelques statistiques locales, sur les données interpolées. On peut ainsi exiger que la variance empirique sur une carte interpolée soit approximativement égale à l'exterieur et à l'intérieur des régions masquées.
L'algorithme d'interpolation de données manquantes décrit ci-après pour résoudre au moins approximativement le problème (6.11) est naturellement calqué sur l'algorithme MCA. L'interpolation des données résulte de la construction par blocs d'une représentation parcimonieuse des données : l'algorithme consiste en un seuillage itératif alternant sur les composantes s_k, avec un seuil décroissant.

Algorithme

Nous proposons d'approcher une solution au problème (6.11) en utilisant le même processus de seuillage itératif que celui de l'algorithme MCA. La seule modification requise réside dans la multiplication du résidu global par le masque M apèrs chaque estimation du résidu ce qui permet de relacher la contrainte d'adéquation aux données lorsque ces dernières ne sont pas valides. En notant $\mathbf{T}_k$ et $\mathbf{R}_k$ les transformations directes et inverses associées à $\mathbf{\Phi_k}$, l'algorithme *MCA-inpainting* prend la forme suivante :

1. Fixer le nombre d'itérations maximum $I_{\max}$ et le seuil initial $\lambda^{(0)}$
2. Tant que $\lambda_k^{(t)} > \lambda_{\min}$ ($\lambda_{\min}$ égal à zéro ou dépendant de niveau de bruit),
   – Réaliser les itérations suivantes pour estimer les composantes $(s_k)_{k=1,\ldots,K}$ à l'itération t :
   Pour $k = 1, \cdots, K$
   - Calculer le terme résiduel $r^{(t)}$:
     $r^{(t)} = y - \sum_k \tilde{s}_k^{(t-1)}$
   - Estimer les coefficients à l'itération t de $\tilde{s}_k^{(t)}$ en seuillant avec $\lambda_k^{(t)}$:
     $\tilde{\alpha}_k^{(t)} = \delta_{\lambda_k^{(t)}}\left(\mathbf{T}_k\left(M \odot r^{(t)} + \tilde{s}_k^{(t-1)}\right)\right)$
   - Reconstruire une estimation de s_k en reconstruisant à partir des coefficients sélectionnées $\tilde{\alpha}_k^{(t)}$:
     $\tilde{s}_k^{(t)} = \mathbf{R}_k \tilde{\alpha}_k^{(t)}$
   – Faire décroître le seuil λ_k en suivant une stratégie donnée.

Le même processus de raffinement progressif que dans l'algorithme MCA est ici à l'œvre : des détails de plus en plus fins sont incorporés alternativement dans chacune des composantes morphologiques s_k, au fur et à mesure des itérations. C'est une grande force de l'algorithme que de procéder par seuillage itératif avec un seuil décroissant. Ceci confère de la robustesse à l'algorithme en particulier en présence de bruit sur les données. Les différentes stratégies de seuillage décrites dans Bobin et al. (2007b) sont applicables selon les circonstances. Le seuil final est déterminé en fonction des propriétés statistiques du signal et du bruit.
De nouveau, lorsque les transformations utilisées sont non-unitaires, l'approche proposée n'est plus strictement valable. Néanmoins les résultats dans M.Elad (2006); Chaux et al. (2007); Elad et al. (2007); Combettes and Wajs (2005) suggèrent que l'extension est possible. Répétons enfin que l'opérateur de seuillage doux apparait dans ce qui précède comme une conséquence de l'utilisation de la norme ℓ_1. Vers la fin du processus itératif, l'utilisation d'un opérateur de seuillage dur peut conduire à de meilleurs résultats (Starck et al. 2004b).
Différentes expérinces illustrent cet algorithme dans le prochain paragraphe. Nous nous attardons en particulier sur une application en astrophysique : la carte du fond diffus cosmologique que fournira la mission Planck sera partiellement masquée par des résidus d'emissions galactiques d'avant plan. Nous proposons de corriger l'effet des trous dans les données sur l'estimation de certaines statistiques en interpolant préalablement les données manquantes avec l'algorithme *MCA-inpainting*.

6.2.2 Experiences Numériques - Application en Astrophysique

Expérience de principe

La figure 6.10 reproduit une expérience numérique simple ou l'algorithme MCA-inpainting sur la sphère est utilisé pour combler des trous créés artificiellement dans une carte. L'image sphèrique initiale est une vue satellite de

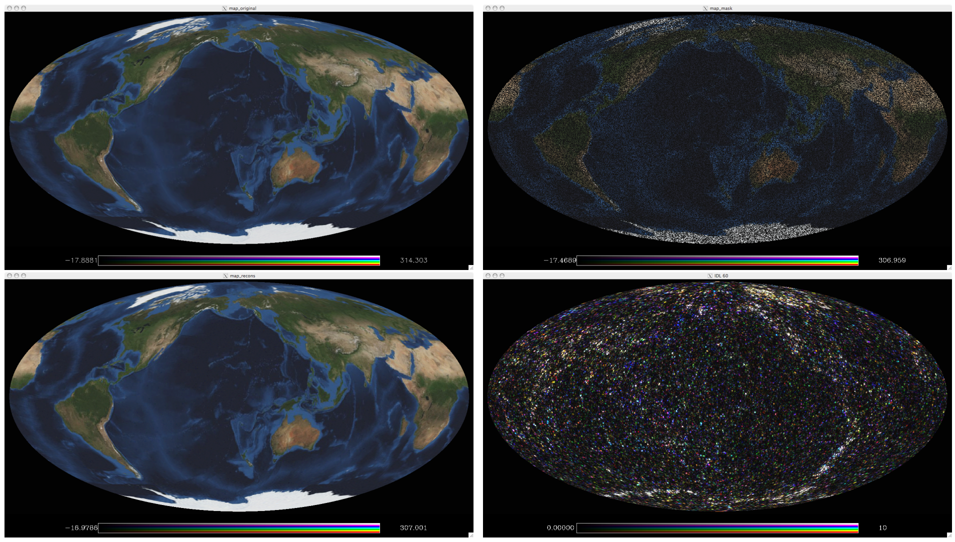

Fig. 6.10 – Application de l'algorithme MCA-inpainting sur la sphère - En haut à gauche : vue satellite originale de la Terre (*moyenne* = 76.9, σ = 47.7). En haut à droite : carte imcomplète obtenue en supprimant aléatoirement 60% des pixels de la carte originale. En bas à gauche : carte interpolée avec l'algorithme MCA-inpainting. Le dictionnaire utilisé ici est formé de la transformation en curvelets et de la transformation en harmoniques sphériques. En bas à droite : carte de l'erreur de reconstruction (*moyenne* = 0.0, σ = 2.86 estimée empiriquement à partir des pixel reconstruits uniquement.)

la Terre dans le domaine visible [2]. L'image retenue est une image couleur : nous traitons séparément les trois canaux rouge, vert et bleu. Vraisemblablement, de meilleurs résultats pourraient être obtenus avec une approche multicanale telle que décrite dans Bobin et al. (2007b).
Une carte incomplète est obtenue en masquant aléatoirement les pixels sur chacune des trois cartes : en l'occurrence, 60% des pixels ont été supprimés. L'application de l'algorithme MCA-inpainting avec simultanément les transformations en harmoniques sphériques et en curvelets permet d'interpoler les pixels manquants de manière à les rendre invisibles à l'oeil nu. Les erreurs de reconstruction, dont la figure 6.10 reproduit la carte, restent faibles en dépit du pourcentage plutot élevé de données manquantes. Au-delà du simple aspect visuel de la carte interpolée, on peut se demander dans quelle mesure procéder à une interpolation préalable des données manquantes permet de faciliter l'estimation de statistiques non locales sur des cartes incomplètes.

[2]disponible sur http ://www.nasa.gov/vision/earth/features/bmng_gallery_4.html

Application à l'analyse du fond diffus cosmologique

Un problème majeur en cosmologie observationnelle contemporaine est la caractérisation statistique des fluctuations du fond diffus cosmologique. En particulier la confirmation du caractère gaussien du CMB et l'estimation précise de son spectre de puissance sont des objectifs importants de l'analyse des données de la prochaine mission d'observation de Planck de l'ESA. Le satellite Planck fournira une cartographie du rayonnement micro-onde sur toute la voûte céleste dans plusieurs bandes de fréquences avec une résolution encore inégalée. L'analyse de ces données permettra de fortement contraindre certains paramètres cosmologiques intervenant dans les modèles actuels d'évolution de l'Univers.
Il y a néanmoins, pour l'analyse des données Planck, des difficultés notables à surmonter. En particulier, le CMB à proprement parler n'est pas la seule source de radiation dans la gamme de fréquence de l'instrument et d'autres sources astrophysiques contribuent au signal observé (R.Bouchet and R.Gispert 1999). La séparation des différentes contributions en vue notamment d'isoler le CMB, est un problème inverse difficile qui concentre les efforts de nombreuses équipes travaillant à la conception de méthodes plus ou moins dédiées (voir Patanchon et al. 2005; Bobin et al. 2006a; Moudden et al. 2005; Pires et al. 2006, et références ci-incluses).

Inévitablement, la carte du fond diffus cosmologique, obtenue après application d'une méthode de séparation, sera contaminée par quelques résidus d'avant-plan, en particulier dans les régions où les émissions galactiques sont fortes ou à l'emplacement de sources radio intenses. Une pratique courante pour l'analyse du CMB est alors de masquer cette partie des données par exemple en utilisant le masque fourni par l'équipe WMAP[3] team) reproduit en haut à gauche de la figure 6.13. Ainsi, l'étude de la non-gaussianité du CMB au moyen de statistiques d'ordre supérieurs estimées dans différentes représentations (*e.g.* ondelettes, curvelets) (Starck et al. 2004a; Jin et al. 2005), ou l'estimation du spectre de puissance du CMB, s'affranchiraient de l'impact de données de mauvaise qualité. Reste que les *trous* introduits de cette manière dans la carte du CMB doivent être manipulés avec soin. Nous illustrons ici l'utilisation de l'algorithme *MCA-inpainting* d'interpolation de données manquantes sur la sphère décrit plus haut pour rétablir la stationnarité d'une carte incomplète du CMB et ainsi réduire l'effet négatif des trous sur l'estimation de statistiques non-locales.

L'application des méthodes développées pour le traitement de données parcimonieuses à l'analyse des cartes du fond diffus cosmologique telles qu'en produira la mission Planck a déjà rencontré quelques succès. Mentionnons en particulier le problème de sépartion de composantes à partir de données

[3]http ://map.gsfc.nasa.gov

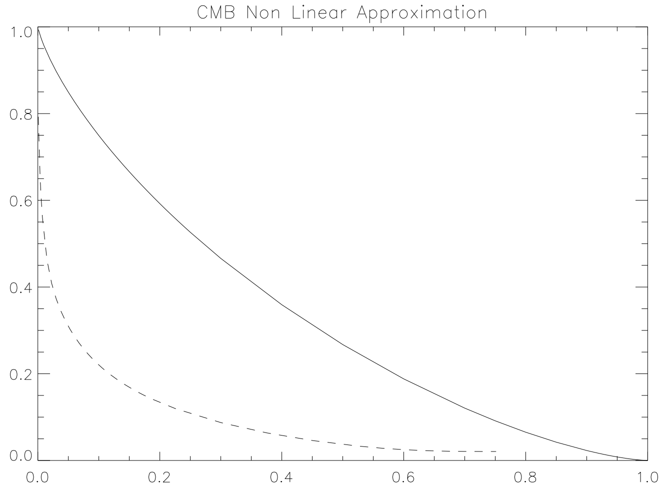

FIG. 6.11 – Courbe représentant l'erreur d'approximation non-linéaire ℓ_2 pour des cartes simulées de CMB dans la représentation en harmoniques sphériques (ligne pointillée) et dans la représentation *pixel* (ligne continue) en fonction de la fraction de coefficients forts gardés pour la reconstruction. La carte de CMB s'avère avoir une représentation parcimonieuse dans une représentation en harmonique sphérique.

multivaluées, qui est approché dans Bobin et al. (2006a,b) sous l'angle de la décomposition parcimonieuse dans un dictionnaire *spatio-spectral* ou *morpho-spectral* multicanal. L'interpolation de données manquantes dans une carte de CMB peut également bénéficier de ce type de méthode. En effet, bien que le CMB à proprement parler soit bien modélisé comme une réalisation particulière d'un champ aléatoire gaussien stationnaire sur la sphère, il n'est pas dépourvu de structure et le concept de parcimonie demeure valable.

