

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES

Spécialité : Physique subatomique et astroparticules

Arrêté ministériel : 25 mai 2016

Présentée par

Florian KÉRUZORÉ

Thèse dirigée par **Frédéric MAYET, Professeur,**
Université Grenoble Alpes

préparée au sein du **Laboratoire de Physique Subatomique et Cosmologie (LPSC)**

dans l'**École Doctorale de Physique de Grenoble**

Cosmologie à partir des observations Sunyaev-Zeldovich d'amas de galaxies avec NIKA2

Manuscrit proposé pour une soutenance le **27 septembre 2021**, devant le jury composé de :

Dominique BOUTIGNY

Directeur de recherche, LAPP Annecy, Examinateur

Delphine HARDIN

Professeure, LPNHE Paris, Examinatrice

Sophie MAUROGORDATO

Directrice de recherche, OCA Nice, Rapporteuse

Jean-Baptiste MELIN

Chercheur, CEA Saclay, Rapporteur

Pierre SALATI

Professeur émérite, LAPTh Annecy, Examinateur

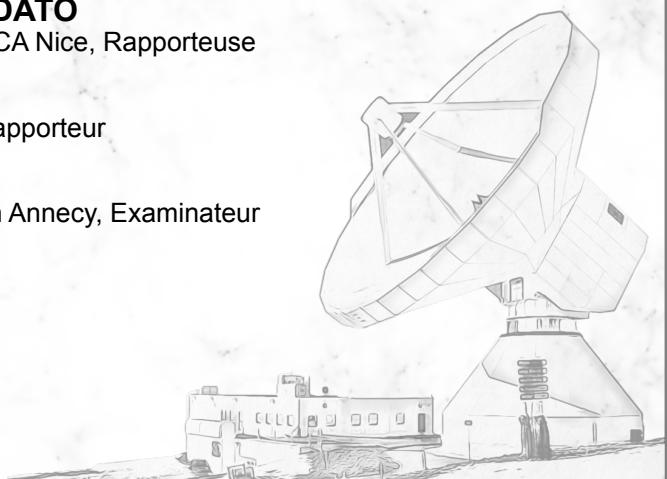


Image de couverture : toile cosmique dans la simulation Millenium, crédit : Springel *et al.* (2005) [1],

<https://wwwmpa.mpa-garching.mpg.de/galform/virgo/millennium/>

Télescope de 30 mètres de l'IRAM

Table des matières

Introduction	7
1 Évolution de l'Univers et des grandes structures	11
1.1 Le modèle standard de la cosmologie	12
1.1.1 Les équations de Friedmann	12
1.1.2 Expansion de l'Univers, accélération et énergie sombre	13
1.1.3 Histoire thermique de l'Univers	15
1.2 Origine des grandes structures	16
1.2.1 Limitations du modèle standard de la cosmologie	16
1.2.2 Fluctuations primordiales et inflation	17
1.3 Évolution des grandes structures	19
1.3.1 Croissance des inhomogénéités de densité	19
1.3.2 Spectre de puissance de la distribution de matière	21
1.3.3 Fonction de masse et distribution de halos	22
1.4 Conclusions	23
2 Sonder la cosmologie à l'aide d'amas de galaxies	25
2.1 Composition et masse des amas de galaxies	26
2.1.1 Définition d'un amas de galaxies	26
2.1.2 Masse et rayon des amas de galaxies	26
2.1.3 Le biais hydrostatique	28
2.2 Observations des amas de galaxies	29
2.2.1 Observations en visible et infrarouge	29
2.2.2 Observations en X	31
2.2.3 Observations en millimétrique : l'effet Sunyaev-Zeldovich	32
2.2.4 Relevés d'amas de galaxies par effet SZ	37
2.3 Cosmologie par comptage d'amas de galaxies	40
2.3.1 Fonction de masse et comptage d'amas	40
2.3.2 Relation d'échelle masse-observable	42
2.4 État de l'art et perspectives	44
2.4.1 Relevés d'amas et résultats cosmologiques actuels	44

2.4.2	Relevés futurs	45
2.4.3	Suivis dédiés d'échantillons d'amas	46
2.4.4	Conclusions	47
3	NIKA2 et son Grand Programme SZ	49
3.1	Le télescope de 30 mètres de l'IRAM	50
3.1.1	Localisation et conditions météorologiques	50
3.1.2	Caractéristiques instrumentales	51
3.2	La caméra NIKA2	52
3.2.1	Détecteurs à inductance cinétique	52
3.2.2	Éléments clés de la caméra NIKA2	54
3.3	Observations avec NIKA2	56
3.3.1	Balayage du ciel et données ordonnées en temps	56
3.3.2	Reconstruction du plan focal	57
3.3.3	Pointage et focalisation	57
3.3.4	Performances instrumentales de NIKA2	59
3.4	Le grand programme SZ de NIKA2	63
3.4.1	La caméra NIKA2 : un instrument idéal pour la mesure de l'effet SZ	63
3.4.2	Échantillon du grand programme SZ	65
3.4.3	Objectifs scientifiques du grand programme SZ	67
3.5	La base de données du grand programme SZ de NIKA2	72
3.5.1	Contexte	72
3.5.2	Contenu et interface de la base de données	73
3.5.3	Développements futurs	74
4	Construction de cartes de l'effet SZ avec NIKA2	77
4.1	Composantes des données en temps	78
4.2	Soustraction du bruit corrélé	80
4.2.1	Estimation de mode commun	81
4.2.2	Masquage de source et analyse itérative	84
4.3	Des données en temps aux cartes du ciel	86
4.3.1	Projection et cartographie du signal d'intérêt	86
4.3.2	Cartes de couverture et de bruit	87
4.3.3	Coaddition de cartes	88
4.4	Performance de la décorrélation	89
4.4.1	Bruit corrélé résiduel	90
4.4.2	Filtrage du signal	91
4.5	Contamination par des sources ponctuelles	93
4.5.1	Données d'entrée	94
4.5.2	Mesures de coordonnées et flux à 260 GHz	95

4.5.3	Ajustement du spectre de chaque source	96
4.5.4	Inférence du flux à 150 GHz	97
4.6	Conclusions et perspectives	98
5	Pipeline pour la mesure des propriétés thermodynamiques du milieu intra-amas	101
5.1	Modélisation du signal SZ	103
5.1.1	Profil de pression et effet Sunyaev-Zeldovich	103
5.1.2	Modèle de carte de brillance de surface SZ	104
5.1.3	Contamination par des sources ponctuelles	105
5.1.4	Signal SZ intégré	106
5.2	Ajustement par méthode Monte Carlo à Chaînes de Markov	106
5.2.1	Fonction de vraisemblance	106
5.2.2	Distribution <i>a priori</i> des paramètres	107
5.2.3	Échantillonnage de la distribution de probabilités	107
5.2.4	Évaluation des propriétés thermodynamiques	110
5.3	Validation de PANCO2 sur des amas simulés	112
5.3.1	Simulations d'amas simples	112
5.3.2	Optimisation de la distribution radiale	118
5.3.3	La simulation MUSIC et l'échantillon jumeau du grand programme SZ	120
5.3.4	Ajustement du profil de pression	122
5.4	Conclusions	122
6	Mesure des propriétés thermodynamiques d'un amas distant de faible masse	125
6.1	Observations de ACT-CL J0215.4+0030	126
6.1.1	Détection de l'amas par l'Atacama Cosmology Telescope	126
6.1.2	Observations de l'amas avec NIKA2	126
6.1.3	Observations de l'amas avec <i>XMM-Newton</i>	128
6.2	Contamination par des sources ponctuelles	128
6.2.1	Identification des sources ponctuelles	128
6.2.2	Estimation de la contamination	129
6.2.3	Résultats	130
6.3	Mesures des propriétés thermodynamiques du milieu intra-amas	131
6.3.1	Modélisation du milieu intra-amas	133
6.3.2	Analyse MCMC	133
6.3.3	Auto-consistance des estimations de flux des sources ponctuelles	135
6.4	Résultats	135
6.4.1	Thermodynamique du milieu intra-amas	136
6.4.2	Grandeurs intégrées	138
6.5	Conclusions	139
7	La relation d'échelle SZ – Masse	141

7.1	Modélisation et état de l'art	142
7.1.1	Paramètre de Compton intégré et masse	142
7.1.2	État de l'art des mesures	144
7.2	Modélisation probabiliste de la relation d'échelle	146
7.2.1	Relation linéaire masse-observable	146
7.2.2	Effets de sélection et représentativité	149
7.3	Ajustement de la relation d'échelle par MCMC	152
7.3.1	Distribution de probabilité <i>a priori</i> des paramètres	152
7.3.2	Analyse MCMC par échantillonnage de Gibbs	153
7.4	Validation sur données simulées	155
7.4.1	Incertitudes de mesure et dispersion intrinsèque	155
7.4.2	Estimateur de masse dispersé	157
7.4.3	Estimateur de masse biaisé	157
7.4.4	Données tronquées	158
7.5	Application au grand programme SZ de NIKA2	158
7.5.1	Simulation d'échantillons	158
7.5.2	Modélisation et impact de la sélection	160
7.5.3	Incertitudes sur les paramètres d'intérêt	163
7.5.4	Extensions du grand programme SZ	164
7.6	Conclusions et perspectives	166
Conclusions		169
Perspectives		173
Bibliographie		188

Introduction

Les vingt dernières années ont été l'époque de l'essor de la cosmologie observationnelle. Le développement de grandes expériences permettant de réaliser des relevés profonds du ciel avec une grande précision a permis d'accroître notre connaissance des propriétés de l'Univers. En particulier, les paramètres cosmologiques ont pu être contraints avec une grande précision en ajustant les paramètres libres du modèle standard de la cosmologie à partir de ces données. L'utilisation de plusieurs sondes pour mesurer ces paramètres cosmologiques a permis de tirer profit des différentes sensibilités des observables aux propriétés de l'Univers. Si la plupart de ces mesures concordent, certains désaccords subsistent, pouvant être le signe de physique encore inconnue, ou bien d'effets systématiques non-maîtrisés. Avec l'augmentation de la quantité des données disponibles et la diminution des incertitudes statistiques résultante, ces effets deviendront de plus en plus dominants. La cosmologie observationnelle devient donc une science de grande précision, ce qui crée un besoin d'études rigoureuses des systématiques associées aux analyses cosmologiques.

Dans ce contexte, les amas de galaxies sont des objets d'un grand intérêt. Leur abondance en masse et en redshift est étroitement liée à la distribution de masse dans l'Univers. Cette dernière étant intimement liée aux paramètres cosmologiques, le comptage d'amas dans l'Univers par intervalles de masse et redshift peut être utilisé comme une sonde cosmologique. De grands catalogues d'amas ont pu être compilés au cours des dernières années, qui exploitent les observables des amas dans différentes longueurs d'onde pour les détecter dans de grands relevés du ciel. Parmi les observables utilisées, l'effet Sunyaev-Zeldovich (SZ) permet de détecter les amas dans le domaine millimétrique, grâce à leur empreinte sur le fond diffus cosmologique. Les catalogues les plus récents construits grâce à l'effet SZ ont été obtenus à partir des télescopes ACT (Atacama Cosmology Telescope, [2]), SPT (South Pole Telescope, [3, 4]) et du satellite *Planck* [5], et comportent plusieurs milliers d'amas. Dans le futur, le Simons Observatory [6] et CMB-S4 [7] augmenteront ce nombre par un à deux ordres de grandeur, et seront accompagnés par les relevés du Vera Rubin Observatory [8] et d'*Euclid* [9] en optique et infrarouge, et d'eROSITA [10] en X.

Ainsi, à l'instar de la cosmologie observationnelle dans laquelle elle s'inscrit, la cosmologie avec des amas entre dans une ère de précision, avec un grand nombre de catalogues comportant chacun de grands nombres d'amas. Le contrôle des systématiques s'avère être d'une importance cruciale, afin de pouvoir tirer pleinement profit de ces catalogues d'amas comme sondes cosmologiques. Cette maîtrise est actuellement limitée par une faible connaissance des propriétés physiques des amas de galaxies, en particulier des amas distants. En effet, une caractérisation précise de la physique des amas est nécessaire à leur exploitation cosmologique. Dans le cadre des relevés de l'effet SZ mentionnés précédemment, cette connaissance est limitée aux objets proches du fait de la faible résolution angulaire des instruments. Il est donc nécessaire de compléter ces relevés avec des suivis d'amas distants, en utilisant des instruments à haute résolution. Parmi ces études, le grand programme SZ de NIKA2

réalise un suivi d'amas distants détectés par *Planck* et ACT avec la caméra millimétrique à grande résolution angulaire NIKA2, installée au télescope de 30 mètres de l'IRAM [11]. L'objectif de ce suivi est d'exploiter les observations SZ résolues pour mesurer deux outils nécessaires à la cosmologie à partir de relevés SZ : le profil de pression moyen des amas de galaxies, et la relation d'échelle masse-observable. Ces outils sont actuellement étalonnés sur des amas proches, en utilisant des observations SZ à basse résolution ou X.

Le travail développé au cours de cette thèse s'inscrit dans le cadre du grand programme SZ de NIKA2, et de l'étude des effets systématiques en cosmologie avec des amas détectés par effet SZ. Ce manuscrit présente le travail effectué, de l'analyse des données NIKA2 aux prédictions des contraintes pouvant être mises sur la relation d'échelle masse-observable grâce au grand programme SZ, en passant par la mesure des propriétés thermodynamiques des amas de galaxies en tenant compte des systématiques caractéristiques des observations millimétriques. Il se découpe en sept chapitres.

- Le chapitre 1 présente le contexte du modèle standard de la cosmologie, en décrivant le lien entre le contenu de l'Univers et ses propriétés, et la formation des grandes structures. Il permet d'introduire les éléments importants des analyses cosmologiques basées sur des amas de galaxies.
- Le chapitre 2 est consacré aux amas de galaxies et à leur utilisation en cosmologie. La définition des amas de galaxies, leur composition et leurs caractéristiques sont présentées, de même que les différentes possibilités pour les observer. Nous décrivons le lien entre la distribution d'amas en masse et en redshift et les paramètres cosmologiques, et les analyses par comptage d'amas permettant de contraindre ces derniers à partir de catalogues d'amas. Nous présentons quelques-uns des résultats les plus récents de telles études, et les expériences à venir.
- Le chapitre 3 présente la caméra NIKA2 au télescope de 30 mètres de l'IRAM, ainsi que son grand programme SZ. Nous décrivons les caractéristiques techniques du télescope et de NIKA2, notamment les détecteurs à inductance cinétique de cette caméra. Les performances de l'instrument sont également présentées, de même que les observations au télescope, que j'ai pu réaliser en participant à cinq semaines d'observation au télescope au cours des 18 premiers mois de ma thèse, puis trois semaines d'observations à distance. Nous présentons les caractéristiques du grand programme SZ de NIKA2, son échantillon d'amas, et ses objectifs scientifiques.
- Le chapitre 4 décrit la procédure utilisée pour construire des cartes de l'effet SZ à partir des données brutes de NIKA2. Nous présentons tout d'abord la procédure employée pour séparer le signal astrophysique d'intérêt du bruit dans les données. Cette procédure, nommée décorrélation, nécessite un traitement complexe du bruit dans les données, car celui-ci est plusieurs ordres de grandeur supérieur au signal d'intérêt. Nous présentons l'évaluation de la performance de la décorrélation et des systématiques qu'elle induit, qui sont prises en compte au cours de l'analyse des cartes SZ résultantes. Nous décrivons également la procédure mise en place pour estimer la contamination des cartes SZ par des sources ponctuelles.
- Le chapitre 5 présente le logiciel PANCO2, que j'ai développé au cours de ma thèse. L'objectif de ce programme est de contraindre les propriétés thermodynamiques des amas de galaxies à partir d'observations SZ par analyse Monte Carlo à chaînes de Markov. Le développement d'un nouveau logiciel pour de telles analyses était nécessaire pour des raisons de performance : les outils existants ne permettaient pas de réaliser l'analyse d'un échantillon d'amas en un temps raisonnablement court. Son fonctionnement permet de tenir compte des effets systématiques caractéristiques des observations SZ, comme la corrélation du bruit résiduel, le filtrage du signal, et la contamination par des

sources ponctuelles. Nous décrivons le fonctionnement du logiciel et les propriétés physiques qu'il permet de sonder, puis présentons sa validation sur des données simulées. PANC02 est le *pipeline* officiel pour l'analyse des données du grand programme SZ de NIKA2, et une généralisation est en cours pour qu'il puisse être utilisé pour d'autres instruments. Ce logiciel sera alors rendu public pour la communauté scientifique.