Le spectre de puissance des corrélations spatiales du CMB n'est pas plat et les harmoniques sphériques $a_{\ell,m}$ avec $0 \geq m \geq \ell$, du CMB sont bien modélisées par des variables aléatoires gaussiennes indépendantes dont la variance C_ℓ décroît rapidement vers zéro à mesure que l'ordre du multipôle ℓ augmente. Ainsi a distribution empirique de ces coefficients pris dans leur ensemble est parcimonieuse. Une manière pratique de visualiser la parcimonie d'une représentation particulière d'une carte y est de tracer l'évolution de l'erreur quadratique ℓ_2 d'approximation en fonction du nombre N de coefficients -les N plus grands dans ladite représentation- retenus pour cette approximation. La figure 6.11 montre que le CMB à une représentation parcimonieuse dans le domaine des harmoniques sphériques. L'application de l'algorithme *MCA-inpainting* à une carte incomplète du CMB s'appuie sur cette propriété. L'algorithme prend la forme suivante :

1. Fixer le nombre d'itérations maximum $I_{\max}$, le seuil initial $\lambda^{(0)}$ et $\tilde{y}^{(0)} = 0$.

2. Tant que $\lambda^{(t)} > \lambda_{\min}$ réaliser les itérations suivantes :
   - $\tilde{y}^{(t)} = \Phi\delta_{\lambda^{(t)}}\left(\Phi^{-1}\left(\tilde{y}^{(t-1)} + M(y - \tilde{y}^{(t-1)})\right)\right)$
   - Appliquer des contraintes supplémentaires sur la carte reconstruite $\tilde{y}^{(t)}$
   - Faire décroître le seuil $\lambda^{(t+1)} < \lambda^{(t)}$ en suivant une stratégie donnée.

oú Φ désigne en l'occurrence la transformation en harmoniques sphériques. Comme évoqué plus haut, et afin de stabiliser l'algorithme, la carte reconstruite est projétée à chaque itération sur un espace de contraintes supplementaires. Ceci est une conséquence de la précision insuffisante de la transformation en harmoniques sphériques associée au schéma de pixelisation Healpix. Dans le cas de Glesp, la difficulté subsiste mais pour des nombres d'itération plus grands. Pour stabiliser l'algorithme itératif proposé, une contrainte de stationnarité consistant à forcer la variance de la carte interpolée à être égale à l'intérieur et à l'extérieur des régions masquées à chaque échelle d'une décomposition non décimée en *paquets d'ondelettes*. Pour cela, nous utilisons une base (*i.e.* c'est à dire une partition de l'intervalle des ℓ) adaptée à la connaissance *a priori* que nous avons du spectre du CMB. Une telle partition réalise un compromis entre l'exigence de bandes larges pour une meilleure localisation des coefficients à l'intérieur ou à l'extérieur du masque, et l'exigence de bandes étroites pour une meilleure représentation *a priori* du spectre de puissance du CMB.

Finalement, l'algorithme ci-dessus a été appliqué avec succès pour interpoler les pixels manquants dans des cartes de CMB incomplètes. La figure 6.12 reproduit la carte du CMB délivrée par le consortium WMAP[4] ainsi que la carte resultant du processus d'inpainting permettant une première évalution visuelle de la méthode proposée sur ce type de données. Les pixels masqués dans la carte initiale le sont du fait d'un niveau attendu élevé de contamination par des résidus d'avant-plan. Après interpolation, le masque est indétectable à l'œil. Il ne l'est pas davantage aux différentes échelles d'une décomposition non decimée en ondelettes isotropes de la carte interpolée reproduites sur la figure 6.13.

L'algorithme a ensuite été appliqué à une centaine de cartes de CMB simulées sur la sphère masquées en partie en utilisant le même masque $kp0$ de WMAP que ci-dessus visible sur la figure 6.13, avec une centaine d'itérations dans le domaine des harmoniques sphériques. La figure 6.14 compare le spectre de puissance moyen estimé sur les cartes complètes au spectre estimé sur les cartes incomplètes après une simple correction au premier ordre par le facteur de couverture du ciel ainsi qu'au spectre estimé sur les cartes interpolées grâce à l'algorithme proposé. Il apparait clairement que l'interpo-

[4]http ://map.gsfc.nasa.gov/

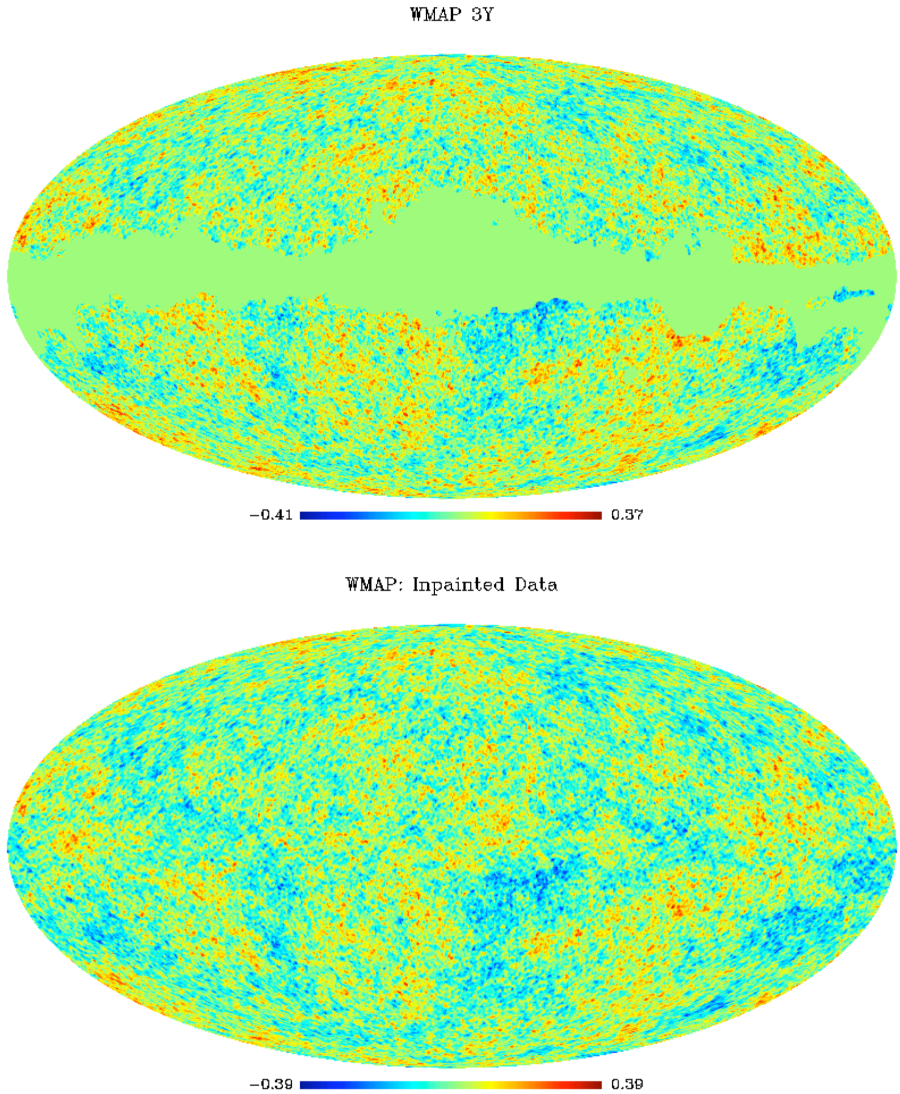

Fig. 6.12 – A gauche Carte de CMB obtenue par WMAP. La zone galactique sujette à d'importantes contaminations par les sources d'avant-plan ainsi que les sources ponctuelles radio ont été masquées. A droite : Carte obtenue en appliquant l'algorithme MCA-inpainting proposée sur la carte masquée de WMAP.

lation préalable des données permet une réduction significative du biais sur l'estimation spectrale que la simple correction naive par le facteur de couverture. En revanche, le prix à payer est une légère augmentation de la variance de l'estimateur. Pour juger du caractère gaussien d'un champ, il est commun de recourir à des tests faisant intervenir des moments d'ordres supérieurs ty-

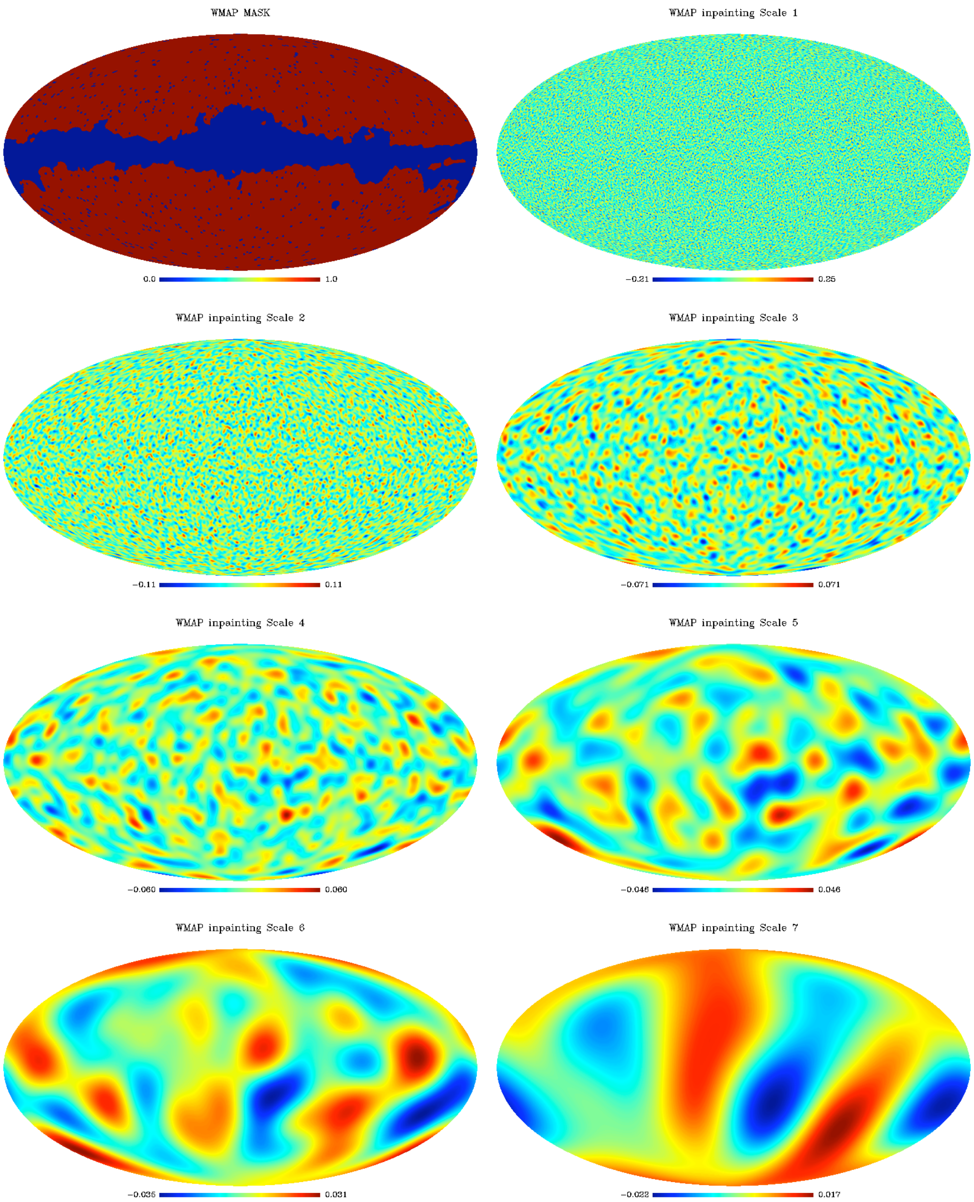

FIG. 6.13 – En haut à gauche : la zone masquée. De haut en bas et de gauche à droite, les sept décompositions ondelette de la carte inpaintée. D'un point de vue visuel, la zone masquée n'est pas detectable d'aucune façon dans les differentes échelles ondellette de la carte inpaintée.

piquement 3 ou 4 pour le kurtosis. Changer les données de represéntation avant d'estimer ces moments peut être un moyen de faire ressortir des structures autrement indécelables. Ces tests statistiques non-locaux sont difficiles

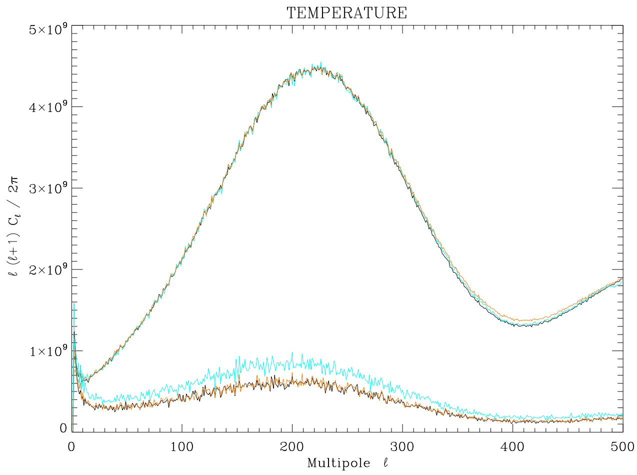

Fig. 6.14 – Spectre de puissance du CMB estimé à partir de données du CMB simulée complète (en noir), à partir de données incomplètes mais corrigé par un facteur de couverture du ciel (en orange), à partir de données inpaintées (en bleu). Les courbes inférieures représentent la variance empirique (multipliée par 2 pour des raisons visuelles) en fonction de ℓ pour les 3 estimateurs du spectre de puissance.

à mettre en œuvre sur des cartes incomplètes de sorte qu'une interpolation préalable peut s'avérer bénéfique. En reprenant les cartes simulées complètes et interpolées utilisées ci-dessus, nous avons pu montré que l'algorithme proposé n'affecte pas notablement ce type de genre de test de gaussianité. Ainsi nous pensons, bien qu'il s'agisse d'une étude à approfondir, que la méthode d'interpolation de données manquantes sur la sphère introduite dans ce chapitre permettra de distinguer un CMB qui serait véritablement non gaussien d'une non gaussianité qui résulterait de la non-stationarité d'une carte incomplète.

6.3 Conclusion

Nous avons discuté et illustré dans ce chapitre différentes applications des transformations multi-échelles sur la sphère construites dans les premiers chapitres de ce mémoire. En particulier, la variété des dictionnaires désormais disponibles pour la représentation de données sur la sphère nous a permis d'ouvrir la voie de l'analyse de données parcimonieuses sur la sphère. En pratique, nous avons élargi le domaine d'application de l'algo-

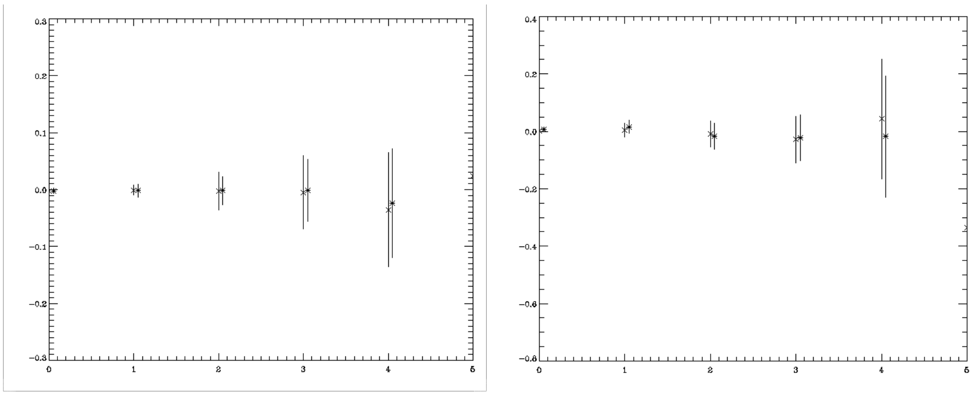

FIG. 6.15 – Horizontalement se trouve le nombre d'échelles croissant à partir des basses fréquences. à gauche : la skewness des coéficients d'ondelette à une échelle des données sphériques de CMB originale (o) et des données inpaintés (x).à droite : le kurtosis des coéficients d'ondelette à une echelle donnée de la carte originale (O) et de la carte inpaintée (x). Les barres d'erreur sont estimées sur un petit jeu de quinze simulations de CMB complet.

rithme MCA à ce type de données et illustré son bon fonctionnement sur des données réelles pour la caractérisation de défauts surfaciques sur des objets sphériques. Nous avons également étendu l'algorithme *MCA-inpainting* de manière à permettre l'interpolation de données manquantes sur la sphère. L'application à l'analyse du CMB cartographié sur la sphère est prometteuse : les résultats préliminaires obtenus pour l'estimation du spectre de puissance ou l'étude de la gaussianité du CMB justifient de poursuivre dans cette voie. Les domaines où les données sont naturellement récoltés sur la sphère sont nombreux : en géophysique, astrophysique ou imagerie médicale. Dans tous les cas ou ces données sont structurées et admettent des représentations parcimonieuses dans l'un ou l'autre des dictionnaires de fonctions sur la sphère, les algorithmes décrits sont des outils numériques performants, rapides et en fin de compte assez aisés à mettre en œuvre.

Chapitre 7
Détection de non-Gaussianités dans le CMB sur la sphère

7.1 Introduction

La recherche de signes d'une éventuelle présence de non-Gaussianité dans les cartes de fluctuations de température du fond diffus cosmologique (CMB) est d'un grand interêt pour la cosmologie. Des non-Gaussianités ont été vues par WMAP [1] (Komatsu et al. 2003; McEwen et al. 2008), mais restent controversées et on attend maintenant beaucoup des futures données Planck.

Certaines questions fondamentales sur les débuts de l'Univers (théories de l'inflation, théorie des cordes, ...) sont intrinsèquement reliées à la présence de non-Gaussianités dans le CMB, qui peuvent prendre différentes formes. La non-gaussianité du CMB peut également avoir pour origine des effets cosmologiques telles que l'effet Sunyaev-Zel'dovich (SZ) (Sunyaev and Zeldovich 1980) (qui correspond à un effet Compton inverse) produit par des nuages de gaz chauds et ionisés intergalactique (Aghanim and Forni 1999; Cooray 2001), ou encore les effets de lentille gravitationnelle dues à des structures de très grandes tailles (Refregier 2003). Enfin, elles peuvent également être simplement liées à des résidus d'émissions d'avant plan (Jewell 2001), telles que les émissions de notre propre Galaxie, ou encore à des bruits instrumentaux non-Gaussiens et à des systématiques (Banday et al. 2000).

Détecter la non-Gaussianité dans le CMB et en déterminer ses caractéristiques est un enjeu important pour contraindre les processus de formation de l'Univers et sa morphologie. Un grand nombre d'études se sont portées sur cette problématique. De nombreuses approches ont été utilisées : fonctionnelles de Minkowski et statistiques morphologiques (Novikov et al. 2000; Shandarin 2002), le bispectre (Bromley and Tegmark 1999; Verde et al. 2000; Phillips and Kogut 2001), le trispectre (Kunz et al. 2001), les transformées ondelettes (Aghanim and Forni 1999; Forni and Aghanim 1999; Hobson et al. 1999; Barreiro and Hobson 2001; Cayón et al. 2001b; Jewell 2001; Starck et al. 2004a), et curvelets (Starck et al. 2004a; Jin et al. 2005).

[1] http ://map.gsfc.nasa.gov/

Dans ce chapitre, nous montrons comment les transformées muti-échelles sur la sphère présentées précédemment peuvent être des outils remarquables pour la recherche de non-Gaussianité.

7.2 Les ondelettes en action

Dans Aghanim et al. (2003); Starck et al. (2004a), il est montré que la transformation en ondelettes est un outil très puissant pour détecter certaines signatures de non-Gaussianité. Des études récentes ont montré la présence plausible de non-Gaussianité dans les données du satellitte WMAP (Vielva et al. 2004; Mukherjee and Wang 2004; Cruz et al. 2005, 2007; McEwen et al. 2008). Le détecteur de non-Gaussianité utilisé consiste d'abord à appliquer une transformée en ondelettes, puis à étudier la distribution des coefficients obtenus. Pour caractériser cette distribution, plusieurs outils statistiques sont disponibles :

- Le coefficient d'asymétrie (S) :
  Le coefficient d'asymétrie (souvent appelé *skewness*) est le moment d'ordre 3 normalisé d'une distribution. Il est couramment utilisé pour mesurer la non-Gaussianité. Le coefficient d'asymétrie S_κ d'une carte κ est défini de la manière suivante :

$$S_\kappa = \sum_{i=1}^{N} \frac{(\kappa_i - \overline{\kappa})^3}{(N-1)\sigma^3}. \tag{7.1}$$

  C'est une mesure de l'asymétrie de la distribution.
- Le coefficient d'aplatissement (K) :
  Le coefficient d'aplatissement (souvent appelé *kurtosis*) est le rapport entre le moment centré d'ordre 4 et le carré du moment centré d'ordre 2. Il est également souvent utilisé pour caractériser la non-Gaussianité. L'excès d'aplatissement K_κ est défini de la manière suivante :

$$K_\kappa = \sum_{i=1}^{N} \frac{(\kappa_i - \overline{\kappa})^4}{(N-1)\sigma^4} - 3. \tag{7.2}$$

  C'est une mesure de l'aplatissement de la distribution. Un excès d'aplatissement positif correspond à une distribution pointue et un excès d'aplatissement négatif à une distribution aplatie.
- Le *Higher Criticism* (HC) :
  Le *Higher Criticism* (HC) est une statistique qui a été récemment développée par Donoho and Jin (2004). C'est un estimateur de non-Gaussianité obtenue en cherchant la déviation maximum des *p-values* triées par rapport à une distribution Gaussienne. Pour définir le HC, il faut tout d'abord convertir chaque pixel de la carte κ en *p-values*.

Soit $p_i = P\{N(0,1) > \kappa_i\}$, la $i$ème *p-value*, et soit $p_{(i)}$ les *p-values* triées par ordre croissant. Le HC est défini comme suit :

$$HC^* = \max_i \left| \frac{\sqrt{n}[i/n - p_{(i)}]}{\sqrt{p_{(i)}(1 - p_{(i)})}} \right|. \tag{7.3}$$

Ou sous sa forme modifiée :

$$HC^+ = \max_{i:1/n \leqslant p_{(i)} \leqslant 1-1/n} \left| \frac{\sqrt{n}[i/n - p_{(i)}]}{\sqrt{p_{(i)}(1 - p_{(i)})}} \right|. \tag{7.4}$$

Ainsi une grande valeur de HC suppose la présence de non-Gaussianité. HC permet de détecter même une faible quantité de non-Gaussianité.

En pratique, plusieurs difficultés doivent être prises en compte :
- sources ponctuelles : les sources ponctuelles ne peuvent pas être enlevées par des algorithmes de séparation de composantes, et il est nécessaire de les détecter et de les masquer dans la carte de CMB.
- résidu de la séparარion de composantes : comme nous l'avons expliqué dans le premier chapitre, la carte de CMB est dérivée d'une méthode de séparation de composante. La séparation n'est jamais parfaite, et des résidus galactiques subsistent. Les zones contaminées doivent donc être masquées avant toute étude de non-Gaussianité. On ne dispose donc pas de données sur toute la sphère, et il faut alors contrôler les effets du masque sur la mesure de non-Gaussianité.
- bruit : le bruit est en général considéré comme Gaussien, mais il est non stationnaire. Sa variance varie suivant la position spatiale. Cet effet peut pertuber les mesures sur les echelles fines, s'il n'est pas correctement étudié.
- *map-making* : si le *map-making* n'est pas parfait, des stries seront présents dans la carte de CMB.

Pour contourner ces probèmes, on utilise des simulations qui modélisent les données au mieux, et on compare en fait la distribution des coefficients de la carte observée à la distribution des coefficients des cartes simulées. La simulation consiste à génerer des cartes de CMB issues de réalisations d'un modèle cosmologique de référence, compatible avec les données, à ajouter à ces cartes tous les effets connus du ciel et de l'instrument, puis à appliquer la même chaine de traitement de données que pour le CMB observé. Un point important à mentionner est que si un écart entre la distribution des coefficients du CMB et celle des simulations est mesuré, cela ne signifie pas forcément que le CMB est non-Gaussien, mais uniquement que le CMB n'est pas comptatible avec nos simulations. Dans le cas où les simulations sont suffisamment réalistes, alors on peut prétendre à une réelle détection.

7.3 Discrimination de non-Gaussianité

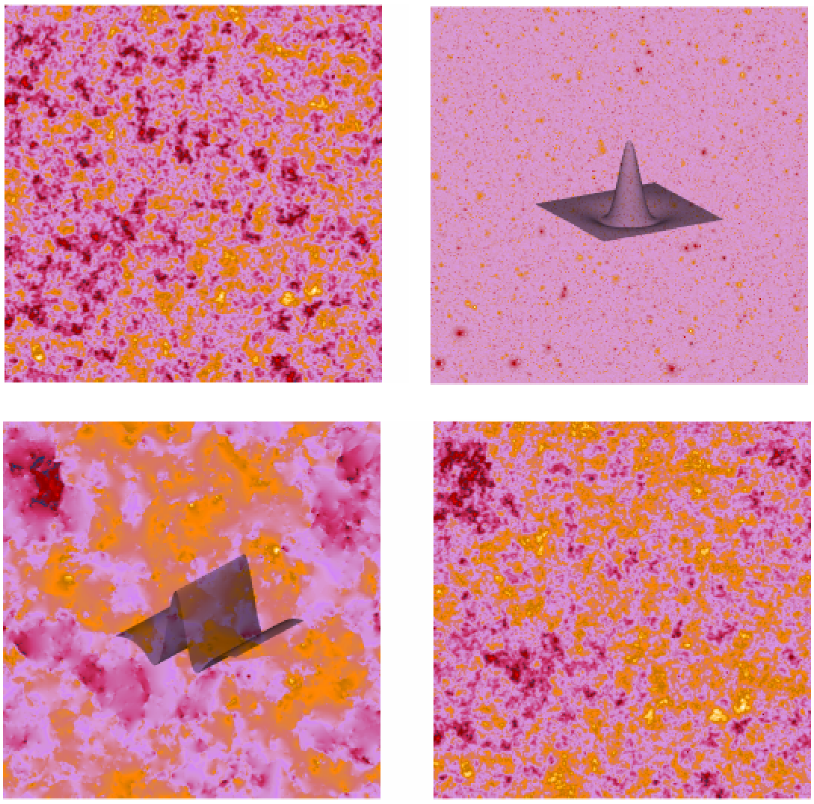

FIG. 7.1 – En haut à gauche : carte simulée du CMB. En haut à droite : fluctuations Sunyaev-Zel'dovich cinétiques (SZ) . En bas à gauche : carte de cordes cosmiques. En bas à droite : simulation d'observation contenant les trois composantes. Une fonction ondelette est affichée par dessus la carte SZ et une fonction curvelet par dessus celle de cordes cosmiques.

7.3.1 Sources ponctuels

Des nombreuses méthodes ont été proposées ces dernières années pour la détection de sources pontuelles dans le CMB, telles que l'ondelette chapeau mexicain (Cayón et al. 2000, 2001b), le pseudo filtre (Sanz et al. 2001), ou encore le filtre biparamétrique à echelle adaptative (López-Caniego et al. 2005). Une méthode simple et robuste, qui maximise le rapport signal-sur-bruit est le "Matched Filter" (Vio et al. 2002). En considérant que l'instrument est à symetrie de révolution et de réponse spectrale connue $\tau(q)$, et en notant

$P(q)$ le spectre de puissance du CMB, le *matched filter* est donné par Vio et al. (2002) :

$$\widehat{\psi}_{MF}(q) = \frac{1}{2\pi\alpha}\,\frac{\tau(q)}{P(q)}, \qquad \alpha \equiv \int_0^{+\infty} q\frac{\tau^2}{P}\,dq, \tag{7.5}$$

avec une variance minimale :

$$\sigma^2 = \frac{1}{2\pi\alpha}. \tag{7.6}$$

Si la PSF est inconnue (ou variable spatialement), l'ondelette chapeau mexicain peut être une bonne alternative. Il s'agit alors de convoluer les données par l'ondelette chapeau mexicain (voir 3.2.2) $\psi_{a,b}(x) = \psi(\frac{x-b}{a})$, où $\psi(x) = \frac{1}{\sqrt{2\pi}}(1-x^2)e^{-x^2/2}$. a est ici le paramètre d'échelle et b le paramètre de position. Une implémentation rapide est possible en passant par le domaine des harmoniques sphériques (López-Caniego et al. 2005).