- Le chapitre 6 est dédié à l'analyse des propriétés thermodynamiques de l'un des amas les plus complexes du grand programme SZ, ACT-CL J0215.4+0030. Il représente l'un des amas de plus faible masse et de plus haut redshift de l'échantillon, et le premier amas du programme pour lequel les observations NIKA2 sont de qualité standard. De plus, le signal SZ de l'amas est fortement contaminé par des sources ponctuelles, dont le flux masque le signal de l'amas. Nous présentons l'extraction des propriétés thermodynamiques de l'amas en considérant les sources ponctuelles comme des paramètres de nuisance de l'analyse, dont le flux est connu *a priori* grâce à des observations du satellite *Herschel*. Nous combinons les observations SZ de NIKA2 avec des données X obtenues par le satellite *XMM-Newton*, ce qui permet une caractérisation complète des propriétés de l'amas.
- Enfin, le chapitre 7 présente la relation d'échelle liant la masse des amas à leur observable dans les relevés SZ. Nous présentons une approche probabiliste basée sur un modèle bayésien hiérarchique. Cette approche permet de tenir compte d'un grand nombre d'effets systématiques caractéristiques des relations masse-observable, comme la dispersion ou le biais des estimateurs de masse, la dispersion intrinsèque autour de la relation, et les effets de sélection. Nous appliquons cette modélisation à des simulations réalistes d'échantillons d'amas similaires au grand programme SZ de NIKA2, ce qui permet d'identifier les biais et systématiques et de prévoir les incertitudes attendues sur les paramètres de la relation à l'issue du grand programme SZ. Nous utilisons également cet outil pour évaluer l'intérêt d'extensions du grand programme SZ.

CHAPITRE 1

Évolution de l’Univers et des grandes structures

1.1	Le modèle standard de la cosmologie	12
1.1.1	Les équations de Friedmann	12
1.1.2	Expansion de l’Univers, accélération et énergie sombre	13
1.1.3	Histoire thermique de l’Univers	15
1.2	Origine des grandes structures	16
1.2.1	Limitations du modèle standard de la cosmologie	16
1.2.2	Fluctuations primordiales et inflation	17
1.3	Évolution des grandes structures	19
1.3.1	Croissance des inhomogénéités de densité	19
1.3.2	Spectre de puissance de la distribution de matière	21
1.3.3	Fonction de masse et distribution de halos	22
1.4	Conclusions	23

Le modèle standard de la cosmologie offre un formalisme décrivant l’évolution de l’Univers, du *Big Bang* à aujourd’hui. Ses succès observationnels sont nombreux, de l’expansion de l’Univers à l’existence du fond diffus cosmologique, en passant par la nucléosynthèse primordiale ; aujourd’hui, il n’existe pas d’observations le mettant en échec. Dans ce cadre, les propriétés de l’Univers sont décrites par un nombre restreint de quantités, nommées paramètres cosmologiques, dont la mesure est aujourd’hui l’enjeu de la cosmologie observationnelle. Cette thèse s’inscrit dans ce cadre, et propose en particulier de contribuer à l’étude des mesures des paramètres cosmologiques en utilisant les amas de galaxies comme sondes cosmologiques.

Ce chapitre décrit les éléments nécessaires à la compréhension des analyses cosmologiques basées sur des amas de galaxies. Nous commençons par présenter le modèle Λ CDM, ou modèle standard de la cosmologie, décrivant le lien entre le contenu de l’Univers et sa dynamique. Nous décrivons ensuite la formation des grandes structures, de leur origine à l’issue de l’inflation à leur développement récent par croissance des instabilités.

1.1 Le modèle standard de la cosmologie

1.1.1 Les équations de Friedmann

La cosmologie moderne décrit l’Univers comme un objet dynamique obéissant aux lois de la relativité générale. Ce cadre permet de lier la structure de l’espace-temps avec le contenu de l’Univers au travers des équations d’Einstein [12] :

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \quad (1.1)$$

La partie gauche de l’équation encode la géométrie de l’espace-temps au travers du tenseur d’Einstein $G_{\mu\nu}$ et de la métrique $g_{\mu\nu}$, et la partie droite décrit la dynamique des composants de l’Univers par leur tenseur énergie-impulsion $T_{\mu\nu}$. G est la constante gravitationnelle Newtonienne, c la vitesse de la lumière dans le vide, et Λ la constante cosmologique, qui sera discutée en section 1.1.2. Dans l’hypothèse d’un Univers globalement homogène et isotrope, la solution dynamique la plus simple de l’équation (1.1) est donnée par la métrique de Friedmann–Lemaître–Robertson–Walker (FLRW) :

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = c^2 dt^2 - a^2(t) \left[\frac{1}{1 - kr^2} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2 \right], \quad (1.2)$$

où (t, r, θ, φ) sont les coordonnées sphériques à quatre dimensions.

Le paramètre de courbure k caractérise la géométrie globale de l’Univers, qui peut être hyperbolique ($k = -1$), sphérique ($k = 1$), ou plate ($k = 0$, cas favorisé par les observations). La fonction $a(t)$ est le facteur d’échelle de l’Univers, décrivant son expansion. Son expression peut être écrite en considérant le contenu de l’Univers comme un fluide parfait comobile de densité d’énergie ρ et de pression p , dont le tenseur énergie-impulsion s’écrit en fonction de la quadri-vitesse du fluide u^μ

$$T^{\mu\nu} = (p/c^2 + \rho) u^\mu u^\nu - p g^{\mu\nu}. \quad (1.3)$$

Alors, la combinaison des équations (1.1), (1.2) et (1.3) permet d’aboutir aux équations de Friedmann :

$$H^2(t) \equiv \left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2} \quad (1.4)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \left(\rho + \frac{3}{c^2} p \right) + \frac{\Lambda c^2}{3} \quad (1.5)$$

où \dot{a} dénote la dérivée temporelle de a , et $H(t)$ est le paramètre de Hubble, quantifiant le taux d’expansion de l’Univers.

Ces équations différentielles lient la dynamique de l’Univers à l’évolution de la densité d’énergie et de pression de son contenu. Cette évolution dépend du paramètre d’équation d’état $w = p/(\rho c^2)$ du fluide : la combinaison des équations (1.4) et (1.5) donne l’équation de conservation de l’énergie pour un fluide parfait comobile,

$$\dot{\rho} = -3(\rho + p/c^2) \times H(t), \quad (1.6)$$

et donc l’évolution avec le facteur d’échelle de la densité d’énergie ρ :

$$\rho(a) \propto a^{-3(1+w)} \quad (1.7)$$

Composante i	w_i	$\Omega_{i,0}$
Matière m	0	0.3111 ± 0.0056
Rayonnement r	$1/3$	$\sim 10^{-5}$
Constante cosmologique Λ	-1	0.6889 ± 0.0056
Courbure k	$-1/3$	0.0007 ± 0.0019

Table 1.1 : Paramètres d'état w et abondances actuelles Ω_0 des constituants de l'Univers issues de l'analyse des anisotropies du fond diffus cosmologique mesurées par *Planck* [13].

Le contenu de l'Univers peut être séparé en plusieurs fluides parfaits comobiles, notés j , représentant les différents constituants de l'Univers, et possédant chacun un paramètre d'état w_j . Ces constituants sont la matière, comprenant l'ensemble des baryons de l'Univers, ainsi que la matière sombre, et le rayonnement, composé des photons et autres espèces relativistes de l'Univers. L'équation de Friedmann (1.4) peut s'écrire en fonction des densités ρ_j des différents constituants j de l'Univers :

$$H^2(a) = \frac{8\pi G}{3} \sum_j \rho_j(a) + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2}. \quad (1.8)$$

Ainsi, en introduisant la densité critique ρ_c comme :

$$\rho_c(a) \equiv \frac{3H^2(a)}{8\pi G}, \quad (1.9)$$

on peut définir les abondances relatives de chaque composante comme $\Omega_j \equiv \rho_j/\rho_c$. La constante cosmologique et la courbure peuvent donc être considérées comme des constituants supplémentaires de l'Univers, d'abondances relatives respectives $\Omega_\Lambda \equiv \Lambda c^2/3H^2$ et $\Omega_k \equiv -kc^2/(aH)^2$. La première équation de Friedmann (1.4) s'écrit finalement :

$$H^2(a) = H_0^2 \sum_i \Omega_{i,0} a^{-3(1+w_i)}, \quad (1.10)$$

où H_0 est le taux d'expansion actuel de l'Univers, et où les composantes i comprennent les constituants précédemment notés j (matière et rayonnement), ainsi que la courbure et la constante cosmologique. Leurs abondances relatives actuelles sont notées $\Omega_{i,0}$. Les différents constituants de l'Univers, leurs paramètres d'équation d'état, et leur abondances actuelles mesurées par l'analyse des anisotropies du fond diffus cosmologique (ou CMB pour *Cosmic Microwave Background*) par le satellite *Planck* [13] sont résumés dans la table 1.1.

1.1.2 Expansion de l'Univers, accélération et énergie sombre

Les équations de Friedmann permettent de lier l'évolution du facteur d'échelle de l'Univers aux densités d'énergie de ses différents constituants, et d'étudier cette évolution dans le temps. D'après la combinaison de la première équation de Friedmann (1.10) et de l'équation (1.7), il apparaît que le facteur d'échelle d'un Univers de densité d'énergie dominée par un constituant i d'équation d'état $w_i \geq -1$ ne peut qu'augmenter avec le temps¹. Chacun des constituants de notre Univers vérifiant cette condition (comme montré en table 1.1), nous pouvons en déduire que celui-ci est en expansion. Cette expansion se caractérise par une dilatation de l'espace-temps avec le temps, et une augmentation des

¹Dans la mesure où $\Omega_{i,0}$ est positif.

distances séparant les différents objets présents dans l’Univers. Il en découle que par le passé, l’Univers entier était contenu dans un volume bien plus faible qu’aujourd’hui, diminuant en remontant le temps jusqu’à atteindre une singularité de l’espace-temps : c’est le modèle du *Big Bang*, élément majeur de notre compréhension de la cosmologie.

Historiquement, la constante cosmologique Λ fût introduite par Einstein dans les équations du champ (1.1) afin de pouvoir décrire un modèle d’Univers statique. Cette hypothèse fût mise en échec par les confirmations observationnelles de l’expansion par Lemaître [14] et Hubble [15]. Ces études ont mis en évidence le décalage vers le rouge des spectres de galaxies l’expansion : un photon émis à un instant t_1 et à une longueur d’onde λ_1 , verra sa longueur d’onde augmenter au cours de son voyage dans l’Univers de par la dilatation de l’espace-temps. Ainsi, à un instant postérieur t_0 , sa longueur d’onde λ_0 sera décalée vers les grandes longueurs d’onde – et donc vers le rouge, décalage quantifié par le *redshift* z :

$$1 + z = \frac{\lambda_0}{\lambda_1} = \frac{a(t_0)}{a(t_1)}. \quad (1.11)$$

Einstein retira alors son hypothèse d’Univers statique et la constante cosmologique, menant aux travaux sur le modèle cosmologique Einstein-de Sitter [16] décrivant un univers en expansion continue avec $\Lambda = 0$.

Les modèles d’univers à constante cosmologique non-nulle ne refirent leur apparition que soixante ans plus tard, lorsque les observations de supernovas de type Ia par Riess *et al.* [17] et Perlmutter *et al.* [18] mirent en évidence l’accélération récente de l’expansion de l’Univers. En vertu de la deuxième équation de Friedmann (1.5), une telle accélération, caractérisée par $\ddot{a} > 0$, ne peut avoir lieu sans la présence d’un constituant de l’Univers de paramètre d’équation d’état $w < -1/3$. Une nouvelle espèce fût donc ajoutée au contenu de l’Univers, génériquement nommée énergie sombre (*dark energy*), d’abondance relative Ω_{DE} et d’équation d’état w_{DE} . Dans le cas où $w_{\text{DE}} = -1$, l’énergie sombre se comporte comme la constante cosmologique d’Einstein. Les analyses cosmologiques actuelles, basées sur l’exploitation de différentes sondes pour mesurer les différents paramètres cosmologiques, sont toutes en accord avec la valeur de $w_{\text{DE}} = -1$, faisant de la constante cosmologique un candidat à l’explication de l’origine de l’accélération de l’expansion. Des modèles plus élaborés permettent au paramètre d’état de varier dans le temps, par exemple dans la paramétrisation CPL [19, 20],

$$w(a) = w_0 + (1 - a)w_a, \quad (1.12)$$

mais les incertitudes sur les mesures ne permettent actuellement pas de différencier de tels modèles d’une constante cosmologique. De même, des modèles de gravitation modifiée cherchent à expliquer l’accélération en modifiant la relativité générale aux échelles cosmologiques, mais ne sont pour l’instant pas contraignables observationnellement (voir [21, 22] pour des revues).

Le modèle actuellement adopté comme modèle standard de la cosmologie est le modèle ΛCDM pour *Lambda Cold Dark Matter*, dans lequel l’énergie sombre est une constante cosmologique ($w_{\text{DE}} = -1$), la matière sombre est non-relativiste, et l’Univers est plat ($\Omega_k = 0$). Ses extensions les plus couramment considérées en cosmologie observationnelles sont :

- le modèle $w\text{CDM}$, dans lequel la densité d’énergie sombre peut varier dans le temps ($w_{\text{DE}} \neq -1$) ;
- le modèle $w_0w_a\text{CDM}$, extension du modèle $w\text{CDM}$ dans laquelle le paramètre d’état de l’énergie sombre peut également varier dans le temps (voir équation 1.12) ;
- le modèle $k\Lambda\text{CDM}$, dans lequel l’Univers n’est pas nécessairement plat ;

- le modèle $\nu\Lambda\text{CDM}$, dans lequel on considère la masse des neutrinos comme un paramètre cosmologique, du fait de l'importance de ceux-ci dans la formation des structures (voir par exemple [23] pour une revue).

1.1.3 Histoire thermique de l'Univers

Le modèle standard de la cosmologie fournit une description du contenu de l'Univers et de son évolution dans le temps. Quelques instants après le Big Bang, l'intégralité des constituants de l'Univers est contenue dans un volume très faible. Il est alors rempli d'un gaz chaud ionisé de pression et température extrêmes, nommé plasma primordial, dans lequel les particules élémentaires sont toutes libres et en interaction constante. La densité d'énergie est alors dominée par le rayonnement, et le facteur d'échelle évolue comme $a \propto t^{1/2}$.