7.3.2 Analyse statistique des coefficients ondelettes et curvelets

Nous avons vu que les propriétés statistiques des anisotropies de température du CMB nous informent sur la physique de l'univers primordial. Si leur distribution est gaussienne, elles sont générées par des modèles simples d'inflation, sinon elles sont issues de défauts topologiques comme des cordes cosmiques. Mais des anisotropies peuvent également provenir de l'interaction des photons du CMB avec des électrons libres du gaz chaud intra-amas (effet Sunyaev-Zel'dovich). Une des problématiques des études de non-Gaussianité du CMB est donc notre capacité à démêler les différentes sources de non-Gaussianité les unes des autres.

Il a été montré que l'utilisation combinée de deux transformées multi-échelles, une décomposition sur une base d'ondelettes et une transformée en curvelets, permet non seulement de détecter de manière optimale les anisotropies dans le CMB, mais aussi de déterminer leur origine, ce qui est impossible avec les méthodes traditionnelles (Starck et al. 2004a). En effet, les ondelettes sont bien adaptées à l'analyse des structures spatialement isotropes et permettent donc de bien détecter l'effet Sunyaev-Zel'dovich, tandis que les fonctions curvelets sont optimales pour la recherche de structures spatialement anisotropes. Chacune de ces transformations fournit un estimateur bien adapté à une type particulier de non-Gaussianité. Une combinaison des deux estimateurs donne alors des informations à propos de la nature des signaux non-Gaussiens.

Pour illustrer ceci, nous montrons à la figure 7.1 des cartes simulées de CMB (en haut à gauche), de SZ cinétique (en haut à droite), de cordes cosmiques (en bas à gauche), ainsi que le mélange des trois cartes précédentes (en bas à droite). On peut remarquer aisément que le CMB domine tout le signal, sauf aux très petites échelles. Une fonction ondelette est affichée par dessus la carte de SZ cinétique et une fonction curvelet par dessus celle de cordes cosmiques. Ces deux transformations sont celles qui s'adaptent le mieux à chaque type de données respectivement.

Analyse statistique combinée sur la sphère

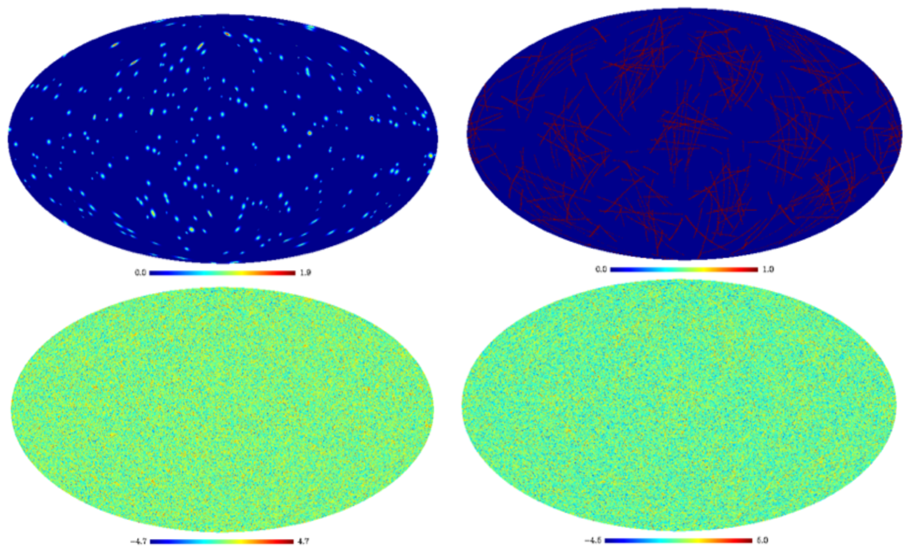

FIG. 7.2 – En haut, images contenant des Gaussiennes et des droites. En bas, mêmes images, mais contaminées par un bruit Gaussien. Le rapport signal-sur-bruit est égal à 1.

Le même principe peut être appliqué pour une analyse de données sur la sphère. Grâce aux transformées en ondelettes et curvelets sur la sphère développées dans cette thèse, nous pouvons calculer les kurtosis par échelle avec les deux transformées et observer laquelle permet de mieux détecter la non-Gaussianité. Une meilleure détection dans l'espace ondelette indiquerait que la morphologie des structures générant la non-Gaussianité est plutôt isotrope, tandis que des structures anisotropes (stries résiduelles, cordes cosmiques, etc) seraient mieux détectées avec des curvelets.

Les deux images en haut à gauche et en haut à droite sur la figure 7.2 contiennent respectivement des Gaussiennes et des lignes. Nous avons creer un ensemble d'images simulées en ajoutant du bruit Gaussien, avec différents niveaux de bruit à ces deux images. Le rapport signal-sur-bruit (SNR) varie de 0 à 1. Pour l'image avec des lignes, le SNR est défini comme le rapport de la valeur du pixel le long de la ligne sur l'écart type du bruit, et pour l'image contenant les Gaussiennes, le SNR est défini comme le rapport entre le maximum de la Gaussienne et l'écart type du bruit. La figure 7.2 montre,

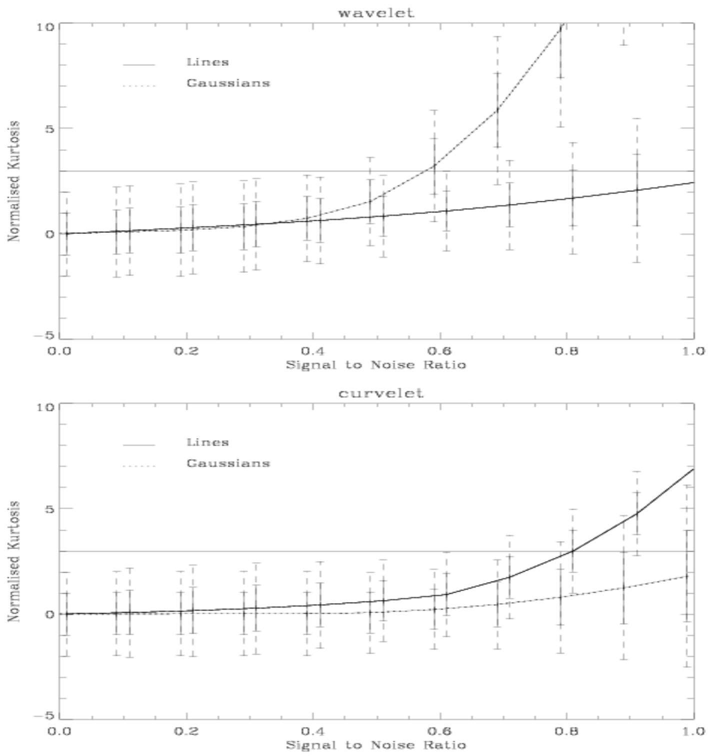

FIG. 7.3 – Kurtosis normalisé en fonction du rapport signal-sur-bruit pour les coefficients ondelettes (en haut) et curvelets en bas). La ligne continue correspond à l'analyse des images contenant des droites, et la ligne en pointillé celle des images contenant des Gaussiennes. On peut voir que les ondelettes détectent mieux les Gaussiennes que les curvelets, mais que les curvelets sont plus sensibles pour trouver les droites noyées dans du bruit. Les barres d'erreur correspondent à 1σ et celles avec des traits à 2σ.

en bas à gauche et à droite, les deux images bruitées avec un SNR égal à 1. Pour chaque valeur de SNR, nous avons généré trente réalisations du bruit, que nous avons ajoutées aux deux images, puis nous avons calculé le kurtosis de chaque bande de la transformée en ondelettes et de la transformée en curvelets. Ces valeurs de kurtosis ont ensuite été normalisées par l'ecart type des kurtosis obtenues sur les transformées de trente autres réalisations d'un bruit blanc Gaussien. On note $\mathcal{K}_{\mathcal{W}}^{X}[j, i, n]$ le kurtosis de la $j$ème échelle en ondelettes, pour la $i$ème réalisation d'un bruit avec un SNR indéxé par n. $X = G, L$ ou N pour une carte qui contient respectivement des Gaussiennes plus du bruit, des lignes plus du bruit ou du bruit blanc seulement.

Les kurtosis moyens normalisés $\bar{\mathcal{K}}_{\mathcal{W}}^{G}$ et $\bar{\mathcal{K}}_{\mathcal{W}}^{L}$ par bande j et pour un SNR

n pour l'image des lignes et Gaussiennes sont alors :

$$\begin{aligned}
\bar{\mathcal{K}}^G_{\mathcal{W}}[j,n] &= \frac{< \mathcal{K}^G_{\mathcal{W}}[j,*,n] >}{\sigma\left(\mathcal{K}^N_{\mathcal{W}}[j,*,n]\right)} \\
\bar{\mathcal{K}}^L_{\mathcal{W}}[j,n] &= \frac{< \mathcal{K}^L_{\mathcal{W}}[j,*,n] >}{\sigma\left(\mathcal{K}^N_{\mathcal{W}}[j,*,n]\right)}
\end{aligned} \qquad (7.7)$$

Finallement, on ne garde pour chaque niveau de bruit que le kurtosis moyen maximum le long des échelles, et on a :

$$\begin{aligned}
\mathrm{K}^G_{\mathcal{W}}[n] &= \max_j \left(\bar{\mathcal{K}}^G_{\mathcal{W}}[j,n]\right) \\
\mathrm{K}^L_{\mathcal{W}}[n] &= \max_j \left(\bar{\mathcal{K}}^L_{\mathcal{W}}[j,n]\right)
\end{aligned} \qquad (7.8)$$

On définit de la même manière les kurtosis normalisés dans l'espace curvelet :

$$\begin{aligned}
\mathrm{K}^G_{\mathcal{C}}[n] &= \max_j \left(\bar{\mathcal{K}}^G_{\mathcal{C}}[j,n]\right) \\
\mathrm{K}^L_{\mathcal{C}}[n] &= \max_j \left(\bar{\mathcal{K}}^L_{\mathcal{C}}[j,n]\right)
\end{aligned} \qquad (7.9)$$

$\mathrm{K}^G_{\mathcal{W}}$, $\mathrm{K}^L_{\mathcal{W}}$ (resp. $\mathrm{K}^G_{\mathcal{C}}$, $\mathrm{K}^L_{\mathcal{C}}$) nous donnent alors la puissance de détection en fonction du niveau de bruit de la transformée en ondelettes (resp. curvelets) pour détecter des Gaussiennes et des lignes. La figure 7.3 à haut (resp. en bas) montre les courbes $\mathrm{K}^G_{\mathcal{W}}$ et $\mathrm{K}^L_{\mathcal{W}}$ (resp. $\mathrm{K}^G_{\mathcal{C}}$ et $\mathrm{K}^L_{\mathcal{C}}$). La ligne continue correspond à l'analyse des images contenant des droites (c'est à dire $\mathrm{K}^L_{\mathcal{W}}$ en haut et $\mathrm{K}^L_{\mathcal{C}}$ en bas), et la ligne en pointillés celle des images contenant des Gaussiennes (c'est à dire $\mathrm{K}^G_{\mathcal{W}}$ en haut et $\mathrm{K}^G_{\mathcal{C}}$ en bas). On peut voir que les ondelettes détectent mieux les Gaussiennes que les curvelets, mais que les curvelets sont plus sensibles pour trouver les droites noyées dans du bruit. Les barres d'erreur correspondent à 1σ et celles avec des traits à 2σ. La ligne horizontale avec une ordonnée égale à 3 correspond à un niveau de détection à 3σ.

7.4 Débruitage et Non-Gaussianité

La superposition d'un signal non-Gaussien à un signal Gaussien, peut être modelisée par $y = z + x$, où y est l'image observée, x la composante non-Gaussienne, et z une composante Gaussienne. En pratique, la composante Gaussienne est la superposition du CMB et du bruit. C'est donc un champ Gaussien, mais qui n'est pas blanc. Si le signal non-Gaussien x est très faible, alors on ne peut pas l'estimer, et seule une étude statistique de la distribution de y (ou des coefficients de y dans une décomposition donnée) peut permettre une detection. Si x est suffisamment grand, alors on peut tenter d'estimer x connaissant y et les propriétés statisitiques de z. Le problème de la détection de non-Gaussianité devient alors un problème de débruitage, mais où le *bruit* contient aussi le CMB.

Suppréssion des résidus galactiques et des sources ponctuels dans le CMB

Dans le cadre du *Working Group 2* du projet PLANCK, des simulations ont été faites pour tester les algorithmes de séparation de composantes. Il reste en général dans la carte de CMB un résidu galactique et des sources ponctuelles qui sont bien visibles. Ces non-Gaussianités entrainent forcément un biais sur l'estimation du spectre de puissance de la carte de CMB. On peut alors essayer de les estimer en appliquant un algorithme de filtrage qui élimine le CMB et le bruit. Il est évident que cette tâche est très délicate, car on ne veut pas dégrader notre estimée du CMB. Pour cela, pour estimer x, nous allons adopté la stratégie suivante :

– Nous allons utilisé des ondelettes isotropes sur la sphère.
– Le bruit va être considéré comme Gaussien et corrélé. Comme on l'a décrit au chapitre 5, l'écart type du bruit à léchelle j est alors calculé par $\sigma_j = \text{median}(\mid w_j \mid)/0.6745$.
– Le seuil doit être suffisamment élevé pour éliminer tout le CMB. On préfère clairement laisser des non-Gaussianités dans le CMB, plutôt que de commettre l'erreur d'attribuer un coefficient en ondelettes du CMB au résidu galactique et aux sources ponctuelles. Des expériences ont montré qu'un seuil à 5σ est un bon choix, suffisamment conservateur.
– Le nombre d'échelles doit être limité. Les coefficients ondelettes des grandes échelles sont tous attribués au CMB, et la dernière échelle ondelette (c'est à dire le plan lissé) est considéré comme n'étant que du CMB. Nous devons donc la supprimer. Dans le cadre de PLANCK ($nside = 2048$), nous avons utilisé six échelles (cinq échelles ondelettes et le plan lissé).

En appliquant l'algorithme de débruitage en ondelettes présenté au chapitre 5 et en utilisant l'agorithme de reconstruction itératif, nous avons donc mis en place un mécanisme d'élimination de la composante z, et donc d'estimation de x. Pour avoir une carte du CMB *nettoyée* , il nous suffit maintenant de soustraire $\tilde{x}$ à la carte de CMB y : $y_{clean} = y - \tilde{x}$, où $\tilde{x}$ est notre estimation de x.

la figure 7.4 montre le résultat de cette méthode sur une carte issue de la séparation de composantes dans le cadre de PLANCK-WG2. Les résidus galactiques sont très visibles sur l'image du haut. Apres filtrage, nous avons une estimée de x (carte du mileu). Le CMB nettoyé est montré sur la carte du bas.

7.5 WMAP et le Cold Spot

Comme cela a été présenté dans le premier chapitre, le satelite WMAP a été lancé en 2001, et a permis d'obtenir les premières cartes du fond diffus cosmologique couvrant tout le ciel avec une résolution de 10' d'arc. En utili-

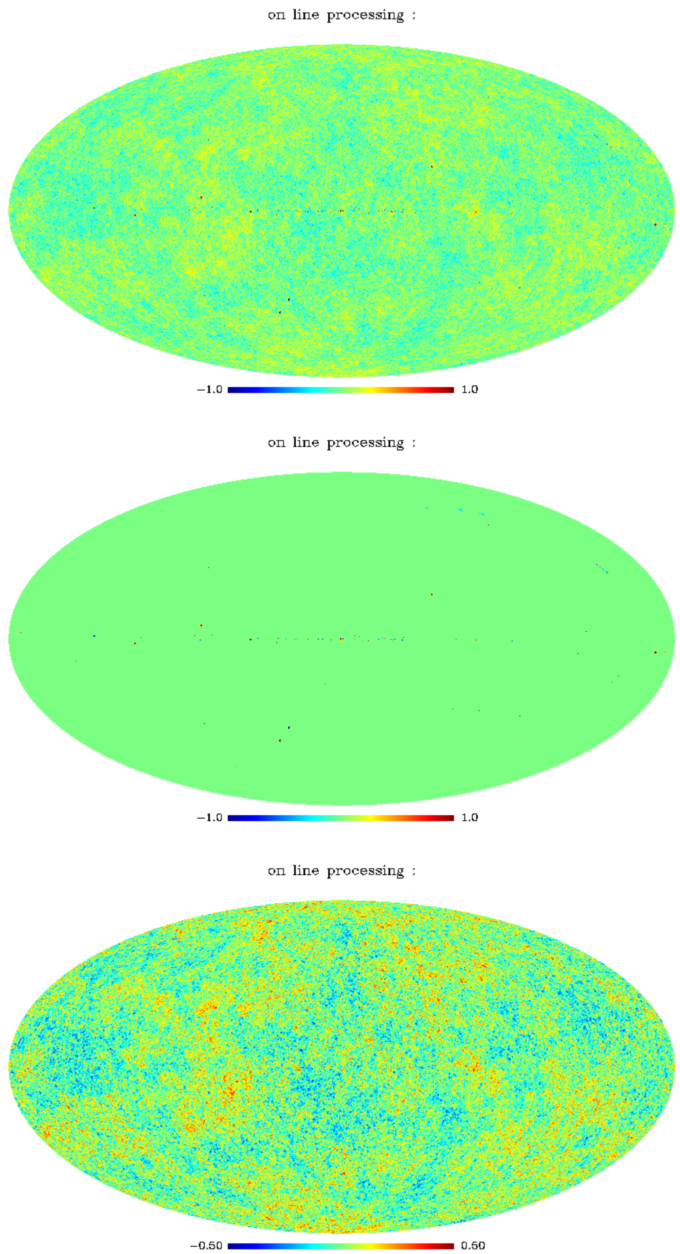

Fig. 7.4 – En haut carte de CMB après la séparation de composantes. Au milieu, estimée du résidu galactique et des sources ponctuelles. En bas, différence entre les deux cartes précédentes.

sant des ondelettes continues sur la sphère, une détection d'un *cold spot* a été faite à une latitude galactique $b = -57°$ et à une longitude $l = 209°$ (Vielva et al. 2004; Cruz et al. 2005; Vielva et al. 2006; Cruz et al. 2007; McEwen

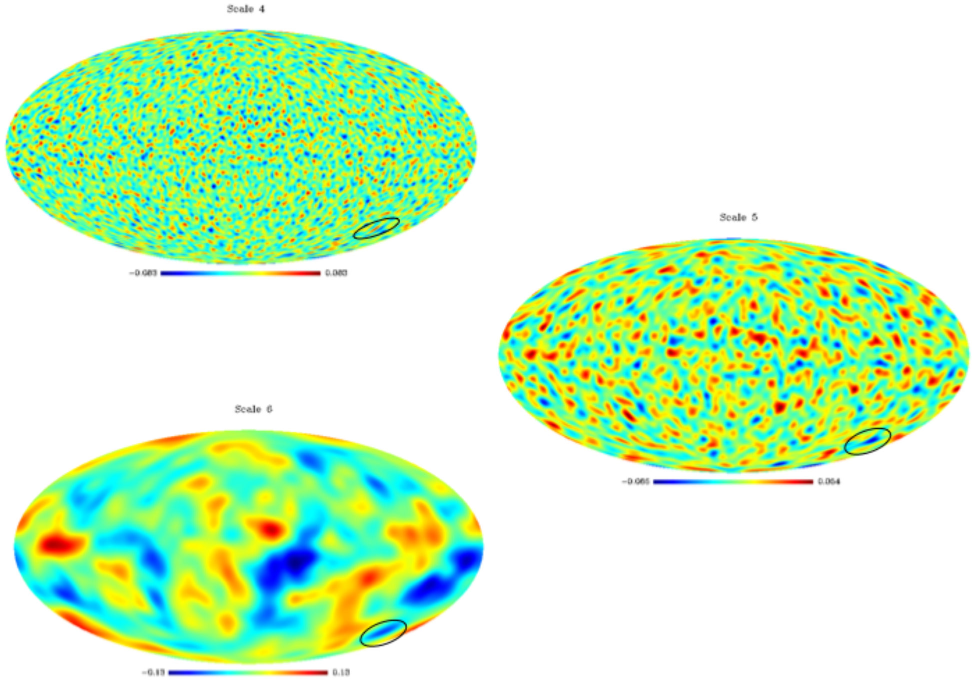

Fig. 7.5 – Trois échelles de la transformée en ondelettes de la carte de CMB-WMAP, avec la position du *cold spot* indiquée par une ellipse.

et al. 2007). Si la détection semble être confirmée par d'autres observations (Genova-Santos et al. 2008), elle reste tout de même controversée (Smith and Huterer 2008). La figure 7.5 montre trois échelles de la transformée en ondelettes de la carte de CMB-WMAP, avec la position du *cold spot* indiquée par une ellipse.

Cette détection a été faite à partir de simulations de cartes de CMB, comme nous l'avons expliqué au début de ce chapitre. Le filtrage nous offre une voie alternative pour l'analyse du CMB. On peut en effet essayer d'extraire du CMB les zones non-Gaussiennes, d'une manière identique à ce que l'on a fait pour les résidus galactiques.

Nous avons donc tenter d'évaluer la probabilité d'existence réel de ce fameux *cold spot*, en regardant à partir de quel niveau de détection on peut l'extraire du CMB par le filtrage. La figure 7.6 montre le résultat après un filtrage à 4σ. Le *spot* est donc bien détecté. Toutefois, la probabilité d'obtenir un tel évènement n'est pas simple à dériver, car le test de détection à 4σ est répété sur toutes les échelles. On est donc dans un cas de tests multiples, bien connu des statisticiens, et la probabilité d'apparition d'un tel spot sous une hypothèse de Gaussianité est bien plus grande que ce qu'on aurait dans le cas d'un test simple (c'est à dire non multiple). A titre d'exemple, il nous a fallu

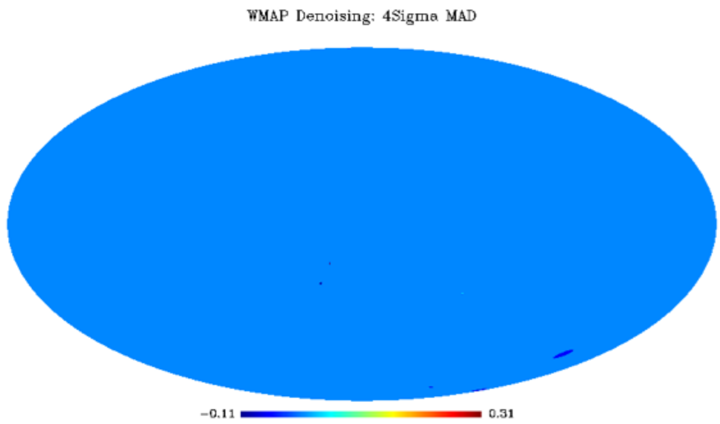

FIG. 7.6 – Débruitage de la carte WMAP à 4σ.

filtrer seulement une vingtaine de cartes de CMB (purement Gaussienne) pour arriver à reproduire un tel spot. Une manière statistique plus saine de détecter les coefficients d'ondelettes est d'utiliser le *False Discovery Rate* introduit au chapitre 5.

Avec le FDR, le seuil de détection est fixé par l'algorithme, et l'utilisateur ne fournit que le paramètre α, qui fixe le pourcentage de fausses détections. En appliquant le filtrage avec FDR, il nous a fallu augmenter le paramètre α jusquà 0.6 (c'est à dire 60% de fausses détections), pour faire apparaitre le *spot*! La conclusion de notre étude est donc que nous ne pouvons PAS confirmer l'existence du *cold spot* avec nos outils.

MCA

L'Analyse en Composante Morphologique (MCA) est un autre outil d'investigation qui peut être utilisé pour extraire un *spot*. En effet, comme nous l'indiquons au chapitre 6, le principe de la MCA est de mettre en compétition deux transformées pour représenter le contenu d'une image. Les structures qui sont bien représentées par une transformée vont être capturées et être ajoutées à la composante associée à cette transformée. Ainsi, nous avons utilisé la transformée en ondelettes isotrope sphérique et les harmoniques sphériques. L'idée est que si le *spot* est vraiment un *spot*, il devrait être mieux représenté par les ondelettes que par les harmoniques sphériques, tandis que si le *spot* n'est qu'un minimum de la fluctuation CMB à une échelle donnée, alors les harmoniques sphériques devraient mieux le capturer. La première expérience que nous avons faite n'a, encore une fois, pas permis d'extraire le *spot*. La carte de CMB a été entièrement représentée par les harmoniques sphériques.

Nous avons alors une conçu une seconde expérience. Nous avons extrait

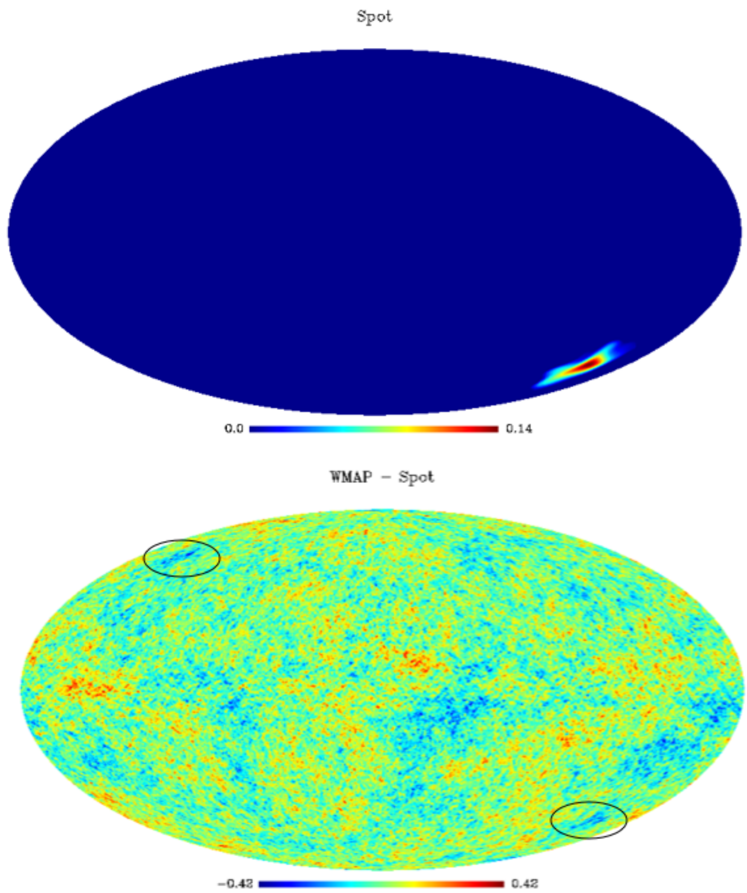

FIG. 7.7 – En haut, *cold spot* extrait de la carte WMAP ; en bas, carte WMAP où on a ajouté un second *cold spot*.

dans les échelles ondelettes (échelles 4,5 et 6) de la carte WMAP-CMB le *spot*, puis nous l'avons ajouté ailleurs dans l'image. La figure 7.7 (en haut) montre le *cold spot* qui a été extrait, et la figure 7.7 (en bas) montre la carte WMAP où on a ajouté le *spot*. Cette carte contient donc maintenant deux spots, très semblables.