Avec l'expansion, le volume de l'Univers augmente, permettant au plasma primordial de se diluer. Sa densité, sa température et sa pression décroissent, diminuant ainsi les taux d'interaction des particules. Les espèces interagissant le plus faiblement, comme les neutrinos et la matière sombre, commencent alors à se découpler du plasma primordial alors que leur taux d'interaction devient plus faible que le taux d'expansion de l'Univers. Lorsque la température atteint ~ 100 keV, les premiers noyaux se forment par interaction entre protons et neutrons : des noyaux de deutérium, hélium, puis lithium sont alors formés, marquant la nucléosynthèse primordiale (pour une revue du processus et de résultats récents, voir [24]).

Alors que l'expansion continue son cours, la densité d'énergie du rayonnement décroît en $\rho_r \propto a^{-4}$, plus vite que celle de la matière qui diminue en $\rho_m \propto a^{-3}$ (cf. section 1.1.1). L'Univers arrive donc dans une phase où la densité d'énergie de la matière domine, à partir d'un redshift de $z \sim 3400$ ($T \sim 10^4$ K), soit quelques soixante mille ans après le Big Bang, et l'expansion continue avec $a \propto t^{2/3}$. Par la suite, alors que le taux d'interactions continue de chuter, les premiers atomes d'hydrogène se forment par combinaison d'électrons avec les noyaux formés lors de la nucléosynthèse primordiale, constituant la recombinaison. L'Univers devient alors neutre, et les photons peuvent se propager librement et se découplent : c'est l'émission du fond diffus cosmologique (ou CMB), à $z \sim 1000$. Ce rayonnement constitue la première lumière émise dans l'Univers jusqu'alors opaque, et il constitue l'une des sondes majeures de la cosmologie. Il se manifeste observationnellement comme un rayonnement dont le spectre est celui d'un corps noir de température $T_{\text{CMB}} = 2.725$ K [25], remarquablement homogène sur tout le ciel, présentant des fluctuations en température de l'ordre de 10^{-4} K. Le spectre de puissance de ces fluctuations est intimement lié aux propriétés de l'Univers au moment du découplage des photons, et donc aux paramètres cosmologiques.

La matière s'arrange alors par effondrement gravitationnel autour des régions les plus denses, donnant naissance à la structure à grande échelle de l'Univers. L'origine de ces fluctuations, de même que le processus de formation des structures, sera détaillée dans la section 1.2. Enfin, à un redshift $z \sim 0.5$, la densité d'énergie de la matière devient plus faible que celle de l'énergie sombre, qui commence à dominer le contenu énergétique de l'Univers. C'est alors que l'accélération de l'expansion commence.

Les grandes étapes de l'histoire thermique de l'Univers sont résumées sur la figure 1.1. Celle-ci montre une représentation schématique de l'Univers au fur et à mesure de ces grandes étapes. L'évolution du facteur d'échelle est représentée par l'agrandissement avec le temps (de gauche à droite) de la représentation de l'Univers. Entre l'inflation et la recombinaison, l'Univers est opaque aux photons et très homogène. Par la suite, la matière le composant se structure pour former la toile cosmique, composée de filaments séparés par des vides, et dont les intersections forment des pics de surdensité

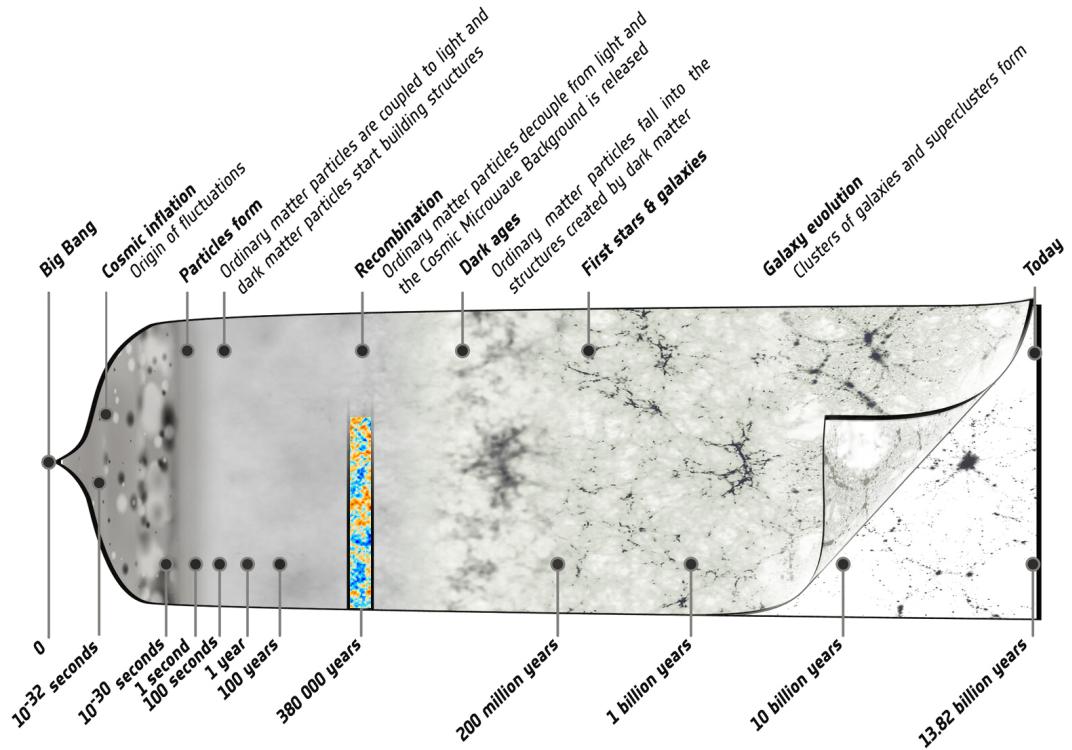


Figure 1.1 : Résumé des étapes de l'histoire thermique de l'Univers, du Big Bang à aujourd'hui. Crédit : [ESA – C. Carreau](#)

qui donneront naissance aux amas de galaxies : c'est la formation des grandes structures, qui sera détaillée dans la section suivante. Les régions denses continuent par la suite à accréter de la matière, et deviennent de plus en plus denses, donnant lieu à la structure de l'Univers contemporain.

1.2 Origine des grandes structures

Les équations de Friedmann permettent l'étude de la dynamique d'un univers homogène et isotrope. La faible amplitude des fluctuations en température du CMB nous confirme l'isotropie à grande échelle de l'Univers. Cependant, l'Univers contemporain est hautement hétérogène aux petites échelles, comme en témoigne l'existence de structures telles que les galaxies, amas de galaxies, étoiles et autres astres. Notre Univers est donc homogène aux grandes échelles et hétérogène aux petites, ce qui n'est pas expliqué par le modèle standard de la cosmologie. Nous décrivons dans cette section le paradigme inflationnaire, permettant de pallier aux lacunes du modèle standard, et offrant un cadre théorique pouvant expliquer l'origine de la formation des grandes structures de l'Univers.

1.2.1 Limitations du modèle standard de la cosmologie

Bien que capable d'expliquer une grande variété de phénomènes observés dans l'Univers, le modèle Λ CDM est limité et ne fournit pas de réponse à certaines questions fondamentales de la cosmologie, parmi lesquelles :

Le problème de l'homogénéité à grande échelle

Les observations des anisotropies en température du fond diffus cosmologique nous indiquent que l'Univers était extraordinairement homogène au moment de son découplage 380 000 ans après le Big Bang, avec des fluctuations quatre ordres de grandeurs plus faibles que sa température moyenne. Cette homogénéité peut s'expliquer par un équilibre thermique dans le plasma primordial. Cependant, un tel équilibre requiert un contact causal entre les différentes régions de l'Univers, contact n'ayant pu exister au moment de l'émission du CMB au vu de la valeur du facteur d'échelle à cette époque. Quantitativement, alors que le CMB est homogène sur tout le ciel (et donc pour des points séparés par des angles allant jusqu'à 180°), l'angle tendu par le rayon d'horizon de Hubble à cette époque est de l'ordre du degré. Le modèle standard de la cosmologie ne permet donc pas d'expliquer l'homogénéité en température du fond diffus cosmologique.

Le problème de l'hétérogénéité à petite échelle

Bien que remarquablement homogène à grande échelle, l'Univers est très hétérogène à petite échelle. Le modèle cosmologique standard ne permet d'expliquer que partiellement ce constat : de faibles inhomogénéités dans l'Univers primordial peuvent croître avec l'expansion jusqu'à devenir les structures que nous connaissons. Cependant, l'origine des fluctuations de densité primordiales nécessaires à ce développement n'est pas expliquée par le modèle cosmologique standard.

Le problème de la platitude

Les analyses cosmologiques récentes, basées sur l'exploitation de différentes sondes, s'accordent à dire que la géométrie de l'Univers est extrêmement plate, correspondant à une densité d'énergie dans l'Univers très proche de la densité critique (équation 1.9). Une courbure aussi faible que celle mesurée par l'analyse des anisotropies du CMB avec *Planck* ($\Omega_{k,0} \sim 10^{-3}$) nécessite une courbure inférieure à 10^{-60} à l'issue de l'ère de *Planck*². Des contraintes aussi fortes requièrent donc un ajustement fin des propriétés physiques de l'Univers primordial pour qu'il puisse se développer en l'Univers contemporain que nous observons, sans que le modèle cosmologique standard n'offre d'explication.

1.2.2 Fluctuations primordiales et inflation

Les problèmes d'homogénéité à grande échelle et de platitude peuvent être résolus par un même mécanisme : l'inflation. Dans ce scénario, l'Univers primordial a subi une phase d'expansion accélérée très rapide, caractérisée par $\ddot{a} > 0$, soit :

$$\frac{d}{dt} \frac{1}{aH} < 0 \quad (1.13)$$

La présence de faibles hétérogénéités dans l'Univers primordial, à l'origine du processus de formation des structures, peut également être expliquée dans ce cadre.

Le paradigme de l'inflation comprend un grand nombre de modèles visant à expliquer l'origine de l'expansion accélérée dans l'Univers primordial. Dans sa forme la plus simple, l'inflation est due à un champ scalaire homogène ϕ nommé *inflaton*, faiblement couplé et de potentiel $V(\phi)$, dont la densité d'énergie domine celles de la matière et du rayonnement dans l'Univers primordial, avant de

²Premières 10^{-43} secondes de l'Univers.

s'effondrer. Le Lagrangien de ce champ s'écrit

$$\mathcal{L} = \frac{1}{2} \partial^\mu \phi \partial_\mu \phi - V(\phi) = \frac{\dot{\phi}^2}{2} - V(\phi), \quad (1.14)$$

et le tenseur énergie-impulsion de ce champ $T^{\mu\nu} = \partial^\mu \phi \partial^\nu \phi - \mathcal{L}g^{\mu\nu}$ prend alors la forme de celui d'un fluide parfait dont la densité et la pression s'écrivent

$$\begin{aligned} \rho_\phi c^2 &= \frac{\dot{\phi}^2}{2} + V(\phi) + \frac{1}{2} \left(\frac{\nabla \phi}{a} \right)^2, \\ p_\phi &= \frac{\dot{\phi}^2}{2} - V(\phi) - \frac{1}{6} \left(\frac{\nabla \phi}{a} \right)^2. \end{aligned} \quad (1.15)$$

Dans une phase d'expansion rapide de l'Univers (comme au cours du *slow-roll* discuté dans le paragraphe suivant), les termes proportionnels au gradient du champ d'inflaton $\nabla\phi$ deviennent rapidement négligeables, du fait de la présence au dénominateur du facteur d'échelle. La conservation du tenseur énergie-impulsion donne alors l'équation du mouvement de ce champ :

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV(\phi)}{d\phi} = 0. \quad (1.16)$$

On peut alors distinguer deux régimes. Dans le cas où $\ddot{\phi}$ est négligeable devant les autres termes, l'énergie potentielle du champ d'inflaton domine son énergie cinétique : on dit que le champ se trouve dans une phase de *slow-roll*³. On peut alors écrire $\dot{\phi}^2 \ll V(\phi)$, et le rapport de la densité d'énergie et de la pression du champ d'inflaton (équation 1.15) donne un paramètre d'équation d'état $w_\phi = -1$, caractéristique d'une expansion accélérée de l'Univers. À l'inverse, si le terme cinétique domine par rapport au potentiel, l'équation du mouvement est celle d'un oscillateur harmonique amorti, et le champ se trouve dans une phase de *reheating*. Dans cette phase, l'oscillation du champ d'inflaton permet de convertir son énergie par production de particules élémentaires (voir par exemple [26, 27]).

Le comportement de l'inflaton est illustré en figure 1.2. S'il se trouve dans une région de potentiel élevé, il se dirigera vers un potentiel plus faible en *slow-roll*, générant l'inflation dans la région grise. Une fois proche d'un minimum de potentiel, le champ se trouvera dans une phase d'oscillations périodiques amorties avant de se stabiliser, mettant un terme à l'inflation.

La dynamique de l'inflation permet de comprendre la raison pour laquelle elle résout les problèmes de la platitude, de l'horizon et de l'hétérogénéité. La densité d'énergie correspondant à la courbure de l'Univers étant définie par $\Omega_k = -kc^2/(aH)^2$, celle-ci diminue rapidement avec le facteur d'échelle. Quantitativement, en introduisant le nombre de *e-folds* N_e , quantifiant l'augmentation du facteur d'échelle au cours de la phase inflationnaire, comme

$$N_e = \ln \frac{a_f}{a_i} = \int_{t_i}^{t_f} H(t) dt, \quad (1.17)$$

il apparaît que $N_e \sim 60$ permet de passer d'une courbure $\Omega_k \sim 1$ au début de l'inflation à $\Omega_k \sim 10^{-60}$ à son issue, condition requise pour observer une courbure actuelle aussi faible que celle mesurée par

³L'équation du mouvement (1.16) est similaire à celle d'une balle dévalant une colline avec une friction non-nulle. Le *slow-roll* correspond au roulement initial, où la friction domine et ralentit le mouvement – voir figure 1.2, panneau gauche.

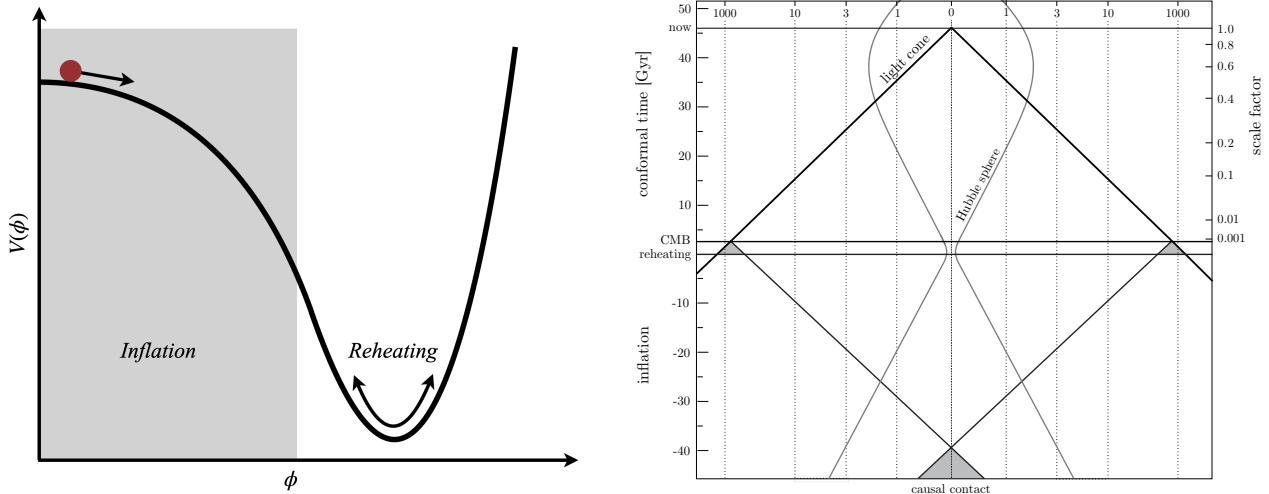


Figure 1.2 : **Gauche** : Exemple de comportement du champ d'inflaton illustrant les deux régimes de l'équation (1.16) : le champ commence en *slow roll* dans une région de potentiel élevé, puis se stabilise dans une région de bas potentiel, avec des oscillations rapides correspondant au *reheating*. **Droite** : Diagramme d'espace-temps illustrant le contact causal de régions opposées dans le CMB au début de l'inflation. Figure extraite de [28].

le satellite *Planck*. De plus, le rayon de Hubble diminue pendant la phase inflationnaire, permettant à des points diamétralement opposés dans le CMB d'avoir été en contact causal. Cela permet donc d'expliquer l'homogénéité du CMB, comme illustré sur le panneau droit de la figure 1.2.