Nous avons ensuite lancer le programme de séparation MCA, et le résultat est montré sur la figure 7.8. Le *spot* que nous avons ajouté *à la main* a bien été retrouvé par la MCA, mais pas le vrai *spot*. Cette expérience semblerait montrer que le *spot* serait plutôt un minimum de la fluctuation. On peut évidemment aussi objecter à notre approche que le *spot* extrait et rajouté n'est qu'une estimation du vrai *spot*, et que par conséquence, notre expérience ne signifie rien.

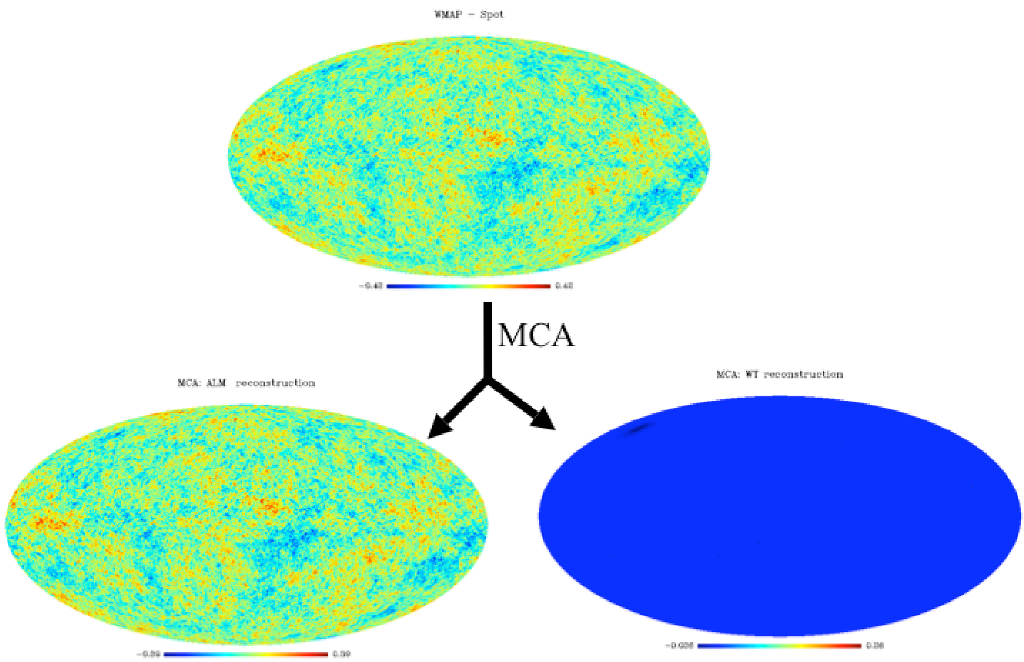

FIG. 7.8 – Décomposition de la carte WMAP+*cold spot* (en haut et également figure 7.7 en bas) en deux composantes, une Gaussienne (en bas à gauche), et une non-Gaussienne (en bas à droite) .

7.6 Conclusion

La validité des hypothèses de Gaussianité et d'isotropie du CMB sont des questions importantes au regard des modèles actuels, qu'il faut confronter à l'analyse des observations du fond diffus cosmologique. Les statistiques d'ordre supérieur sur des coefficients en ondelettes, ou plus généralement sur des coefficients dans un dictionnaire donné, sont particulièrement efficaces pour effectuer cette confrontation. Dans ce chapitre, nous avons montré comment les méthodes multi-échelles sur la sphère permettent de détecter la présence de signaux très faibles noyés dans du bruit. En fonction de la nature de la non-Gaussianité, une décomposition multi-échelles peut être meilleure qu'une autre. C'est pourquoi il est interressant d'utiliser plusieurs transformées plutôt qu'une, pour l'analyse des données. Nous avons montré également que le débruitage est un outil précieux pour le dépouillement de cartes de CMB. En utilisant l'hypothèse d'ergodicité, nous pouvons en effet nous affranchir de lourdes simulations de réalisations de cartes de bruit, et analyser différemment les données,

Toutefois, bien que les travaux de Vielva et al. (2004) tendent à montrer la présencé d'un *cold spot* dans les données WMAP, l'étude de ces mêmes données à l'aide de nos outils ne nous permet pas de confirmer cette détection.

Conclusion

A l'ère de la cosmologie de précision, il est nécessaire d'extraire de manière optimale l'information contenue dans les données. Les transformées multi-échelles géométriques sont clairement des outils indispensables pour analyser les signaux, en particulier ceux qui sont à un très faible rapport signal-sur-bruit.

Nous avons montré dans cette thèse comment construire des transformées géométriques sur la sphère. A la différence des transformées multi-échelles existentes, ces nouvelles transformées sont inversibles et elles peuvent être utilisées dans des applications de filtrage, déconvolution, ou encore de séparation de composantes. Nous avons également montré que le concept de diversité morphologique est très intéressant pour étudier les données PLANCK, et que l'utilisation combinée de plusieurs dictionnaires peut permettre d'obtenir une qualité d'analyse impossible à atteindre en utilisant une seule transformée.

Nous nous sommes intéressés plus particulièrement à deux types d'applications, la restauration et l'inpainting. Le filtrage basée sur les ondelettes et/ou les curvelets permet de bien séparer le signal du bruit, quand le signal est non-stationnaire. Ce type de filtrage est donc très intéressant pour débruiter les cartes de poussières ou de synchrotron. La méthode d'inpainting que nous avons proposée est une alternative aux approches traditionnelles qui consistent à gérer le masque des données manquantes. Nous avons ouvert une nouvelle voie qui est maintenant étudiée par d'autres groupes travaillant dans le cadre du projet PLANCK et en dehors (Pires et al. 2008; McAuley and Caetano 2007; Inoue et al. 2008).

Les codes développés au cours de cette thèse ont été documentés et le logiciel appeé **MRS** (*Multi-Resolution on the Sphere*) est maintenant disponible pour la communauté scientifique aux adresses :
http ://irfu.cea.fr/en/Phocea/Vie_des_labos/Ast/ast_visu.php ?id_ast=895
et
http ://jstarck.free.fr/mrs.html.

Ce logiciel a été intégré dans le *Data Processing Center* (DPC) du projet PLANCK.

Au-delà des multiples développements ayant fait l'objet de cette thèse, il est possible de dresser une liste des différents points qu'il sera intéressant d'explorer par la suite :

- Suppression du bruit dans le CMB : les méthodes de filtrage que nous avons développées sont basées sur du seuillage de coefficients ondelettes ou curvelets. Cette approche est extrêmement pertinente pour la classe de signaux "lisses par morceau". Elle s'applique donc très bien pour la plupart des composantes que PLANCK cherche à reconstruire (SZ, poussière, synchrotron, etc), mais pas du tout pour un champ Gaussian comme le CMB. Le filtre de Wiener actuellement utilisé fait l'hypothèse que le signal et le bruit sont des champs Gaussiens et stationnaires, mais le bruit dans PLANCK n'est pas stationnaire. Une méthode adéquate reste donc à imaginer.
- Polarisation : les cartes de PLANCK sont polarisées. Nous n'avons abordé dans cette thèse que la composante température. Toutes les méthodes que nous proposées doivent donc être étendues pour prendre en compte la polarisation.
- Autres applications : Si le package **MRS** a été concçu dans le cadre du projet PLANCK, son utilité dépasse bien évidemment ce seul projet. Des projets futurs comme EUCLIDE auront besoin d'outils multi-échelles sur la sphère, et **MRS** représente une base parfaite pour de les développements à venir.

Annexe

Mes articles

1. P. Abrial, Y. Moudden, J.L. Starck, M.J. Fadili, J. Delabrouille, and M. Nguyen, "CMB Data Analysis and Sparsity", Statistical Methodology, Vol 5, No 4, pp 289-298, 2008
2. P. P. Abrial, Y. Moudden, J.L. Starck, B. Afeyan, J. Bobin, M.J. Fadili and M. Nguyen, "Morphological Component Analysis and Inpainting on the Sphere : Application in Physics and Astrophysics", Journal of Fourier Analysis and Applications (JFAA), special issue on "Analysis on the Sphere", 13, 6, pp 729-748, 2007.
3. J.-L. Starck, Y. Moudden, P. Abrial and M. Nguyen, "Wavelets, Ridgelets and Curvelets on the Sphere", Astronomy and Astrophysics, 446, 1191-1204, 2006.

Mes Conférences

1. Y. Moudden, P. Abrial, JL Starck, J. Bobin, "Wavelets and Curvelets on the sphere for polarized data", Proceedings of SPIE conference on Signal and Image Processing : Wavelet Applications in Signal and Image Processing XII, SPIE Vol 6701, pp.1601-1615 , August 2007, San Diego, USA.
2. P. Abrial, J.L. Starck, Y. Moudden, M.J. Fadili, J. Delabrouille "CMB inpainting on the Sphere", In proceedings of The Fourth Conference on Astronomical Data Analysis, ADA IV, Marseille (France), 2006.
3. P. Abrial, J.-L. Starck, Y. Moudden, N. Aghanim, O. Forni, "Descriminating secondary from primary non-Gaussian signals", CMB and Physics of the Early Universe, April 18-20, Ishia, 2006.
4. P. Abrial, J.L. Starck, Y. Moudden and M. Nguyen, "Curvelet transform on the Sphere", IEEE International Conference on image Processing 2005, ICIP'05, Volume 1, pp 737-740, Genova, September 11-14, 2005.

5. Y. Moudden, P. Abrial, P. Vielva, J.-B. Melin, J.L. Starck, J.-F. Cardoso, J. Delabrouille, et M. Nguyen, "Independent Component Separation from Incomplete Spherical Data using Wavelets : Application to CMB Analysis", Proceedings of the 4th International Conference on Physics in Signal and Image Processing, PSIP'2005, pp 23-28, Toulouse (France), Jan 31- Feb 2, 2005.
6. Y. Moudden, P. Abrial and J.L. Starck, "Wavelets, ridgelets, curvelets on the sphere and applications " Proceedings of SPIE conference on Signal and Image Processing : Wavelet Applications in Signal and Image Processing XI, SPIE Vol 5914, pp 217-231, August 2005, San Diego, USA.
7. J.-L. Starck and P. Abrial, "Non Gaussianity detection and discrimination using several multiscale transforms", Workshop on Nongaussianity in Cosmology, Trieste, July 24-28, 2006.
8. J.-F. Cardoso, P. Abrial, Y. Moudden, J.-L. Starck and J. Delabrouille, "Statistiques direction multipole pour la separation de composantes astrophysiques dans le fonds de rayonnement cosmologique", Proceedings of GRETSI 2005, Louvain-la-Neuve, Belgique, Septembre 2005.

Logiciel

MRS (Multi-Resolution on the Sphere) : Le code et une documentation complète sont disponibles à l'adresse
http ://irfu.cea.fr/en/Phocea/Vie_des_labos/Ast/ast_visu.php ?id_ast=895
et
http ://jstarck.free.fr/mrs.html.

Bibliographie

Abrial, P., Moudden, Y., Starck, J., Bobin, J., Fadili, J., Afeyan, B., and Nguyen, M. : 2007, Morphological Component Analysis and Inpainting on the sphere : Application in Physics and Astrophysics, *Journal of Fourier Analysis and Applications (JFAA)* **13(6)**, 729

Abrial, P., Moudden, Y., Starck, J., Fadili, M., Delabrouille, J., and Nguyen, M. : 2008, CMB data analysis and sparsity, *Statistical Methodology* **5**, 289

Afeyan, B., Won, K., Starck, J. L., Stevens, R., Mapoles, E., Johnson, M., and Haan, S. : 2006, Modem : Morphological diversity extraction method to identify, classify and characterize surface defects of spherical icf targets via afm and phase shifting diffraction interferometry, in *Proceedings of the 17th Target Fabrication Meeting, San Diego, CA*, pp 1–5

Aghanim, N. and Forni, O. : 1999, Searching for the non-Gaussian signature of the CMB secondary anisotropies, *Astronomy and Astrophysics* **347**, 409

Aghanim, N., Kunz, M., Castro, P. G., and Forni, O. : 2003, Non-Gaussianity : Comparing wavelet and Fourier based methods, *Astronomy and Astrophysics* **406**, 797

Alliney, S. : 1992, Digital filters as absolute norm regularizers, *IEEE Transactions on Signal Processing* **40(6)**, 1548

Alliney, S. and Ruzinsky, S. : 1994, An algorithm for the minimization of mixed l1 and l2 norms with application to bayesian estimation, *IEEE Transactions on Signal Processing* **42(3)**, 618

Antoine, J., Demanet, L., Jacques, L., and Vandergheynst, P. : 2002, Wavelets on the sphere : Implementation and approximation, *Appl. Comput. Harmon. Anal.* **13**, 177

Antoine, J.-P. : 1999, The 2-D wavelet transform, physical applications and generalizations, in *Wavelets in Physics*, pp 23–+

Ashdown, M. A. J., Baccigalupi, C., Balbi, A., Bartlett, J. G., Borrill, J., Cantalupo, C., de Gasperis, G., Górski, K. M., Hivon, E., Keihänen, E., Kurki-Suonio, H., Lawrence, C. R., Natoli, P., Poutanen, T., Prunet, S., Reinecke, M., Stompor, R., Wandelt, B., and The Planck CTP Working Group : 2007, Making sky maps from Planck data, *AA* **467**, 761

Atzeni, S. and ter Vehn, J. M. : 2004, *The Physics of Inertial Fusion*, Oxford University Press

Averbuch, A., Coifman, R., Donoho, D., Israeli, M., and Waldén, J. : 2001, Fast Slant Stack : A notion of Radon transform for data in a cartesian grid which is rapidly computible, algebraically exact, geometrically faithful and invertible, *SIAM J. Sci. Comput.*