Enfin, le champ d'inflaton permet de comprendre l'origine des fluctuations de densité dans l'Univers primordial : considérons des fluctuations quantiques du champ comme une perturbation au premier ordre, soit

$$\phi(\vec{x}, t) = \bar{\phi}(t) + \delta\phi(\vec{x}, t), \quad (1.18)$$

où $\bar{\phi}(t)$ est la valeur moyenne du champ dans l'espace, ne dépendant que du temps dans l'hypothèse d'homogénéité.

Il apparaît alors que la valeur du champ et de son potentiel ne sont pas égales en tout point de l'espace. En vertu du lien entre densité d'énergie et espace-temps décrit par les équations d'Einstein (1.1), ces inhomogénéités créent des perturbations dans la métrique de l'Univers, qui génèrent à leur tour des fluctuations dans la densité d'énergie du plasma primordial. Ainsi, en plus des problèmes de l'horizon et de la latitude, l'inflation offre une origine au processus de formation des structures, qui sera détaillée dans la section suivante.

1.3 Évolution des grandes structures

1.3.1 Croissance des inhomogénéités de densité

Nous avons vu dans la section précédente comment l'inflation permettait de générer des fluctuations de densité dans l'Univers primordial. Celles-ci seront ensuite amplifiées au cours de l'évolution de l'Univers par effondrement gravitationnel de la matière vers les régions les plus denses. Ce processus de formation des structures est à l'origine des galaxies et de la distribution à grande échelle de la matière dans l'Univers contemporain, formant la toile cosmique. Cette section détaille ce formalisme dans le but de comprendre l'origine des amas de galaxies, qui constituent la sonde cosmologique étudiée dans cette thèse.

Dans le régime linéaire, les fluctuations de densité issues de l'inflation peuvent s'écrire comme une perturbation au premier ordre de la densité moyenne,

$$\rho(\vec{x}, t) = \bar{\rho}(t) [1 + \delta(\vec{x}, t)], \quad (1.19)$$

où $\bar{\rho}(t)$ est la densité moyenne de l'Univers, \vec{x} les coordonnées comobiles⁴ tri-dimensionnelles, et $\delta(\vec{x}, t)$ est appelé le paramètre de contraste.

La densité moyenne de l'Univers ne dépend que du temps, traduisant l'hypothèse d'un univers homogène à grande échelle et hétérogène seulement aux petites échelles. Le fluide parfait comobile composant l'Univers doit alors satisfaire aux équations d'Euler, de la continuité, et de Poisson :

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \vec{\nabla}) \vec{v} = -\vec{\nabla} \phi - \frac{\vec{\nabla} p}{\rho} \quad (1.20)$$

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) = 0 \quad (1.21)$$

$$\nabla^2 \Phi = 4\pi G \rho \quad (1.22)$$

où p est la pression des fluides, \vec{v} leur vitesse, et Φ le potentiel gravitationnel.

En injectant l'expression de la densité perturbée (1.19) dans ces équations, et en les combinant, on trouve l'expression de l'évolution des perturbations dans le temps :

$$\frac{\partial^2 \delta(\vec{x}, t)}{\partial t^2} + 2H(t) \frac{\partial \delta(\vec{x}, t)}{\partial t} = \left[4\pi G \bar{\rho}(t) + \frac{c_s^2}{a^2} \nabla^2 \right] \delta(\vec{x}, t), \quad (1.23)$$

où c_s est la vitesse du son dans le fluide.

Dans l'espace de Fourier, et en définissant $\delta_k \sin(\vec{k} \cdot \vec{r}) = \delta(\vec{x}, t)$, cette équation prend la forme

$$\ddot{\delta}_k + 2H\dot{\delta}_k = \left[4\pi G \bar{\rho} - \frac{c_s^2}{a^2} k^2 \right] \delta_k \quad (1.24)$$

ou encore, en introduisant l'échelle de Jeans $\lambda_J \equiv c_s / \sqrt{\pi/G\rho}$ et le mode associé $k_J = 2\pi a / \lambda_J$,

$$\ddot{\delta}_k + 2H\dot{\delta}_k = \frac{c_s^2}{a^2} (k_J^2 - k^2) \delta_k. \quad (1.25)$$

Deux régimes asymptotiques apparaissent. Pour les perturbations plus grandes que l'échelle de Jeans, soit $k^2 \ll k_J^2$, la gravitation domine l'évolution des perturbations, et leur croissance est permise. À l'inverse, lorsque $k^2 \gg k_J^2$, la pression domine l'effondrement gravitationnel et la croissance des perturbations est atténuée. Ainsi, lorsque la densité d'énergie est partagée entre matière et rayonnement (dans l'Univers primordial, à $z \lesssim 3000$), les inhomogénéités de densité oscillent, augmentant sous domination de la force gravitationnelle, puis diminuant lorsque cette dernière devient faible devant les forces de pression. Ce mécanisme empêche le développement des inhomogénéités au cours de l'ère du rayonnement.

Lorsque la matière domine la densité d'énergie de l'Univers, soit $a(t) \propto t^{2/3}$ d'après l'équation de Friedmann (1.4), les forces de pression deviennent négligeables, autorisant la croissance des inhomogénéités et la formation des structures. La résolution de l'équation (1.25) montre alors que le

⁴Le système de coordonnées comobiles se dilate avec le temps suivant l'expansion de l'Univers.

contraste augmente comme $\delta(t) \propto t^{2/3}$. Plus tard, lors de la domination de l'énergie sombre, le facteur d'échelle croît comme $a(t) \propto \exp(Ht)$, et la résolution de (1.25) montre une diminution du paramètre de contraste comme $\delta(t) \propto \exp(-2Ht)$.

1.3.2 Spectre de puissance de la distribution de matière

Comme nous l'avons vu dans la section précédente, la distribution de densité de l'Univers primordial résulte d'un processus aléatoire, et la structure de l'Univers contemporain du développement de ces inhomogénéités. Dans le cas de fluctuations gaussiennes, la distribution de densité primordiale peut se résumer à la fonction de corrélation du paramètre de contraste,

$$\xi(\vec{r}) = \langle \delta(\vec{x} + \vec{r}) \delta(\vec{r}) \rangle, \quad (1.26)$$

où $\langle \dots \rangle$ représente la valeur moyenne prise en toute position \vec{x} de l'Univers.

On peut encore définir le spectre de puissance primordial $P_{\text{prim.}}(k)$ comme la transformée de Fourier de cette fonction de corrélation,

$$P_{\text{prim.}}(k) = \int \xi(\vec{r}) e^{-i\vec{k} \cdot \vec{r}} d^3 r. \quad (1.27)$$

Les contraintes observationnelles sur le spectre de puissance primordial, obtenues notamment à partir de l'analyse des anisotropies du CMB, favorisent une forme en loi de puissance de ce dernier, soit $P(k) \propto k^n$, avec $n \simeq 1$ [13]. Le cas où $n = 1$ donne le spectre dit de Harrison-Zeldovich [29, 30], qualifié d'invariant d'échelle. En effet, en vertu de l'équation de Poisson, un spectre de puissance de la densité de matière de la forme $P(k) \propto k$ mène à un spectre de puissance du potentiel gravitationnel plat.

Dans l'Univers contemporain, les inhomogénéités de densité proviennent de l'accroissement des fluctuations primordiales. Le spectre de puissance de la distribution de matière observable à un redshift z est donc étroitement lié au spectre de puissance primordial, à travers

$$P(z, k) = P_{\text{prim.}}(k) \times T^2(z, k), \quad (1.28)$$

où $T(z, k)$ est la fonction de transfert caractérisant l'évolution des perturbations.

La variance de la distribution de matière dans une sphère de rayon r à un redshift z est alors donnée par :

$$\sigma^2(z, r) = \frac{1}{(2\pi)^3} \int W(kr) P(z, k) d^3 k, \quad (1.29)$$

où $W(kr)$ est la fonction fenêtre utilisée pour lisser la distribution de matière ; par exemple, la fonction *top-hat* considérant une sphère de rayon r ,

$$W(kr) = 3 \frac{\sin kr}{(kr)^3} - 3 \frac{\cos kr}{(kr)^2}. \quad (1.30)$$

L'équation (1.29) permet de définir le paramètre cosmologique σ_8 comme l'amplitude moyenne des fluctuations dans une sphère de $8 h^{-1} \text{ Mpc}$, soit

$$\sigma_8 = \sqrt{\sigma^2(z = 0, r = 8 h^{-1} \text{ Mpc})}. \quad (1.31)$$

Ce paramètre a été mesuré par l’analyses des anisotropies primaires du CMB vues par *Planck* à une valeur de $\sigma_8 = 0.8102 \pm 0.006$ [13]. Nous discuterons des différentes mesures obtenues par d’autres sondes en 2.4.1 (page 44).

1.3.3 Fonction de masse et distribution de halos

Le formalisme développé jusqu’ici repose sur l’hypothèse de linéarité des inhomogénéités du champ de matière primordial : d’après l’équation (1.19), les surdensités représentent des perturbations au premier ordre d’un Univers globalement homogène. Cette hypothèse suppose donc que $\delta \ll 1$, condition non-remplie dans les surdensités de l’Univers contemporain. Par exemple, les amas de galaxies peuvent aujourd’hui atteindre des densités centrales de l’ordre de $\sim 10^{-22} \text{ kg} \cdot \text{m}^{-3}$, correspondant à un contraste de l’ordre de 10^4 . La description de l’Univers contemporain par le seul formalisme de la croissance des structures par développement des inhomogénéités en régime linéaire est donc incomplète.

Les phénomènes physiques devant être pris en compte lors du traitement non-linéaire de la croissance des structures sont très riches, et font de ce formalisme un cadre complexe. Une approche possible est la théorie de Press et Schechter [31], dans laquelle l’hypothèse principale est le fait que l’intégralité de la matière est aujourd’hui contenue dans des halos sphériques à l’équilibre hydrostatique. Dans le cadre de cette théorie, le nombre de halos de masse comprise dans un intervalle $[M, M + dM]$ est donné par la fonction de masse dN/dM . Dans sa formulation courante, formalisée par Tinker *et al.* [32], celle-ci s’écrit à un redshift donné :

$$\frac{dN}{dM} = f(\sigma) \frac{\bar{\rho}}{M} \left| \frac{d \ln \sigma^{-1}}{dM} \right|, \quad (1.32)$$

où $\bar{\rho}$ est la densité moyenne de l’Univers, σ la racine de la variance de la distribution de matière, reliée au spectre de puissance par l’équation (1.29), et $f(\sigma)$ est la fonction de multiplicité dépendant de σ , par

$$f(\sigma) = A \left[1 + \left(\frac{b}{\sigma} \right)^a \right] e^{-c/\sigma^2} \quad (1.33)$$

Les valeurs des paramètres (A, a, b, c) de la fonction de multiplicité sont donc des paramètres de la fonction de masse, et sont généralement mesurés dans des simulations numériques (voir par exemple [32-34]). Comme nous le verrons dans le chapitre suivant, la fonction de masse est un élément central des analyses cosmologiques basées sur les amas de galaxies. En effet, elle dépend de la distribution de matière dans l’Univers, et donc des paramètres cosmologiques.

C’est cette dépendance, illustrée en figure 1.3, qui est exploitée pour contraindre la cosmologie à l’aide d’amas de galaxies. Celle-ci montre l’évolution des valeurs de la fonction de masse de Tinker *et al.* [32] avec les paramètres Ω_m et σ_8 . On y observe que la fonction de masse croît avec Ω_m : plus la matière est abondante dans l’Univers (c’est-à-dire plus Ω_m est grand), plus les halos, et donc les amas de galaxies, sont nombreux⁵. On y remarque également un comportement similaire pour σ_8 : plus les fluctuations de la distribution de matière sont grandes (c’est-à-dire plus σ_8 est grand), plus les amas sont nombreux. On note enfin que les deux paramètres cosmologiques ont une influence différente sur l’abondance d’amas. Une variation de Ω_m influera principalement sur l’abondance d’amas de faible masse, alors qu’à l’inverse, une variation de σ_8 modifiera surtout l’abondance d’amas massifs. Les contraintes basées sur l’abondance d’amas sont de fait sensibles à une combinaison des deux

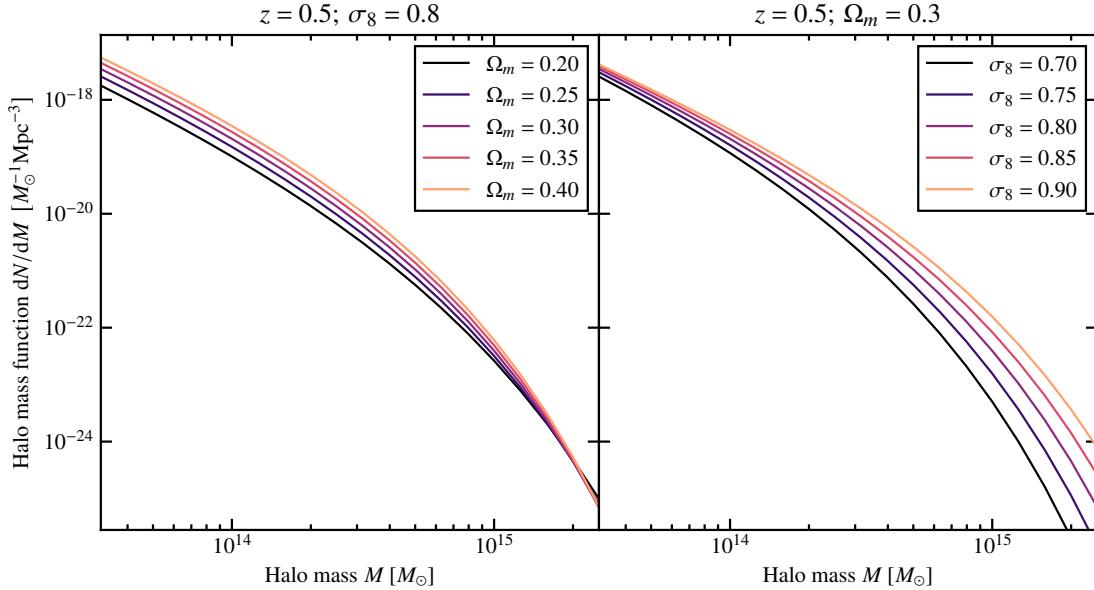


Figure 1.3 : Illustration de l'évolution de la fonction de masse avec les paramètres cosmologiques. La fonction de masse de Tinker *et al.* [32] est représentée pour différentes valeurs de Ω_m (gauche), et de σ_8 (droite). Le redshift est fixé à $z = 0.5$.

paramètres, pouvant varier d'une analyse à l'autre, et souvent exprimée par le paramètre S_8 :

$$S_8 \equiv \sigma_8 \sqrt{\Omega_m / 0.3}. \quad (1.34)$$

La connaissance de la fonction de masse est donc une nécessité pour la justesse des analyses cosmologiques basées sur les statistiques d'amas (voir par exemple [36, 37]). De telles analyses seront détaillées au chapitre suivant, et en particulier en section 2.3 (page 40).