Ballester, C., Bertalmio, M., Caselles, V., Sapiro, G., and Verdera, J. : August 2001, Filling-in by joint interpolation of vector fields and grey levels, *IEEE Trans. Image Processing* **10**, 1200

Banday, A. J., Zaroubi, S., and Górski, K. M. : 2000, On the Non-Gaussianity Observed in the COBE Differential Microwave Radiometer Sky Maps, *Astrophysical Journal* **533**, 575

Barreiro, R. B. and Hobson, M. P. : 2001, The discriminating power of wavelets to detect non-Gaussianity in the cosmic microwave background, *MNRAS* **327**, 813

Barreiro, R. B., Hobson, M. P., Lasenby, A. N., Banday, A. J., Górski, K. M., and Hinshaw, G. : 2000, Testing the Gaussianity of the COBE DMR data with spherical wavelets, *MNRAS* **318**, 475

Barreiro, R. B., Martínez-González, E., and Sanz, J. L. : 2001, Geometrical estimators as a test of Gaussianity in the cosmic microwave background, *MNRAS* **322**, 411

Benjamini, Y. and Hochberg, Y. : 1995, Controlling the false discovery rate - a practical and powerful approach to multiple testing, *J. R. Stat. Soc. B* **57**, 289

Bertalmio, M., Bertozzi, A., , and Sapiro, G. : 2001, Navier-stokes,fluid dynamics, and image and video inpainting, in *in Proc.IEEE Computer Vision and Pattern Recognition (CVPR)*

Bertalmio, M., Sapiro, G., Caselles, V., , and Ballester, C. : July 2000, Image inpainting, *Comput. Graph.(SIGGRAPH 2000)* pp 417–424

Bobin, J., Moudden, Y., Starck, J.-L., and Elad, M. : 2006a, Morphological diversity and source separation, *IEEE Trans. on Signal Processing* **13(7)**, 409

Bobin, J., Moudden, Y., Starck, J.-L., Fadili, J., and Aghanim, N. : 2006b, SZ and CMB reconstruction using generalized morphological component analysis, in *ADA IV*, Proceedings of the 4th Conference on Astronomical Data Analysis, Marseille, (France), September 18-20, 2006.

Bobin, J., Starck, J.-L., Fadili, J., and Moudden, Y. : 2007a, Sparsity and morphological diversity in blind source separation, *IEEE Transactions on Image Processing* **16(11)**, 2662

Bobin, J., Starck, J.-L., Fadili, J., Moudden, Y., and Donoho, D. : 2007b, Morphological component analysis : An adaptive thresholding strategy, *IEEE Transactions on Image Processing* **16(11)**, 2675

Bogdanova, I., Vandergheynst, P., Antoine, J.-P., Jacques, L., and Mrovidone, M. : 2005, Stereographic wavelet frames on the sphere, *Applied and Computational Harmonic Analysis* **19(2)**, 223

Bornemann, F. and März, T. : 2006, *Fast Image Inpainting Based on Coherence Transport*, Technical report, Technische Universität München

Bouchet, R. and Gispert, R. : 1999, Foregrounds and cmb experiments : I. semi-analytical estimates of contamination, *New Astronomy* 4(443)

Bromley, B. C. and Tegmark, M. : 1999, Is the Cosmic Microwave Background Really Non-Gaussian ?, *Ap. J. Letters* **524**, L79

Bruce, A., Sardy, S., and Tseng, P. : 1998, Block coordinate relaxation methods for nonparametric signal de-noising, *Proceedings of the SPIE - The International Society for Optical Engineering* **3391**, 75

Bruckstein, A. and Elad, M. : 2002, A generalized uncertainty principle and sparse representation in pairs of $\mathbf{r}^n$ bases, *IEEE Transactions on Information Theory* **48**, 2558

Candès, E. and Donoho, D. : 1999a, *Curvelets*, Technical report, Statistics, Stanford University

Candès, E. and Donoho, D. : 1999b, Ridgelets : the key to high dimensional intermittency ?, *Philosophical Transactions of the Royal Society of London A* **357**, 2495

Candès, E., Romberg, J., and Tao, T. : 2004, Robust uncertainty principles : Exact signal reconstruction from highly incomplete frequency information, *IEEE Transactions on Information Theory* **52**, 489

Candès, E. and Tao, T. : 2004, Near optimal signal recovery from random projections : Universal encoding strategies ?, *IEEE Transactions on Information Theory* **52**, 5406

Candès, E. and Tao, T. : 2005, *Stable Signal Recovery from noisy and incomplete observations*, Technical report, CalTech, Applied and Computational Mathematics

Candès, E. J. and Donoho, D. L. : 1999c, Curvelets – a surprisingly effective nonadaptive representation for objects with edges, in A. Cohen, C. Rabut, and L. Schumaker (eds.), *Curve and Surface Fitting : Saint-Malo 1999*, Vanderbilt University Press, Nashville, TN

Candès, E. J. and Donoho, D. L. : 2000, *Curvelets and curvilinear integrals*, Technical report, Department of Statistics, Stanford University

Cardoso, J.-F. and Souloumiac, A. : 1996, Jacobi angles for simultaneous diagonalization, *SIAM Journal of Matrix Analysis and Applications* **17(1)**, 161

Cayón, L., Sanz, J. L., Barreiro, R. B., Martínez-González, E., Vielva, P., Toffolatti, L., Silk, J., Diego, J. M., and Argüeso, F. : 2000, Isotropic wavelets : a powerful tool to extract point sources from cosmic microwave background maps, *MNRAS* **315**, 757

Cayón, L., Sanz, J. L., Martínez-González, E., Banday, A. J., Argüeso, F., Gallegos, J. E., Górski, K. M., and Hinshaw, G. : 2001a, Spherical Mexican hat wavelet : an application to detect non-Gaussianity in the COBE-DMR maps, *MNRAS* **326**, 1243

Cayón, L., Sanz, J. L., Martínez-González, E., Banday, A. J., Argüeso, F., Gallegos, J. E., Górski, K. M., and Hinshaw, G. : 2001b, Spherical Mexican hat wavelet : an application to detect non-Gaussianity in the COBE-DMR maps, *MNRAS* **326**, 1243

Chan, T. and Shen, J. : July 2001, Local inpainting models and tv inpainting, *SIAM J. Appl. Math.* **62**, 1019

Chan, T. F., Ng, M. K., Yau, A. C., and Yip, A. M. : 2006, *Superresolution Image Reconstruction Using Fast Inpainting Algorithms*, Technical report, UCLA CAM

Chaux, C., Combettes, P., Pesquet, J.-C., and Wajs, V. : 2007, A variational formulation for frame based inverse problems, *Inverse Problems* **23**, 1495

Chen, S., Donoho, D., and Saunder, M. : 1998, Atomic decomposition by basis pursuit, *SIAM Journal on Scientific Computing* **20**, 33

Coifman, R. and Donoho, D. : 1995, Translation invariant de-noising, in A. Antoniadis and G. Oppenheim (eds.), *Wavelets and Statistics*, pp 125–150, Springer-Verlag

Combettes, P. L. and Wajs, V. R. : 2005, Signal recovery by proximal forward-backward splitting, *SIAM Journal on Multiscale Modeling and Simulation* **4(4)**, 1168

Cooray, A. : 2001, Non-gaussian aspects of thermal and kinetic sunyaevzel'dovich effects, *Physical Review D* **64**, 3514

Crittenden, R. G. : 2000, Igloo Pixelizations of the Sky, *Astrophysical Letters Communications* **37**, 377

Cruz, M., Cayón, L., Martínez-González, E., Vielva, P., and Jin, J. : 2007, The Non-Gaussian Cold Spot in the 3 Year Wilkinson Microwave Anisotropy Probe Data, *ApJ* **655**, 11

Cruz, M., Martínez-González, E., Vielva, P., and Cayón, L. : 2005, Detection of a non-Gaussian spot in WMAP, *MNRAS* **356**, 29

Deans, S. : 1983, *The Radon Transform and Some of Its Applications*, Wiley

Delabrouille, J., Cardoso, J.-F., and Patanchon, G. : 2003, Multi–detector multi–component spectral matching and applications for CMB data analysis, *Monthly Notices of the Royal Astronomical Society* **346(4)**, 1089, to appear, also available as http ://arXiv.org/abs/astro-ph/0211504

Donoho, D. : 1993, Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data, in A. M. Society (ed.), *Proceedings of Symposia in Applied Mathematics*, Vol. 47, pp 173–205

Donoho, D. : 2006a, For most large underdetermined systems of linear equations, the minimal ℓ_1 solution is also the sparsest solution, *Comm. Pure Appl. Math.* **59(7)**, 907

Donoho, D. and Duncan, M. : 2000, Digital curvelet transform : strategy, implementation and experiments, in H. Szu, M. Vetterli, W. Campbell, and J. Buss (eds.), *Proc. Aerosense 2000, Wavelet Applications VII*, Vol. 4056, pp 12–29, SPIE

Donoho, D. and Elad, M. : 2003a, Optimally sparse representation in general (non-orthogonal) dictionaries via ℓ^1 minimization, *Proc. Nat. Aca. Sci.* **100**, 2197

Donoho, D., Elad, M., and Temlyakov, V. : 2006a, Stable recovery of sparse overcomplete representations in the presence of noise, *IEEE Transactions on Information Theory* **52(1)**, 6

Donoho, D. and Flesia, A. : 2002, Digital Ridgelet Transform Based on True Ridge Functions, in J. Schmeidler and G. Welland (eds.), *Beyond Wavelets*, Academic Press

Donoho, D. and Huo, X. : 2001, Uncertainty principles and ideal atomic decomposition, *IEEE Transactions on Information Theory* **47(7)**, 2845

Donoho, D. and Jin, J. : 2004, Higher criticism for detecting sparse heterogeneous mixtures, *Ann. Statist.* **32(3)**, 962

Donoho, D. and Johnstone, I. : 1994, Ideal spatial adaptation via wavelet shrinkage, *Biometrika* **81**, 425

Donoho, D., Tsaig, Y., Drori, I., and Starck, J.-L. : 2006b, Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit, *IEEE Transactions on Information Theory*, submitted

Donoho, D. L. : 2006b, Compressed sensing, *IEEE Transactions on Information Theory* **52(4)**, 1289

Donoho, D. L. and Elad, M. : 2003b, Maximal sparsity representation via l_1 minimization, *the Proc. Nat. Aca. Sci.* **100**, 2197

Doroshkevich, A. G., Naselsky, P. D., Verkhodanov, O. V., Novikov, D. I., Turchaninov, V. I., Novikov, I. D., Christensen, P. R., and Chiang, L.-Y. : 2005, Gauss-legendre sky pixelization (glesp) scheme for cmb maps, *International Journal of Modern Physics D* **14(2)**, 275, also available at http ://arXiv.org/abs/astro-ph/0305537

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. : 2004, Least angle regression, *Annals of Statistics* **32(2)**, 407

Elad, M. and Bruckstein, A. : 2002, A generalized uncertainty principle and sparse representation in pairs of bases, *IEEE Transactions on Information Theory* **48**, 2558

Elad, M., Matalon, B., Shtok, J., and Zibulevsky, M. : 2007, A Wide-Angle View at Iterated Shrinkage Algorithms, in *SPIE Wavelets XII*, San Diego, CA

Elad, M., Starck, J.-L., Donoho, D., and Querre, P. : 2006, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA), *Journal on Applied and Computational Harmonic Analysis* **19**, 340

Elad, M., Starck, J.-L., Querre, P., and Donoho, D. : 2005, Simultaneous Cartoon and Texture Image Inpainting using Morphological Component Analysis (MCA), *J. on Applied and Computational Harmonic Analysis* **19(3)**, 340

Escalera, E. and MacGillivray, H. T. : 1995, Topology in galaxy distributions : method for a multi-scale analysis. A use of the wavelet transform., *Astronomy and Astrophysics* **298**, 1

Fadili, M. and Starck, J.-L. : 2008, Curvelets and ridgelets, in *Encyclopedia of Complexity and System Science*, Springer, in press

Forni, O. and Aghanim, N. : 1999, Searching for non-gaussianity : Statistical tests, *Astronomy and Astrophysics, Supplement Series* **137**, 553

Freeden, W. and Maier, T. : 2002, On multiscale denoising of spherical functions : Basic theory and numerical aspects, *Electronic Transactions on Numerical Analysis (ETNA)* **14**, 40

Freeden, W., Maier, T., and Zimmermann, S. : 2003, A survey on wavelet methods for (geo)applications, *Revista Mathematica Complutense* **16(1)**, 277

Fuchs, J. : 2005, Recovery of exact sparse representations in the presence of bounded noise, *IEEE Trans. On Information Theory* **51**, 3601

Genova-Santos, R., Rubino-Martin, J. A., Rebolo, R., Battye, R. A., Blanco, F., Davies, R. D., Davis, R. J., Franzen, T., Grainge, K., Hobson, M. P., Lasenby, A., Padilla-Torres, C. P., Pooley, G. G., Saunders, R. D. E., Scaife, A., Scott, P. F., Titterington, D., Tucci, M., and Watson, R. A. : 2008, Observations of the Corona Borealis supercluster with the super-extended Very Small Array : further constraints on the nature of the non-Gaussian CMB cold spot, *ArXiv e-prints*

Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., and Bartelmann, M. : 2005, HEALPix : A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, *Astrophysical Journal* **622**, 759

Gorski, K. M., Wandelt, B. D., Hansen, F. K., Hivon, E., and Banday, A. J. : 1999, The HEALPix Primer, *ArXiv Astrophysics e-prints*

Gribonval, R. and Nielsen, M. : 2003, Sparse representations in unions of bases, *IEEE Transactions on Information Theory* **49(12)**, 3320

Hobson, M. P., Jones, A. W., and Lasenby, A. N. : 1999, Wavelet analysis and the detection of non-Gaussianity in the cosmic microwave background, *MNRAS* **309**, 125

Holschneider, M. : 1996, Wavelet analysis on the sphere, *J. Math. Phys.* **37(8)**, 4156

Hopkins, A. M., Miller, C. J., Connolly, A. J., Genovese, C., Nichol, R. C., and Wasserman, L. : 2002, A New Source Detection Algorithm Using the False-Discovery Rate, *AJ* **123**, 1086

Inoue, K. T., Cabella, P., and Komatsu, E. : 2008, Harmonic inpainting of the cosmic microwave background sky : Formulation and error estimate, *Phys. Rev. D* **77(12)**, 123539

Jewell, J. : 2001, A Statistical Characterization of Galactic Dust Emission as a Non-Gaussian Foreground of the Cosmic Microwave Background, *Astrophysical Journal* **557**, 700

Jin, J., Starck, J.-L., Donoho, D., Aghanim, N., and Forni, O. : 2005, Cosmological non-gaussian signatures detection : Comparison of statistical tests, *Eurasip Journal* **15**, 2470

Jungman, G. : 1996, Cosmological parameter determination with microwave backgroung maps, *Phys. Rev. D* **54**, 1332

Komatsu, E., Kogut, A., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Limon, M., Meyer, S. S., Page, L., Spergel, D. N., Tucker, G. S., Verde, L., Wollack, E., and Wright, E. L. : 2003, First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Tests of Gaussianity, *Astrophysical Journal, Supplement Series* **148**, 119

Kunszt, P. Z., Szalay, A. S., and Thakar, A. R. : 2001, The Hierarchical Triangular Mesh, in A. J. Banday, S. Zaroubi, and M. Bartelmann (eds.), *Mining the Sky*, pp 631–+

Kunz, M., Banday, A. J., Castro, P. G., Ferreira, P. G., and Górski, K. M. : 2001, The Trispectrum of the 4 Year COBE DMR Data, *Ap. J. Letters* **563**, L99

Lachièze-Rey, M. : 2005, *Initiation à la cosmologie*, Dunod, 4ème edition

Lindl, J. : 1997, *Inertial Confinement Fusion : The Quest for Ignition and Energy Gain Using Indirect Drive*, AIP Press

López-Caniego, M., Herranz, D., Barreiro, R. B., and Sanz, J. L. : 2005, Filter design for the detection of compact sources based on the Neyman-Pearson detector, *MNRAS* pp 368–+

Mallat, S. and Zhang, Z. : 1993, Matching pursuits with time-frequency dictionaries, *IEEE Transactions on Signal Processing* **41(12)**, 3397

Martínez, V. J., Starck, J.-L., Saar, E., Donoho, D. L., Reynolds, S. C., de la Cruz, P., and Paredes, S. : 2005, Morphology of the Galaxy Distribution from Wavelet Denoising, *Astrophysical Journal* **634**, 744

Masnou, S. and Morel, J. : 1998, Level lines based disocclusion, in *IEEE International Conference on Image Processing*, Vol. III, pp 259–263

Masnou, S. and Morel, J. : 2002, Disocclusion : A variational approach using level lines, *IEEE Trans. Image Process.* **11(2)**, 68

Massey, R., Rhodes, J., Ellis, R., Scoville, N., Leauthaud, A., Finoguenov, A., Capak, P., Bacon, D., Aussel, H., Kneib, J.-P., Koekemoer, A., McCracken, H., Mobasher, B., Pires, S., Refregier, A., Sasaki, S., Starck, J.-L., Taniguchi, Y., Taylor, A., and Taylor, J. : 2007, Dark matter maps reveal cosmic scaffolding, *Nature* **445**, 286

McAuley, J. J. and Caetano, T. S. : 2007, High-Order Nonparametric Belief-Propagation for Fast Image Inpainting, *ArXiv e-prints*

McEwen, J. D., Hobson, M. P., Lasenby, A. N., and Mortlock, D. J. : 2008, A high-significance detection of non-Gaussianity in the WMAP 5-yr data using directional spherical wavelets, *MNRAS* **388**, 659

McEwen, J. D., Hobson, M. P., Mortlock, D. J., and Lasenby, A. N. : 2005, Fast directional continuous spherical wavelet transform algorithms, *ArXiv Astrophysics e-prints*

McEwen, J. D., Vielva, P., Wiaux, Y., Barreiro, R. B., Cayon, L., Hobson, M. P., Lasenby, A. N., Martinez-Gonzalez, E., and Sanz, J. L. : 2007, Cosmological applications of a wavelet analysis on the sphere, *Journal of Fourier Analysis and Applications* **13**, 495

M.Elad : 2006, Why simple shrinkage is still relevant for redundant representations ?, *IEEE Transactions on Information Theory* **52(12)**, 5559

Miller, C. J., Genovese, C., Nichol, R. C., Wasserman, L., Connolly, A., Reichart, D., Hopkins, A., Schneider, J., and Moore, A. : 2001, Controlling the False-Discovery Rate in Astrophysical Data Analysis, *AJ* **122**, 3492

Moudden, Y., Cardoso, J.-F., Starck, J.-L., and Delabrouille, J. : 2005, Blind component separation in wavelet space : Application to CMB analysis, *Eurasip Journal on Applied Signal Processing* **15**, 2437

Mukherjee, P. and Wang, Y. : 2003, Model-independent Reconstruction of the Primordial Power Spectrum from Wilkinson Microwave Anistropy Probe Data, *Astrophysical Journal* **599**, 1

Mukherjee, P. and Wang, Y. : 2004, Wavelets and Wilkinson Microwave Anisotropy Probe Non-Gaussianity, *Astrophysical Journal* **613**, 51

Novikov, D., Schmalzing, J., and Mukhanov, V. F. : 2000, On non-Gaussianity in the cosmic microwave background, *Astronomy and Astrophysics* **364**, 17

Patanchon, G., Cardoso, J.-F., Delabrouille, J., and Vielva, P. : 2005, Cosmic microwave background and foregrounds in Wilkinson Microwave Anisotropy Probe first-year data, *MNRAS* **364**, 1185

Phillips, N. G. and Kogut, A. : 2001, Statistical Power, the Bispectrum, and the Search for Non-Gaussianity in the Cosmic Microwave Background Anisotropy, *Astrophysical Journal* **548**, 540

Pires, S., Juin, J. B., Yvon, D., Moudden, Y., Anthoine, S., and Pierpaoli, E. : 2006, Sunyaev-Zel'dovich cluster reconstruction in multiband bolometer camera surveys, *A&A* **455**, 741

Pires, S., Starck, J. ., Amara, A., Teyssier, R., Refregier, A., and Fadili, J. : 2008, FASTLens (FAst STatistics for weak Lensing) : Fast method for Weak Lensing Statistics and map making, *ArXiv e-prints*

Plumbley, M. D. : 2006, Recovery of sparse representations by polytope faces pursuit, in *ICA2006*, pp 206–213, Proceedings of the 4th International Conference on Independent Component Analysis and Blind Source Separation, Charleston, SC, USA, 5-8 March 2006.

R.Bouchet and R.Gispert : 1999, Foregrounds and cmb experiments i. : Semi-analytical estimates of contamination, *New Astronomer* **4(443)**, 2279

Refregier, A. : 2003, Weak Gravitational Lensing by Large-Scale Structure, *ARA&A* **41**, 645

Rocha, G., Cayón, L., Bowen, R., Canavezes, A., Silk, J., Banday, A. J., and Górski, K. M. : 2004, Topology of the Universe from COBE-DMR - a wavelet approach, *MNRAS* **351**, 769

Sanz, J. L., Herranz, D., and Martínez-Gónzalez, E. : 2001, Optimal Detection of Sources on a Homogeneous and Isotropic Background, *Astrophysical Journal* **552**, 484

Schröder, P. and Sweldens, W. : 1995, Spherical wavelets : Efficiently representing functions on the sphere, *Computer Graphics Proceedings (SIGGRAPH 95)* pp 161–172

Shandarin, S. F. : 2002, Testing non-Gaussianity in cosmic microwave background maps by morphological statistics, *MNRAS* **331**, 865+

Slezak, E., de Lapparent, V., and Bijaoui, A. : 1993, Objective detection of voids and high density structures in the first CfA redshift survey slice, *Astrophysical Journal* **409**, 517

Smith, K. M. and Huterer, D. : 2008, No evidence for the cold spot in the NVSS radio survey, *ArXiv e-prints*

Starck, J.-L., Aghanim, N., and Forni, O. : 2004a, Detecting cosmological non-gaussian signatures by multi-scale methods, *Astronomy and Astrophysics* **416**, 9

Starck, J.-L., Candès, E., and Donoho, D. : 2002, The curvelet transform for image denoising, *IEEE Transactions on Image Processing* **11(6)**, 131

Starck, J.-L., Candes, E., and Donoho, D. : 2003a, Astronomical image representation by the curvelet tansform, *AA* **398**, 785

Starck, J.-L., Donoho, D. L., and Candès, E. J. : 2001, Very high quality image restoration by combining wavelets and curvelets, in *Proc. SPIE Vol. 4478, Wavelets : Applications in Signal and Image Processing IX, Andrew F. Laine ; Michael A. Unser ; Akram Aldroubi ; Eds.*, pp 9–19

Starck, J.-L., Elad, M., and Donoho, D. : 2004b, Redundant multiscale transforms and their application for morphological component analysis, *Advances in Imaging and Electron Physics* 132

Starck, J.-L., Fadili, J., and Murtagh, F. : 2007, The undecimated wavelet decomposition and its reconstruction, *IEEE Transactions on Image Processing* **16**, 297

Starck, J.-L. and Fadili, M. J. : 2007, Numerical issues when using wavelets, in *Encyclopedia of Complexity and Systems Science (to appear)*, Springer

Starck, J.-L., Martinez, V., Donoho, D., Levi, O., Querre, P., and Saar, E. : 2005, Analysis of the spatial distribution of galaxies by multiscale methods, *Eurasip Journal on Applied Signal Processing* **15**, 2455

Starck, J.-L., Moudden, Y., Abrial, P., and Nguyen, M. : 2006, Wavelets, ridgelets and curvelets on the sphere, *Astronomy and Astrophysics* **446**, 1191

Starck, J.-L. and Murtagh, F. : 2006, *Astronomical Image and Data Analysis*, Astronomical image and data analysis, by J.-L. Starck and F. Murtagh. Astronomy and astrophysics library. Berlin : Springer, 2006

Starck, J.-L., Murtagh, F., Candes, E., and Donoho, D. : 2003b, Gray and color image contrast enhancement by the curvelet transform, *IEEE Transactions on Image Processing* **12(6)**, 706

Starck, J.-L., Nguyen, M., and Murtagh, F. : 2003c, Wavelets and curvelets for image deconvolution : a combined approach, *Signal Processing* **83(10)**, 2279

Starck, J.-L., Pires, S., and Réfrégier, A. : 2006, Weak lensing mass reconstruction using wavelets, *A&A* **451**, 1139

Sunyaev, R. A. and Zeldovich, I. B. : 1980, Microwave background radiation as a probe of the contemporary structure and history of the universe, *Annual Review of Astronomy and Astrophysics* **18**, 537

Tegmark, M. : 1996, An Icosahedron-Based Method for Pixelizing the Celestial Sphere, *ApJ* **470**, L81+

Tenorio, L., Jaffe, A. H., Hanany, S., and Lineweaver, C. H. : 1999, Applications of wavelets to the analysis of cosmic microwave background maps, *MNRAS* **310**, 823

Toft, P. : 1996, *The Radon Transform - Theory and Implementation*, *Ph.D. thesis*, Department of Mathematical Modelling, Technical University of Denmark

Tropp, T. : 2006, Just relax : Convex programming methods for subset selection and sparse approximation, *IEEE Transactions on Information Theory* **52(3)**, 1030

Verde, L., Wang, L., Heavens, A. F., and Kamionkowski, M. : 2000, Large-scale structure, the cosmic microwave background and primordial non-Gaussianity, *MNRAS* **313**, 141

Verdera, J., Caselles, V., Bertalmio, M., and Sapiro, G. : 2003, Inpainting surface holes, *ICIP* **2**, 903

Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., and Cayón, L. : 2004, Detection of Non-Gaussianity in the Wilkinson Microwave Anisotropy Probe First-Year Data Using Spherical Wavelets, *Astrophysical Journal* **609**, 22

Vielva, P., Wiaux, Y., Martínez-González, E., and Vandergheynst, P. : 2006, Steerable wavelet analysis of CMB structures alignment, *New Astronomy Review* **50**, 880

Vio, R., Tenorio, L., and Wamsteker, W. : 2002, On optimal detection of point sources in CMB maps, *Astronomy and Astrophysics* **391**, 789

Wiaux, Y., Jacques, L., and Vandergheynst, P. : 2005, Correspondence Principle between Spherical and Euclidean Wavelets, *Astrophysical Journal* **632**, 15

Wiaux, Y., Jacques, L., Vielva, P., and Vandergheynst, P. : 2006a, Fast Directional Correlation on the Sphere with Steerable Filters, *ApJ* **652**, 820

Wiaux, Y., McEwen, J. D., Vandergheynst, P., and Blanc, O. : 2008, Exact reconstruction with directional wavelets on the sphere, *MNRAS* **388**, 770

Wiaux, Y., Vielva, P., Martínez-González, E., and Vandergheynst, P. : 2006b, Global Universe Anisotropy Probed by the Alignment of Structures in the Cosmic Microwave Background, *Physical Review Letters* **96(15)**, 151303

Yamada, I. : 2001, The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings, in D. Butnariu, Y. Censor, and S. Reich (eds.), *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, Elsevier