1.4 Conclusions

Le modèle standard de la cosmologie, accompagné du paradigme de l'inflation, fournit une description de l'Univers et de son évolution basée sur des hypothèses simples. À l'heure actuelle, il est en remarquable accord avec les observations disponibles. Cependant, plusieurs questions restent en suspens. Par exemple, la nature de l'énergie sombre, cause de l'accélération récente de l'expansion, est une des énigmes non-résolues par le modèle Λ CDM. La mesure de son paramètre d'équation d'état, et de son éventuelle évolution dans le temps, peut apporter des réponses à cette question.

La mesure toujours plus précise des valeurs des paramètres cosmologiques est l'un des enjeux majeurs de la cosmologie moderne. Dans ce but, la combinaison de plusieurs observables, nommées *sondes cosmologiques*, s'impose comme un puissant moyen de mitiger les différentes incertitudes systématiques associées à chacune des sondes individuelles. Plusieurs études récentes semblent indiquer un désaccord entre les mesures des paramètres cosmologiques effectuées avec différentes sondes. Par exemple, les mesures du taux actuel d'expansion de l'Univers H_0 réalisées par l'analyse des aniso-

⁵Cette observation n'est plus vraie pour les objets extrêmement massifs, au delà de $2 \times 10^{15} M_\odot$ à $z = 0.5$. On note que de tels objets sont des cas extrêmes ; leur existence seule pousserait le modèle Λ CDM à ses limites, et des objets plus massifs pourrait même permettre de l'invalider – voir par exemple [35].

tropies du CMB sont en désaccord avec les mesures réalisées dans l’Univers contemporain (voir par exemple [38-41]). On note également la différence entre les estimations du paramètre S_8 par l’analyse du CMB et d’autres sondes, qui sera discutée au chapitre suivant. Ces tensions peuvent être un signe de nouvelle physique, ou bien indiquer des effets systématiques non-maîtrisés. Alors même que les incertitudes statistiques s’apprêtent à être réduites de manière conséquente avec les différents relevés cosmologiques du futur, ces systématiques deviendront de plus en plus importantes, et devront être étudiées afin de pouvoir répondre aux questions toujours ouvertes sur les propriétés les plus fondamentales de l’Univers. Cette thèse s’inscrit dans le cadre de l’étude de ces effets systématiques dans l’exploitation cosmologique des amas de galaxies, qui sera décrite dans le chapitre suivant.

CHAPITRE 2

Sonder la cosmologie à l'aide d'amas de galaxies

2.1	Composition et masse des amas de galaxies	26
2.1.1	Définition d'un amas de galaxies	26
2.1.2	Masse et rayon des amas de galaxies	26
2.1.3	Le biais hydrostatique	28
2.2	Observations des amas de galaxies	29
2.2.1	Observations en visible et infrarouge	29
2.2.2	Observations en X	31
2.2.3	Observations en millimétrique : l'effet Sunyaev-Zeldovich	32
2.2.4	Relevés d'amas de galaxies par effet SZ	37
2.3	Cosmologie par comptage d'amas de galaxies	40
2.3.1	Fonction de masse et comptage d'amas	40
2.3.2	Relation d'échelle masse-observable	42
2.4	État de l'art et perspectives	44
2.4.1	Relevés d'amas et résultats cosmologiques actuels	44
2.4.2	Relevés futurs	45
2.4.3	Suivis dédiés d'échantillons d'amas	46
2.4.4	Conclusions	47

Nous avons vu au chapitre 1 que la fonction de masse prédisait l'abondance de halos de matière sombre par unité de masse et de redshift. Étant donné le lien étroit entre le processus de formation des structures et la cosmologie sous-jacente, on peut prédire que la fonction de masse dépend des propriétés de l'Univers et de son évolution depuis l'inflation. Elle fait donc le lien entre la distribution de matière dans l'Univers et les paramètres cosmologiques à travers l'abondance de surdensités par unités de masse et de redshift.

C'est dans ce contexte que les amas de galaxies apparaissent comme une sonde cosmologique. Ils constituent en effet les plus grandes structures gravitationnellement liées de l'Univers, et apparaissent comme l'aboutissement du processus de formation hiérarchique des structures. Le nombre d'amas de galaxies observables dans le ciel est donc directement lié à la fonction de masse. Par conséquent, un catalogue d'amas de galaxies de masse et redshift connus permet de contraindre la fonction de masse, et donc les paramètres cosmologiques.

Ce chapitre décrit les éléments nécessaires à la réalisation de telles analyses. Nous nous intéresserons tout d'abord à la définition des amas de galaxies et à leurs caractéristiques. Nous discuterons ensuite des différentes observables des amas dans différentes longueurs d'onde, puis des contraintes cosmologiques par comptage d'amas de galaxies. Nous terminerons par un tour d'horizon des résultats récents en cosmologie avec des amas de galaxies, et des perspectives de ce domaine.

2.1 Composition et masse des amas de galaxies

2.1.1 Définition d'un amas de galaxies

Les amas de galaxies représentent la dernière étape du processus de formation hiérarchique des structures développé dans la section 1.2, et constituent les plus grandes structures gravitationnellement liées de l'Univers. Ils se forment par effondrement gravitationnel de la matière et par accrétion du milieu environnant autour des pics de densité de la toile cosmique, c'est-à-dire aux intersections de ses filaments. Ils offrent donc un traceur de la distribution de matière dans l'Univers. C'est la raison pour laquelle leur distribution en masse et en redshift constitue une sonde permettant d'étudier les processus de formation des structures, et donc la cosmologie sous-jacente. Ce point sera abordé plus en détail en section 2.3.

La formation des amas de galaxies a lieu dans l'Univers contemporain, à des redshifts $z \lesssim 3$. Les structures de plus haut redshift sont essentiellement représentées par des galaxies seules, ou par des groupes d'un petit nombre de galaxies, désignés comme proto-amas (voir par exemple [42]).

La composition des amas de galaxies est représentative de celle de la matière dans l'Univers (voir voir par exemple [43]). Elle est dominée par la matière sombre, comptant pour $\sim 85\%$ de leur masse. La matière baryonique des amas est principalement contenue dans un gaz chaud ionisé nommé milieu intra-amas (ou ICM pour *intracluster medium*), représentant $\sim 12\%$ du contenu en masse des amas. Ce gaz est principalement composé de noyaux légers et d'électrons libres, à des températures de l'ordre de $10^7 - 10^8$ K. Enfin, le reste de la matière baryonique des amas est contenu dans leurs galaxies, qui ne représentent qu'une faible fraction de la masse totale des amas ($\sim 3\%$).

2.1.2 Masse et rayon des amas de galaxies

Les amas de galaxies sont des objets complexes et diffus par nature, auxquels il est difficile d'attribuer une taille physique. Leur étude passe donc par la définition d'un rayon caractéristique à l'intérieur duquel leurs propriétés physiques seront étudiées. Pour cela, le rayon de viriel peut être utilisé, décrivant un halo à l'équilibre. En pratique, celui-ci est difficile à estimer, du fait des différents processus physiques perturbant leur équilibre, et de sa taille souvent plus grande que les champs couverts par les observations. Par conséquent, il est plus utile de définir un rayon R_Δ par la valeur moyenne du contraste de densité Δ qu'il contient :

$$\Delta \equiv \frac{\langle \rho(< R_\Delta) \rangle}{\rho_{\text{crit}}(z)}, \quad (2.1)$$

ce qui correspond à considérer l'amas comme une boule homogène de masse volumique Δ fois supérieure à la densité critique de l'Univers.

La valeur de contraste considérée est arbitraire ; les valeurs usuelles sont $\Delta \in [200, 500, 2500]$, en fonction de la couverture angulaire des observations utilisées et des régions d'intérêt. Des études portant sur les propriétés des amas dans R_{2500} se focaliseront sur le cœur des amas, alors que des études à R_{200} permettront des contraintes sur leur périphérie. Au cours de cette thèse, nous utiliserons principalement R_{500} , valeur conventionnellement utilisée pour des études sur les régions intermédiaires des amas de galaxies et du milieu intra-amas. Comme nous le verrons en section 2.4.1 (page 44), les amas de galaxies détectés par les instruments actuels ont des masses M_{500} typiquement de l'ordre de $10^{14} - 10^{15} M_\odot$.

Parmi les propriétés fondamentales des amas de galaxies, leur masse joue un rôle crucial dans l'estimation de contraintes sur les paramètres cosmologiques. La connaissance des masses individuelles d'un grand nombre d'amas de galaxies est en effet nécessaire aux études de comptage, qui constituent l'une des sondes cosmologiques majeures basées sur les amas de galaxies, et qui sera détaillée en section 2.3.

Plusieurs méthodes existent pour estimer la masse d'amas individuels ; nous renvoyons le lecteur vers la revue récente de Pratt *et al.* [43] pour une liste détaillée des différentes possibilités ouvertes par les observations récentes d'amas de galaxies à différentes longueurs d'onde. La mesure principale utilisée dans cette thèse repose sur la connaissance des propriétés thermodynamiques du milieu intra-amas, et sur l'hypothèse de l'équilibre hydrostatique. Dans le cadre de cette hypothèse, les forces gravitationnelles subies par un élément de fluide sont compensées par les forces de pression :

$$\frac{1}{\rho} \vec{\nabla} P = \vec{g} = \frac{-GM^{\text{HSE}}}{r^2} \vec{u}_r, \quad (2.2)$$

où \vec{u}_r est un vecteur unitaire radial en coordonnées sphériques.

Alors, dans l'hypothèse de symétrie sphérique de l'amas, le profil de masse, soit la valeur de masse contenue dans une sphère de rayon r , s'écrit

$$M^{\text{HSE}}(< r) = \frac{-r^2}{G\mu m_p n_e(r)} \frac{dP_e}{dr}, \quad (2.3)$$

où μ est le poids moléculaire moyen du milieu intra-amas, $\mu \equiv m/m_p$; et $n_e(r)$ et $P_e(r)$ sont respectivement la densité et la pression des électrons du milieu intra-amas.

L'hypothèse de l'équilibre hydrostatique offre donc un estimateur du profil de masse des amas de galaxies. La masse M_Δ contenue dans un rayon R_Δ peut alors être estimée grâce à la connaissance du profil de masse d'un amas de galaxies. En effet, le rayon R_Δ peut être calculé en considérant une boule homogène de rayon R_Δ et de masse volumique $\rho_c(z)$, soit en résolvant :

$$\frac{M(< R_\Delta)}{\rho_c(z) \times \frac{4}{3}\pi R_\Delta^3} = \Delta. \quad (2.4)$$

La valeur du profil de masse à ce rayon donne donc la masse $M_\Delta = M(< R_\Delta)$. On note au passage la corrélation entre R_Δ et M_Δ par construction. Ce point revêt une importance capitale pour les mesures de relations d'échelle masse-observable, et sera discuté dans les chapitres 5 et 7.

La masse hydrostatique se présente donc comme un estimateur de la masse d'un amas reposant sur la mesure du profil de densité et de pression des électrons de l'ICM. Comme nous le verrons par

la suite, il est possible de mesurer ces propriétés grâce à l'émission des amas de galaxies en X et à l'observation de l'effet Sunyaev-Zeldovich.

2.1.3 Le biais hydrostatique

La validité de l'hypothèse de l'équilibre hydrostatique dans les amas est le sujet de nombreuses études (voir par exemple [44]). En effet, comme nous le verrons par la suite, les analyses cosmologiques par comptage d'amas sont limitées par la connaissance des masses des amas. Une grande variété des processus physiques ayant lieu au sein de ces derniers (fusion de sous-structures, injection d'énergie par des noyaux actifs de galaxies) créent dans le milieu intra-amas des forces de pression non-thermiques. Celles-ci ne sont en général pas contraintes par les observations, comme nous le verrons par la suite. Les masses estimées par observation des propriétés thermodynamiques du milieu intra-amas peuvent donc souffrir d'un biais. Celui-ci est nommé biais hydrostatique b , et permet de quantifier l'état de perturbation des amas :

$$M^{\text{HSE}} = (1 - b)M^{\text{total}}. \quad (2.5)$$

On remarque alors que ce biais se propage au profil de masse hydrostatique discuté dans la section précédente, et donc au calcul du rayon caractéristique R_Δ décrit dans l'équation (2.4). La masse caractéristique M_Δ obtenue grâce au profil de masse hydrostatique est donc affectée de ce biais non-seulement de par la différence entre les profils de masse, mais également par la différence des rayons caractéristiques inférés.

La mesure du biais hydrostatique peut être réalisée par plusieurs moyens. On peut étudier le biais à partir de simulations hydrodynamiques, dans lesquelles la masse réelle des amas est connue et la masse hydrostatique peut être calculée à partir de l'équation (2.3) (voir par exemple [45, 46]). Cependant, cette approche suppose que les simulations sont parfaitement réalistes, et les résultats dépendent fortement de la modélisation des différents processus physiques inclus dans les simulations utilisées [46]. Le biais hydrostatique peut également être étudié à partir d'observations. En effet, il existe plusieurs estimateurs de masse des amas qui ne reposent pas sur l'hypothèse de l'équilibre hydrostatique. Par exemple, les mesures de masse par effet de lentillage gravitationnel du CMB (voir par exemple [47-49]) ou de galaxies d'arrière-plan¹ (voir par exemple [50]). Les différents estimateurs de masse peuvent être comparés sur des échantillons d'amas afin de mesurer l'écart moyen entre les mesures de masse hydrostatique et celles ne reposant pas sur cette hypothèse.

Une compilation de mesures de $(1 - b)$ est présentée en figure 2.1, montrant les différentes valeurs de biais obtenues par comparaison d'estimateurs de masse à partir d'observations et de simulations. La région orange montre la valeur du biais hydrostatique obtenue par l'analyse cosmologique jointe des observations par *Planck* de l'effet Sunyaev-Zeldovich et du CMB réalisée par Salvati *et al.* [51]. Une tension apparaît : ce résultat est incompatible avec les mesures de biais réalisées par la comparaison de différents estimateurs de masse d'amas observés, à partir d'observations (points noirs) et de simulations (région violette). S'il est à noter que cette différence a récemment été réduite par la réanalyse des observations du CMB par *Planck*², la détermination de la valeur du biais hydrostatique reste l'un des enjeux majeurs de la cosmologie avec des amas de galaxies. En effet, elle peut avoir un impact prononcé sur les mesures de paramètres cosmologiques, et une mauvaise estimation de b entraîne un biais sur la mesure du paramètre S_8 (introduit en 1.3.3, et qui sera discuté en 2.4.1).

¹Plus de détails seront donnés en section 2.2.1 (page 29).

²En particulier, la valeur de la profondeur optique jusqu'à la réionisation, τ , a évolué entre les analyses de la collabo-

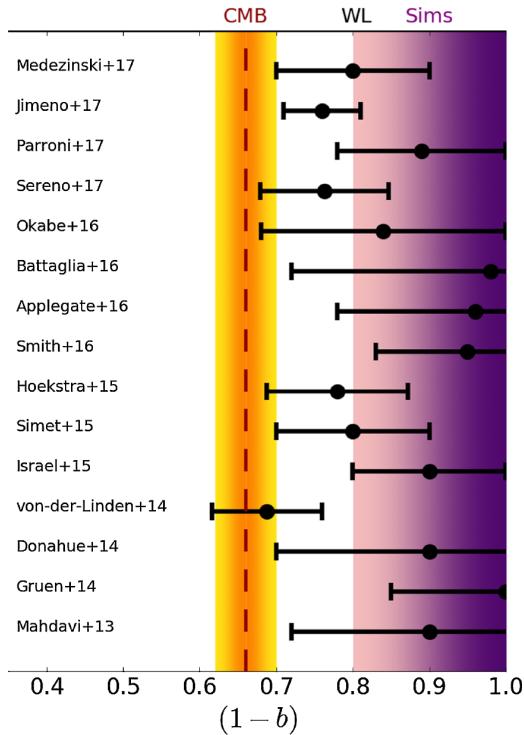


Figure 2.1 : Estimation du biais hydrostatique par comparaison de mesures de masse hydrostatique et par lentillage faible (voir section 2.2.1) pour différentes études (noir). La valeur nécessaire pour concilier les résultats du CMB et des amas de galaxies, mesurée par [51], est indiquée en orange. La région violette représente les valeurs de $(1 - b)$ permises par plusieurs analyses d'amas simulés [52]. Figure extraite de [51].

2.2 Observations des amas de galaxies

La richesse des processus physiques ayant lieu au sein des amas de galaxies font de ceux-ci des objets multilongueur d'onde par nature, pouvant être détectés dans tous des domaines du spectre électromagnétique. Dans cette section, nous présentons les longueurs d'onde principales utilisées pour la caractérisation des amas de galaxies, ainsi que les propriétés physiques des amas qu'elles permettent respectivement de sonder. Une attention particulière sera portée à l'effet Sunyaev-Zeldovich, qui constitue l'observable principale étudiée au cours de cette thèse, en 2.2.3.

À titre d'illustration, les cartes d'un amas de galaxies dans les trois domaines de longueur d'onde discutés ici sont présentées en figure 2.2. On y voit que les amas apparaissent comme des objets diffus dans les observations dans les domaines millimétrique et X. Comme nous le verrons par la suite, la raison en est que de telles observations cartographient le gaz du milieu intra-amas ; c'est pourquoi la structure des amas à ces longueurs d'onde sont souvent similaires. Les observations en visible des amas révèlent quant à elles la distribution des galaxies membres de l'amas.

2.2.1 Observations en visible et infrarouge

De la première détection d'un amas de galaxies à la fin du 18^e siècle à l'avènement des premiers satellites X dans la deuxième moitié du 20^e siècle, les amas de galaxies étaient uniquement observés dans des longueurs d'ondes visibles [53]. L'émission visible et infrarouge des amas provient de leur composante stellaire, contenue dans les galaxies membres des amas. Elle peut être étudiée par observations photométriques ou spectroscopiques, donnant chacune accès à des informations différentes sur les amas de galaxies observés.

Les mesures photométriques, basées sur l'acquisition d'images du ciel dans différentes bandes de longueurs d'onde, permettent l'étude des images des galaxies du champ. La position de ces galaxies

ration Planck en 2015 et 2018. Plus de détails peuvent être trouvés dans [13].

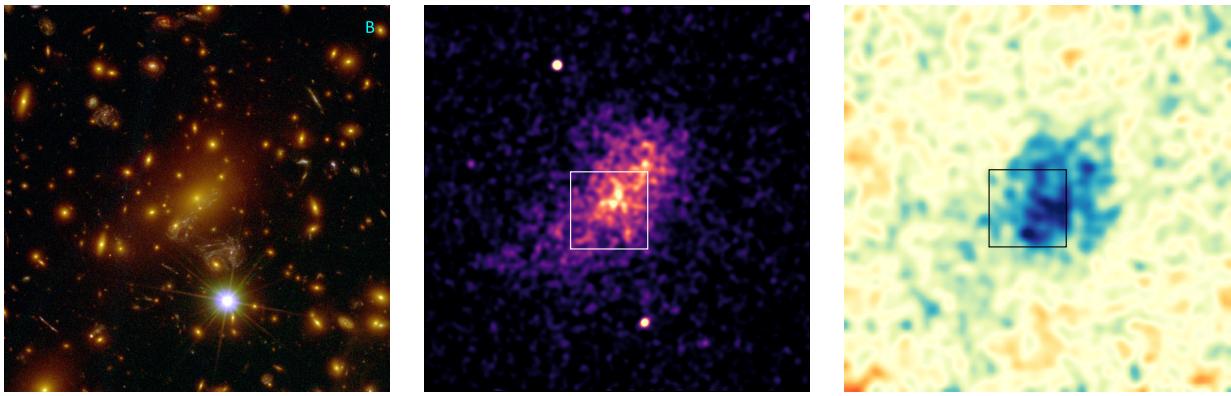


Figure 2.2 : Cartographie de l’amas de galaxies MACSJ1149.5+2223 par le *Hubble Space Telescope* en optique et infrarouge (gauche, extraite de [58]), en X avec *Chandra* (centre), et en millimétrique par l’effet Sunyaev-Zeldovich avec NIKA2 (droite). La taille de la région cartographiée est de 1 arcmin pour la carte optique, et de 5 arcmin pour les cartes X et SZ. La région cartographiée en optique est représentée par un carré sur les cartes X et SZ.

dans un diagramme couleur-magnitude permet de séparer les galaxies membres de l’amas de celles d’avant- ou arrière-plan ; c’est sur ce principe que sont basés certains algorithmes de détection des amas en optique, par exemple redMaPPer [54]. D’autres algorithmes existent, notamment basés sur la géométrie de la distribution de galaxies (voir par exemple [55]). La distribution des galaxies membres peut être utilisée comme un traceur du potentiel gravitationnel, offrant des contraintes sur la distribution de masse et sur l’état dynamique des amas. En particulier, la richesse d’un amas, définie comme le nombre de galaxies membres observables, fait partie des observables étroitement liées à la masse des amas (cf. section 2.3.2). Elle est par conséquent fréquemment utilisée dans les relevés optiques pour estimer la masse d’amas pour lesquels des mesures individuelles de masse ne sont pas possibles³.

L’étude des galaxies d’arrière-plan offre la possibilité de mesurer la masse d’un amas grâce au lentillage gravitationnel. En effet, les photons en provenance de ces sources distantes sont déviés par le fort potentiel gravitationnel des amas, résultant en des distorsions de leurs images. Ces distorsions peuvent être séparées en deux grands régimes. Dans le cas de faibles distorsions, souvent caractérisées par une simple déformation des galaxies d’arrière-plan, on parle de lentillage gravitationnel faible (ou *weak lensing*). Celui-ci n’est souvent détectable que par étude statistique des déformations des galaxies d’arrière-plan, ne permettant pas toujours une mesure des masses individuelles des amas de galaxies (voir [59] pour une revue sur le lentillage faible autour des amas). Le régime de lentillage fort (ou *strong lensing*) est quant à lui caractérisé par l’apparition de plusieurs images d’une même galaxie d’arrière-plan, ou bien d’une déformation extrême comme l’apparition d’arcs, ou plus rarement d’anneaux d’Einstein. Il permet alors une étude plus détaillée de la distribution du potentiel gravitationnel au sein d’amas individuels. Le lentillage gravitationnel n’est pas nécessairement présent et détectable autour d’un amas de galaxies. En particulier, il nécessite la présence de sources d’arrière-plan, qui sont plus rares pour les amas distants, limitant le potentiel des mesures de masses par lentillage à haut redshift. De plus, les mesures de déformations des galaxies d’arrière-plan sont fortement affectées par des effets systématiques, de par le fait qu’elles requièrent des hypothèses sur la forme réelle de ces galaxies, et qu’elles dépendent fortement des conditions d’observations et des caractéristiques des instruments utilisés (voir par exemple [60-63]). Ainsi, si les estimations de masses des amas ne sont pas biaisées⁴, elles sont en général plus dispersées que les estimations basées sur les mesures des propriétés thermodynamiques du milieu intra-amas ([43, 59, 62]).

³Voir par exemple [56, 57] pour des résultats récents.

⁴Au sens où elles mesurent la masse gravitationnelle de l’amas sans reposer sur l’hypothèse de l’équilibre hydrostatique ;

Enfin, les amas de galaxies peuvent être observés aux longueurs d'onde visibles et infrarouges par spectroscopie. Ce sont alors les spectres des galaxies qui sont mesurés, c'est-à-dire l'évolution de leur flux avec la fréquence. De tels relevés sont capables de mesurer les redshifts des galaxies des amas, à l'aide de l'identification des longueurs d'onde observées de raies spectrales connues. Les observations spectroscopiques d'amas de galaxies offrent aussi un moyen d'estimer la dispersion des vitesses des galaxies membres par mesure de l'effet Doppler. Cette dispersion étant liée à la distribution de masse dans les amas, elle peut également être utilisée comme observable reliée à la masse pour des analyses cosmologiques (voir par exemple [64, 65]).

Plusieurs grands relevés dans les domaines visible et infrarouge sont sur le point de commencer ; par exemple, le Legacy Survey of Space and Time réalisé par le Vera Rubin Observatory [66] et le relevé *Euclid* [67]. Il est prévu que ces deux relevés détectent un nombre d'amas supérieur de plusieurs ordres de grandeurs au nombre d'amas actuellement connus, avec des mesures de masses par lentillage d'une grande qualité grâce aux performances des instruments [8, 9].

2.2.2 Observations en X

L'un des domaines de longueur d'onde couramment utilisés pour l'étude des amas de galaxies est le domaine des rayons X. Les amas de galaxies sont émetteurs en X principalement au travers du rayonnement de freinage (*bremsstrahlung*) des électrons dans le gaz chaud du milieu intra-amas (voir [68] pour une revue). La brillance de surface observée en direction d'un amas S_X est liée aux propriétés thermodynamiques du milieu intra-amas :

$$S_X = \frac{1}{4\pi(1+z)^4} \int \Lambda(T_e, Z) n_e^2 dl, \quad (2.6)$$

où n_e est la densité des électrons du milieu intra-amas, et $\Lambda(T_e, Z)$ est la fonction de refroidissement dépendant de la température des électrons du milieu T_e et de sa métallicité Z , quantifiant l'abondance d'éléments plus lourds que l'hélium dans le milieu intra-amas.

La fonction de refroidissement peut être déterminée par étude spectroscopique du signal X, qui sera détaillée au prochain paragraphe. L'intégration a lieu le long de la ligne de visée l . Comme le montre l'équation (2.6), l'amplitude du signal X mesuré en direction d'un amas renseigne sur la densité d'électrons présente le long de la ligne de visée considérée. Les observations d'amas en X permettent donc la mesure de la distribution des électrons dans le milieu intra-amas, nécessaire au calcul de la masse détaillé plus tôt (équation 2.3).

En plus du continuum d'émission X due au *bremsstrahlung* thermique des électrons, l'émission X dans le milieu intra-amas comporte des raies spectrales caractéristiques des transitions entre niveaux d'énergie des éléments présents dans le gaz. Lorsqu'une observation d'un amas a permis d'accumuler suffisamment de photons X, il est alors possible d'étudier leur distribution en énergie, qui représente le spectre du milieu intra-amas. La connaissance de ce spectre permet de quantifier l'abondance des éléments émettant les raies spectrales étudiées, et donc la métallicité Z de l'amas. De plus, le spectre étant intimement lié à la température des électrons du gaz, les études de spectroscopie X permettent également de la mesurer⁵. Ainsi, des observations profondes d'amas de galaxies en X permettent de mesurer à la fois la densité et la température électronique du milieu intra-amas. Ce milieu étant très dilué, avec des densités rarement supérieures à 10^{-2} électrons par centimètre cube, il peut être modélisé

elles sont toutefois affectées d'autres biais – voir voir par exemple [43, 62, 63].

⁵Voir [69] pour une revue de la spectroscopie X

comme un gaz parfait. Les mesures de densité et de pression peuvent alors être combinées pour obtenir une estimation de la pression électronique :

$$P_e = n_e k_B T_e, \quad (2.7)$$

qui peut à son tour être combinée à la densité d'électrons pour calculer la masse des amas par l'équation (2.3).

Étant donnée l'opacité de l'atmosphère aux rayons X, les observations d'amas à ces longueurs d'onde ne peuvent être réalisées que depuis des satellites. La plupart des études d'amas de galaxies en X font actuellement appel à des mesures avec les satellites *XMM-Newton* [70] et *Chandra* [71]. Dans le futur, les mesures spectroscopiques fournies par la mission *Athena* [72] permettront la caractérisation des propriétés du milieu intra-amas avec une grande précision. L'émission X des amas peut également être utilisée pour créer des catalogues d'amas, comme le relevé actuellement réalisé par l'instrument eROSITA [73] qui permettra la détection de plusieurs centaines de milliers d'amas.

Ainsi, l'étude du rayonnement X des amas de galaxies permet de mesurer les propriétés thermodynamiques de leur milieu intra-amas, jusqu'à leur masse, de manière autonome et sans avoir à recourir à des observations dans d'autres longueurs d'onde. Il est toutefois à noter que l'analyse spectroscopique des photons X en provenance d'amas requiert des observations très profondes, et est donc particulièrement coûteuse en temps. C'est l'une des raisons qui motive la combinaison d'observations X peu profondes, ne permettant de mesurer que la distribution de densité du milieu intra-amas, avec des observations effectuées à d'autres longueurs d'onde, sensibles à d'autres propriétés physiques. C'est par exemple le cas de l'effet Sunyaev-Zeldovich, détaillé dans la section suivante.

2.2.3 Observations en millimétrique : l'effet Sunyaev-Zeldovich

L'effet Sunyaev-Zeldovich (SZ, [74-77]) est une distorsion du fond diffus cosmologique due à l'interaction de ses photons avec les électrons des amas de galaxies. L'interaction ayant lieu est une diffusion Compton inverse : les photons du CMB, peu énergétiques ($\sim 10^{-6}$ keV lors de la formation des amas, à $z \lesssim 3$), diffusent sur les électrons libres du milieu intra-amas. Ces derniers possédant une grande énergie cinétique, les photons acquièrent de l'énergie, modifiant la forme du spectre observé du CMB dans la direction des amas.

L'effet SZ peut être séparé en plusieurs composantes, parfois considérées comme des effets SZ différents, selon le processus physique à l'origine de l'énergie cinétique des électrons transmise aux photons du CMB. Dans chacun des cas, le gain en énergie est lié aux propriétés physiques du milieu intra-amas, et l'observation des effets SZ permet de mesurer ces propriétés et donc de caractériser la physique des amas de galaxies. Nous listons ici les trois effets SZ les plus fréquemment utilisés : l'effet SZ thermique (tSZ), cinétique (kSZ) et relativiste (rSZ). Des revues détaillées des différents effets SZ et de leur utilisation en astrophysique et cosmologie peuvent être trouvées, par exemple, dans [78-80].

L'effet Sunyaev-Zeldovich thermique

Le milieu intra-amas est composé d'un gaz chaud ionisé, dont la température est généralement de l'ordre de $10^7 - 10^8$ K. Dans l'hypothèse d'un gaz parfait, ces températures correspondent à une énergie d'agitation thermique $3k_B T/2$ de l'ordre de 1 – 10 keV, bien supérieure à l'énergie des photons du CMB. La diffusion de Compton inverse permet alors aux électrons de céder une partie de cette énergie aux photons du CMB.

La dépendance spectrale de la distorsion peut être calculée à partir de l'équation de Kompaneets [81, 82] : pour un bain d'électrons non-relativistes ($k_{\text{B}}T_{\text{e}} \ll m_{\text{e}}c^2$) traversé par des photons de fréquence réduite $x_{\text{e}} \equiv h\nu/k_{\text{B}}T_{\text{e}}$, on a

$$\frac{\partial n}{\partial y} = \frac{1}{x_{\text{e}}^2} \frac{\partial}{\partial x_{\text{e}}} \left[x_{\text{e}}^4 \left(\frac{\partial n}{\partial x_{\text{e}}} + n^2 + n \right) \right], \quad (2.8)$$

où n est l'indice d'occupation des photons, défini pour une intensité spécifique I_{ν} à une fréquence ν comme

$$n \equiv \frac{I_{\nu}c^2}{2h\nu^3}, \quad (2.9)$$

et y est le paramètre de Compton, défini comme

$$y \equiv \int \frac{k_{\text{B}}T_{\text{e}}}{m_{\text{e}}c^2} d\tau_{\text{e}} \quad (2.10)$$

où $\tau_{\text{e}} \equiv n_{\text{e}}\sigma_{\text{T}}l$ est la profondeur optique⁶ parcourue par les photons le long de la ligne de visée l ; n_{e} est la densité d'électrons, et σ_{T} la section efficace de diffusion Thompson.

Cette équation peut alors aussi s'écrire

$$y = \int \frac{k_{\text{B}}T_{\text{e}}}{m_{\text{e}}c^2} n_{\text{e}}\sigma_{\text{T}} dl = \frac{\sigma_{\text{T}}}{m_{\text{e}}c^2} \int k_{\text{B}}T_{\text{e}} n_{\text{e}} dl. \quad (2.11)$$

Dans le cas de l'effet SZ, l'énergie des photons incidents est très faible par rapport à celle des électrons, soit $x_{\text{e}} \ll 1$ et donc $\partial n/\partial x_{\text{e}} \gg n, n^2$. L'équation (2.8) se simplifie alors, et en y injectant l'expression d'un spectre de corps noir pour les photons du CMB, on trouve l'expression de la variation d'intensité spécifique due à l'effet tSZ :

$$\frac{\Delta I_{\nu}^{\text{tSZ}}}{I_0} = y \times \frac{x^4 e^x}{(e^x - 1)^2} [x \coth(x/2) - 4] [1 + \delta_r(x, T)], \quad (2.12)$$

où $x \equiv h\nu/k_{\text{B}}T_{\text{CMB}}$, et $\delta_r(x, T)$ traduit les corrections relativistes à l'effet tSZ, qui seront discutées par la suite, et $I_0 \equiv 2(k_{\text{B}}T_{\text{CMB}})^3/(hc)^2$ est l'intensité spécifique du CMB.

La distorsion spectrale due à l'effet tSZ est représentée sur la figure 2.3. On peut voir sur le panneau haut que le spectre de photons ayant interagi par effet tSZ est décalé vers les grandes fréquences, correspondant à un gain d'énergie par diffusion Compton inverse. Le panneau bas montre la distorsion spectrale relative provoquée par l'effet tSZ (équation 2.12). Celle-ci se manifeste par un déficit de photons de basse énergie, et une augmentation du nombre de photons à haute énergie, par rapport au spectre du CMB avant interaction. Ainsi, en observant le fond diffus cosmologique en direction d'amas de galaxies, on détectera un décrément dans sa brillance de surface aux fréquences inférieures à 217 GHz, et un incrément aux fréquences supérieures. Cette signature spectrale est très caractéristique de l'effet tSZ, et permet de le différencier des autres sources astrophysiques, comme nous le verrons en 2.4.1.

D'après l'équation (2.12), l'amplitude de la distorsion due à l'effet tSZ est donnée par le paramètre de Compton y . Ce lien est également illustré en figure 2.3, montrant un incrément et un décrément plus importants pour des valeurs de y plus élevées. En supposant que le gaz chaud du milieu intra-amas est bien décrit par un gaz parfait (hypothèse discutée en 2.2.2), l'expression du paramètre de Compton

⁶Dans l'hypothèse d'un milieu optiquement mince.

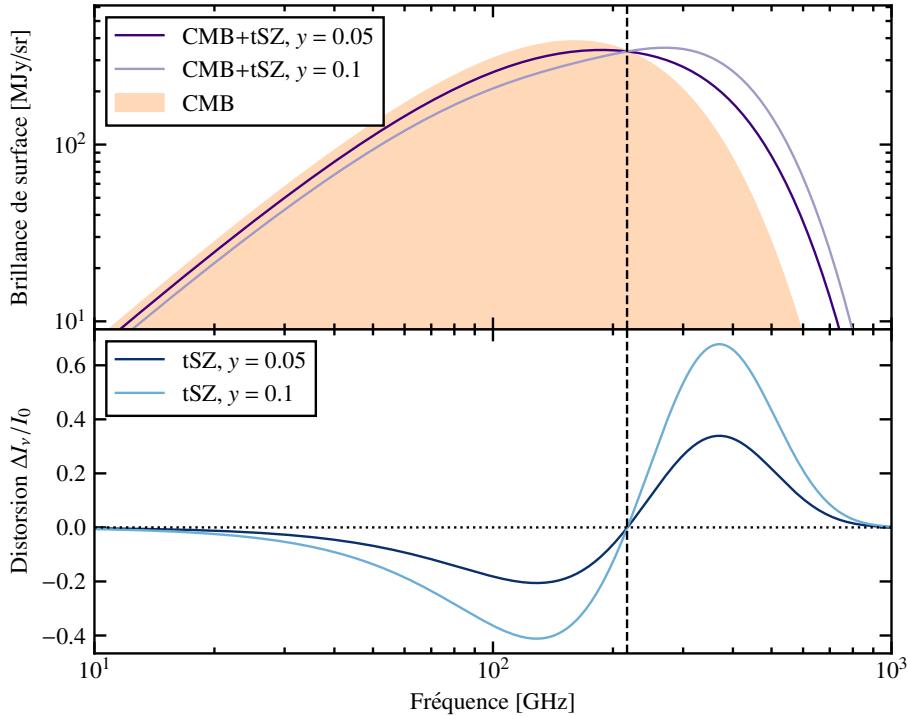


Figure 2.3 : **Haut** : Spectre des photons du CMB (orange) et de photons ayant interagi avec un amas pour deux valeurs de paramètres de Compton y (violet). Au premier ordre, l'effet tSZ représente un décalage du spectre vers les grandes fréquences. **Bas** : Distorsion spectrale due à l'effet tSZ pour les mêmes valeurs de paramètre de Compton. La signature observationnelle de l'effet tSZ apparaît comme un décrément dans la brillance de surface du CMB aux fréquences inférieures à 217 GHz (ligne noire verticale), et un incrément au-delà. Les valeurs de paramètre de Compton utilisées sont grandement exagérées pour des raisons visuelles. Les corrections relativistes à l'effet tSZ sont négligées.

(équation 2.11) prend la forme :

$$y(\theta) = \frac{\sigma_t}{m_e c^2} \int P_e d\ell(\theta), \quad (2.13)$$

où θ représente les coordonnées observées dans le ciel, et P_e est la pression due aux électrons, donnée par l'équation (2.7).

La cartographie de l'effet SZ en direction d'amas de galaxies permet donc de mesurer la distribution de pression électronique dans le milieu intra-amas.

Comme le montre l'équation (2.11), l'amplitude de la distorsion spectrale due à l'effet tSZ ne dépend pas du redshift des amas. Cette indépendance permet à l'effet tSZ d'être l'un des moyens privilégiés pour la détection d'amas de galaxies distants. En effet, alors même que la brillance de surface X décroît comme $(1+z)^{-4}$ (équation 2.6), et que les amas distants sont composés de galaxies moins brillantes, un amas de galaxies de haut redshift ne sera pas moins brillant en millimétrique qu'un amas proche de même masse. Les relevés d'amas de galaxies par effet SZ sont donc capables de détecter des amas plus distants et donc plus anciens que leurs équivalents optiques et X. De plus, les observations de l'effet tSZ permettent la mesure du paramètre de Compton intégré Y ,

$$Y_\Delta = 4\pi \frac{\sigma_t}{m_e c^2} \int_0^{R_\Delta} P_e(r) r^2 dr, \quad (2.14)$$

qui représente une mesure du contenu en énergie thermique dans le milieu intra-amas. Ainsi, l'effet

tSZ permet la détection d'amas de galaxies distants, ainsi qu'une mesure de l'énergie contenue dans leur milieu intra-amas.

L'effet Sunyaev-Zeldovich cinétique

En plus de l'énergie provenant de leur agitation thermique, les électrons libres du milieu intra-amas peuvent posséder une énergie cinétique du fait de mouvements d'ensemble du gaz. Une partie de cette énergie peut alors se transmettre aux photons du CMB traversant le milieu intra-amas, provoquant l'effet SZ cinétique, ou kSZ [77]. Contrairement à l'effet tSZ, le mouvement des électrons ne peut par construction pas être supposé isotrope. La distorsion spectrale du CMB s'écrit alors

$$\frac{\Delta I_\nu^{\text{kSZ}}}{I_0} = y_{\text{kSZ}} \frac{x^4 e^x}{(e^x - 1)^2} [1 + \delta'_r(x, T, v_z)], \quad (2.15)$$

où δ'_r représente les corrections relativistes à l'effet kSZ, discutées par la suite, et v_z est la vitesse de déplacement du gaz le long de la ligne de visée, dont dépend l'amplitude de la distorsion dans une direction θ :

$$y_{\text{kSZ}}(\theta) \equiv \frac{v_z(\theta)}{c} \int d\tau_e = \frac{\sigma_T v_z(\theta)}{c} \int n_e(l) dl(\theta). \quad (2.16)$$

La comparaison des équations (2.12) et (2.15) permet de mettre en évidence la différence entre les signatures spectrales des deux effets. Celle-ci est illustrée en figure 2.4, montrant que l'effet kSZ peut se manifester comme un incrément ou un décrément en brillance de surface du CMB, selon la direction du déplacement de l'amas. En effet, d'après les équations (2.15) et (2.16), un mouvement de gaz en direction de l'observateur ($v_z > 0$) créera un incrément en brillance de surface, et à l'inverse, le gaz s'éloignant de l'observateur ($v_z < 0$) créera un décrément. Ainsi par exemple, dans le cas d'un amas en rotation selon un axe non parallèle à la ligne de visée, l'effet kSZ se manifestera, à une même fréquence, comme un incrément dans les régions où le gaz s'approche de l'observateur, et comme un décrément dans les régions s'en éloignant. On remarque enfin que le maximum de la distorsion due à l'effet kSZ se situe à la fréquence à laquelle celle de l'effet tSZ est nulle, c'est-à-dire 217 GHz. Une mesure du ciel à cette fréquence permet donc *a priori* de mesurer l'effet kSZ sans contamination par le tSZ. Cependant, pour les amas de grande extension spatiale – proches et massifs – l'effet kSZ est confondu avec les anisotropies primaires du CMB, du fait de la similarité entre les spectres de puissance de ces deux types d'anisotropie.

Comme le montre l'équation (2.16), la mesure de l'effet kSZ permet de mesurer le produit de la densité d'électrons et de la vitesse de déplacement du gaz. En combinaison avec des mesures de champs de vitesses dans les amas réalisées en optique (cf. 2.2.1), l'effet kSZ offre donc l'opportunité de mesurer la densité du milieu intra-amas. D'après l'équation (2.3), la combinaison des effets tSZ et kSZ peut donc fournir une mesure de la masse des amas dans l'hypothèse de l'équilibre hydrostatique. Cependant, l'amplitude de l'effet kSZ est bien plus faible que celle de l'effet tSZ. En effet, le rapport entre les deux est de l'ordre de $(k_B T_e / m_e c^2) / (v_z / c) \sim 10$ pour des vitesses de l'ordre de la centaine de km/s et des températures de l'ordre de la dizaine de keV (propriétés caractéristiques d'une grande majorité des amas de galaxies). Cette différence d'amplitude entre les effets tSZ et kSZ pour des propriétés d'amas réalistes est visible sur la figure 2.5.

La mesure d'un signal kSZ significatif requiert donc une grande sensibilité, de même que des mesures à plusieurs longueurs d'onde pour pouvoir séparer sa contribution au signal total de celle de l'effet tSZ. Des détections de l'effet kSZ existent, basées sur l'empilement de signaux en provenance de

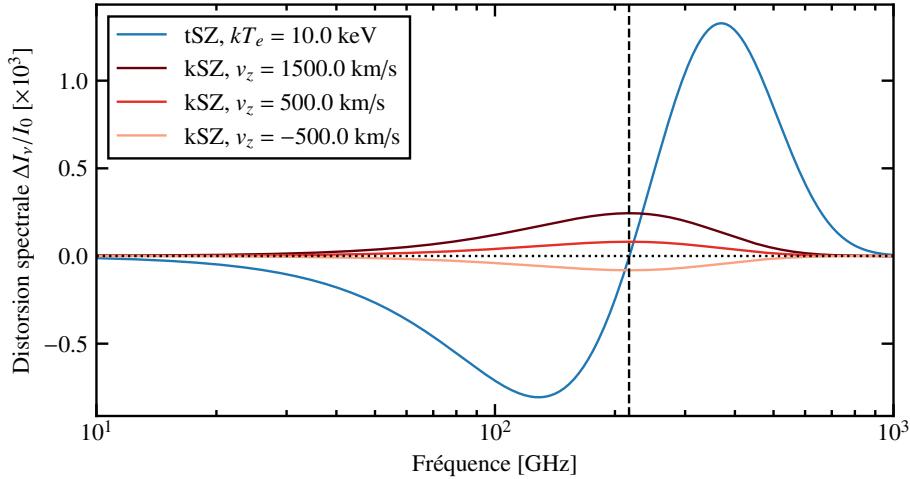


Figure 2.4 : Signatures spectrales de l’effet tSZ (bleu) et kSZ (rouge) pour un amas de profondeur optique $\tau_e = 10^{-2}$. L’amplitude de l’effet tSZ est calculée en supposant un milieu intra-amas isotherme avec $k_b T_e = 10$ keV. Les amplitudes pour l’effet kSZ sont calculées pour différentes valeurs le long de la ligne de visée. Dans les deux cas, les corrections relativistes sont négligées.

plusieurs amas (voir par exemple [83]) ou sur les mesures profondes d’amas dans lesquels des grandes vitesses de mouvements sont attendues (comme l’amas MACS J0717.5+3745, [84, 85]). Dans le cadre des observations traitées au cours de cette thèse, il est souvent négligeable : au vu du faible rapport signal sur bruit des observations, le niveau de signal kSZ sera inférieur au bruit.

L’effet SZ relativiste

Les corrections relativistes aux effets tSZ et kSZ, apparaissant respectivement dans les équations (2.12) et (2.16), sont parfois regroupées sous le terme d’effet SZ relativiste, ou effet rSZ. La raison de cette dénomination est la signature spectrale particulière de ces corrections relativistes, permettant de les différencier des contributions au premier ordre des effets tSZ et kSZ, ainsi que le fait qu’elles encodent une information sur les propriétés du milieu intra-amas différentes de ces derniers. Il est toutefois à noter qu’il s’agit bel et bien seulement de corrections relativistes aux effets susmentionnés : l’effet rSZ ne peut être détecté seul sans la présence des effets principaux.

Nous avons jusqu’à présent négligé les corrections relativistes aux effets SZ. Cette hypothèse est facilement vérifiée dans le cas de l’effet kSZ, où les mouvements d’ensemble du milieu intra-amas ont des vitesses de l’ordre de la centaine de kilomètres par seconde, soit $\beta = v/c \sim 10^{-4}$. Elle est moins évidente dans le cas de l’effet tSZ, où une température de $k_b T_e \sim 1 - 10$ keV conduit à des vitesses⁷ $\beta = \sqrt{3k_b T_e / 2m_e c^2} \sim 0.1$.

Le calcul des corrections relativistes ne sera pas détaillé ici, et nous renvoyons le lecteur vers des revues telles que [80] pour plus de détails. En pratique, l’effet rSZ dépend principalement de la température du milieu intra-amas, et a pour conséquence l’étalement des spectres des effets tSZ et kSZ, ainsi que le décalage de la fréquence à laquelle la distorsion due à l’effet tSZ s’annule. Ces effets sont illustrés sur la figure 2.5 : pour des températures croissantes, les amplitudes des spectres tSZ et kSZ sont diminuées aux fréquences inférieures à ~ 400 GHz, et augmentées au-delà. La fréquence correspondant au zéro de l’effet tSZ est également augmentée, de 217 GHz en négligeant les corrections

⁷D’après la théorie des gaz parfaits.

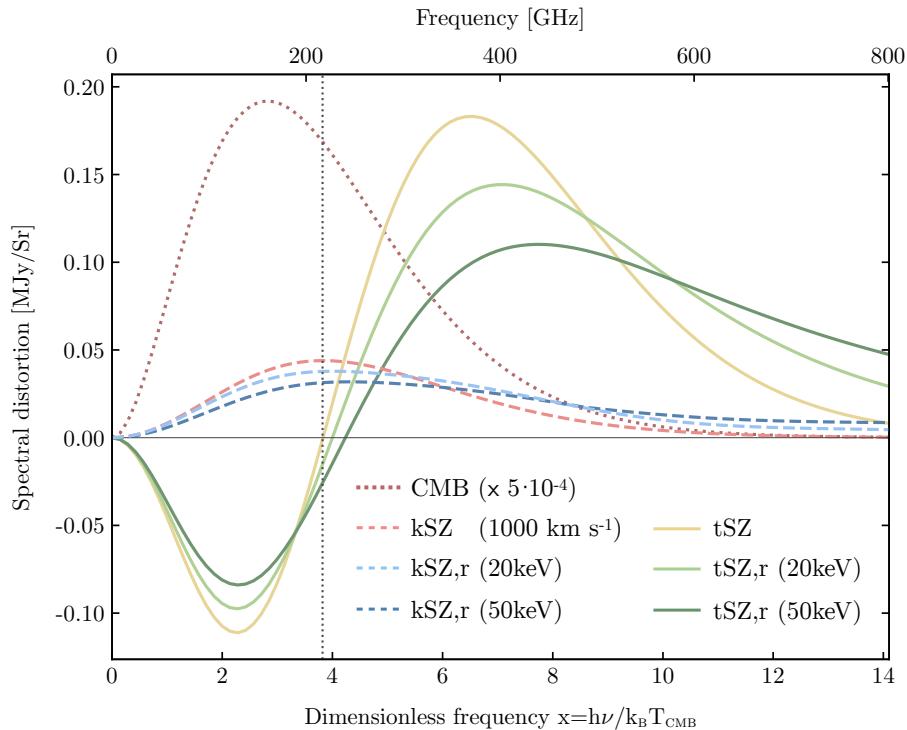


Figure 2.5 : Spectres de l'effet tSZ (courbes pleines) et kSZ (courbes pointillées), incluant les corrections relativistes pour différentes valeurs de température. La profondeur optique considérée est de $\tau_e = 10^{-2}$. Le spectre des photons du CMB sans interaction par effet SZ est également représenté en rouge pour comparaison. Figure extraite de [80].

relativistes, à ~ 240 GHz pour une température de $k_B T_e = 50$ keV. D'un point de vue observationnel, l'effet rSZ induit donc une modification des signaux tSZ et kSZ dépendant de la température du milieu intra-amas. Cette modification peut être détectée en observant les effets SZ en provenance d'amas à un grand nombre de fréquences, afin de pouvoir séparer les contributions des termes thermique, cinétique et relativiste. L'effet rSZ n'a à ce jour été détecté que dans des analyses sur l'empilement du signal SZ de plusieurs amas (voir par exemple [86]), ou à travers son impact sur le spectre de puissance de la cartographie de l'effet SZ sur tout le ciel (voir par exemple [87]). Des études sont en cours pour essayer d'en obtenir des mesures individuelles, qui permettraient une première mesure de la température du milieu intra-amas aux longueurs d'ondes millimétriques, par exemple avec le spectromètre CONCERTO [88].

D'autres composantes de l'effet Sunyaev-Zeldovich existent, telles que l'effet SZ polarisé ou non-thermique. Cette thèse étudie exclusivement la composante thermique ; toute mention ultérieure de l'effet SZ fera référence à l'effet tSZ, à moins que le contraire soit explicitement écrit.

2.2.4 Relevés d'amas de galaxies par effet SZ

Comme nous le verrons dans la section 2.3, la contrainte des paramètres cosmologiques à l'aide d'amas de galaxies requiert la connaissance d'un catalogue d'amas de masses et redshifts connus. De tels catalogues peuvent être construits à partir d'observations de grandes portions du ciel dans l'une des longueurs d'onde auxquelles les amas de galaxies sont détectables, puis par recensement des amas présents dans ce champ.

Détection par filtre adapté

Dans le cas de l'effet SZ, la détection des amas est faite à partir de cartes dans le domaine millimétrique. Les algorithmes de détection d'amas utilisés sont basés sur des techniques de filtre adaptés (*matched filter*, voir par exemple [89-91]). Cette approche consiste en l'exploitation d'une connaissance *a priori* de la distribution attendue du signal typique d'un amas de galaxies, et de la dépendance spectrale de l'effet étudié. Considérons un relevé du ciel à plusieurs fréquences du domaine millimétrique ν_i , $i \in [1 \dots N]$, comme le relevé réalisé à l'aide de l'instrument HFI du satellite *Planck* [92, 93]. Les cartes peuvent être regroupées dans un vecteur \mathbf{M} , dont la composante i représente la carte à la fréquence ν_i :

$$\mathbf{M}(\vec{x}) = y_0 \mathbf{S}(\vec{x}, \theta_c) + \mathbf{N}(\vec{x}), \quad (2.17)$$

où \vec{x} représente les coordonnées célestes observées, y_0 le paramètre de Compton central d'un amas, $\mathbf{S}(\vec{x}, \theta_c)$ la distribution spatiale du signal SZ d'un amas de taille θ_c , et $\mathbf{N}(\vec{x})$ un vecteur analogue à \mathbf{M} contenant les cartes de bruit aux différentes fréquences, incluant le bruit instrumental et astrophysique. On définit alors un filtre adapté $\Psi(x, \theta_c)$, dont la transformée de Fourier s'écrit :

$$\Psi(\vec{k}, \theta_c) = \sigma^2(\theta_c) \mathbf{P}^{-1}(\vec{k}) \mathbf{S}(\vec{k}, \theta_c), \quad (2.18)$$

où $\sigma^2(\theta_c)$ est la variance dans la carte de bruit convoluée par le filtre considéré, et $\mathbf{P}(\vec{k})$ le spectre de puissance de la carte de bruit \mathbf{N} .

La convolution des cartes formant le vecteur \mathbf{M} par ce filtre permet alors d'obtenir un estimateur non-biaisé du paramètre de Compton central des amas,

$$\hat{y}_0 = \int \Psi(\vec{x}, \theta_c) \mathbf{M}(\vec{x}) d\vec{x}. \quad (2.19)$$

La procédure est répétée pour plusieurs valeurs de taille de filtre θ_c , afin de pouvoir détecter aussi bien des amas étendus que compacts.

Les cartes ainsi filtrées permettent d'identifier les positions des pics de paramètre de Compton, correspondant aux positions abritant des amas de galaxies. La significativité de ces pics, donnée par les cartes du rapport $\hat{y}_0/\sigma(\theta_c)$, permet d'associer à chaque pic la probabilité qu'il soit effectivement lié à un amas. Ainsi, seuls les détections de significativité supérieure à un seuil fixé sont considérées comme des détections. Le choix de ce seuil reflète la nécessité d'un compromis entre la pureté (fraction de faux positifs, détections non associées à un amas de galaxies) et la complétude (fractions d'amas existants effectivement détectés, liée au taux de faux négatifs). En effet, plus le seuil sera bas, plus un catalogue sera complet (car incluant les amas de faible significativité), mais moins il sera pur (car composé de plus de détections ne correspondant pas à des amas).

Profil de pression moyen des amas

Les cartes composant les vecteurs \mathbf{M} et \mathbf{N} dans l'équation (2.17) étant connues⁸, l'enjeu est l'utilisation du meilleur modèle possible pour la distribution radiale du signal de l'amas, \mathbf{S} . En effet, une mauvaise connaissance de la distribution du signal SZ en direction d'un amas entraînera de mauvaises performances pour l'algorithme de détection.

⁸Le signal SZ étant largement sous-dominant dans les relevés du ciel par rapport aux signaux contaminants (comme l'atmosphère) ou astrophysiques (comme les anisotropies primaires du CMB), les spectres de puissance des cartes de bruit $\mathbf{N}(\vec{x})$ sont en général supposés égaux à ceux des cartes complètes $\mathbf{M}(\vec{x})$.

Du fait du lien entre l'amplitude de l'effet tSZ et la pression du milieu intra-amas (équation 2.11), la connaissance de la distribution radiale du signal requiert une connaissance *a priori* du profil de pression des amas⁹. En effet, connaître la distribution de pression dans le milieu intra-amas permet d'estimer la distribution de paramètre de Compton par intégration le long de la ligne de visée, et donc la distribution de brillance de surface de l'effet tSZ. La connaissance du profil de pression tridimensionnel moyen des amas $P(r)$ permet donc de remonter à un profil de paramètre de Compton dans le ciel $y(\vec{x})$. Celui-ci peut alors être utilisé pour calculer le modèle de distribution spatiale de signal SZ nécessaire au calcul du filtre adapté (équation 2.18) :

$$\mathbf{S}(\vec{x}, \theta_c) = f(\nu_i) [b_i * y(\vec{x}, \theta_c)], \quad (2.20)$$

où $f(\nu)$ est la dépendance spectrale de l'effet SZ (équation 2.12), et b_i le lobe¹⁰ de l'instrument utilisé pour réaliser les mesures à la fréquence ν_i .

De plus, les cartes obtenues par convolution par le filtre adapté permettent la mesure du paramètre de Compton intégré par intégration polaire du profil de paramètre de Compton. Par exemple, dans un rayon caractéristique R_{500} , on a

$$Y_{500}^{\text{cyl.}} = 2\pi \int_0^{\theta_{500}} \hat{y}(\theta) \theta d\theta, \quad (2.21)$$

où $\theta_{500} = c_{500} \theta_c$ est l'angle sous-tendu par le rayon caractéristique R_{500} , avec c_{500} le paramètre de concentration de l'amas.

On note une différence fondamentale entre les grandeurs Y_{500} et $Y_{500}^{\text{cyl.}}$ définis dans les équations (2.14) et (2.21). Dans la première, l'intégration réalisée est celle du profil de pression dans une sphère de rayon R_{500} , alors que la seconde revient à une intégration dans un cylindre de rayon R_{500} et de profondeur infinie. En pratique, la différence est faible, puisque la pression au-delà de R_{500} est plusieurs ordres de grandeur plus faible qu'à l'intérieur de ce rayon. Pour le profil de pression universel des amas de galaxies mesuré par Arnaud *et al.* [94] (discuté au paragraphe suivant), le rapport entre les deux est de 0.83 (voir section 6 de [94]). Dans la suite, nous considérerons des valeurs calculées par intégration sphérique.

Le profil de pression moyen du milieu intra-amas est donc l'un des ingrédients nécessaires aux études cosmologiques basées sur l'effet SZ. Dans l'hypothèse autosimilaire, la distribution radiale des propriétés thermodynamiques du milieu intra-amas doit pouvoir être bien décrite par un profil universel, mis à l'échelle par rapport à leur masse et taille caractéristique. Les écarts à cette hypothèse proviennent de la diversité des phénomènes physiques liés aux baryons et pouvant modifier le contenu énergétique des amas. Le profil de pression, étant directement lié au potentiel gravitationnel de l'amas par l'équation de Poisson, est supposé être le moins affecté par ces phénomènes, et particulièrement bien décrit par un profil moyen à l'échelle. Cette hypothèse a été vérifiée expérimentalement à plusieurs reprises, menant à plusieurs mesures du profil de pression dit universel des amas de galaxies. L'exemple le plus notable est le profil universel de Arnaud *et al.* [94], mesurant le profil de pression moyen de 33 amas de galaxies à bas redshift ($z < 0.2$), en utilisant des observations X seulement. Le profil résultant est représenté en figure 2.6. Cette étude montre que le profil de pression des amas est remarquablement autosimilaire : les profils des différents amas normalisés par leur masse (courbes

⁹On note qu'il est également possible de s'affranchir de cette connaissance en utilisant un filtre adapté modélisant directement le profil de paramètre de Compton y , par exemple avec un simple modèle β (approche utilisée par exemple dans les relevés réalisés à partir des données SPT, voir par exemple [4]).

¹⁰Le «lobe» d'un instrument est le terme employé en radioastronomie pour désigner sa fonction d'étalement du point, ou PSF pour *point spread function*.

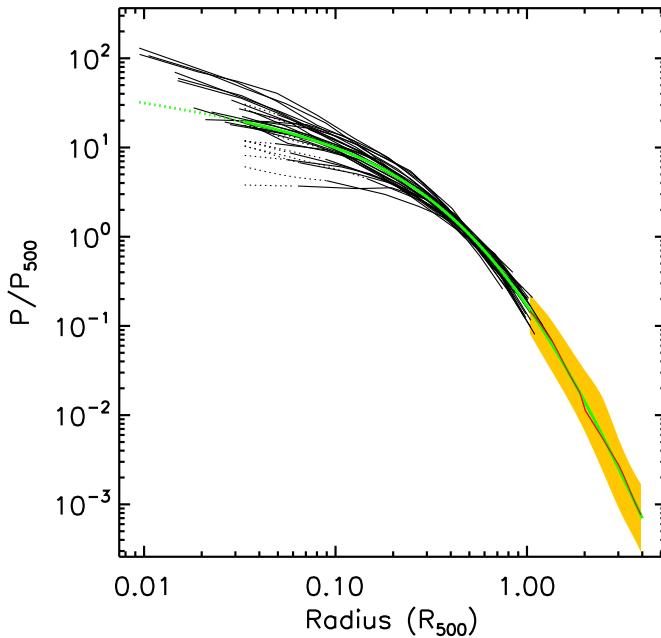


Figure 2.6 : Profils de pression des amas de galaxies de l'échantillon REXCESS mesurés par *XMM-Newton* (noir) et profil de pression moyen de l'échantillon, dit universel (vert). Figure extraite de [94].

noires en 2.6) montrent très peu de dispersion. Cela atteste du fait que les amas se comportent bien comme des répliques à l'échelle les uns des autres.

2.3 Cosmologie par comptage d'amas de galaxies

Le lien étroit entre la formation et l'évolution des amas de galaxies et la cosmologie fait de ces derniers d'excellents outils pour la mesure des propriétés de l'Univers (voir par exemple [95] pour une revue). Ce lien peut être exploité de plusieurs façons complémentaires, du diagramme de Hubble des amas (voir par exemple [96]) au spectre de puissance de l'effet SZ (voir par exemple [97]), en passant par la fraction de gaz dans les amas (voir par exemple [98]). Dans cette thèse, nous nous intéresserons au comptage d'amas de galaxies par intervalles de masse et de redshift. Cette section présente le principe des contraintes cosmologiques par comptage d'amas.

2.3.1 Fonction de masse et comptage d'amas

Comme décrit dans la section 1.3.3, la fonction de masse dN/dM quantifie le nombre de halos par unité de masse dans l'Univers (équation 1.32). Afin de pouvoir l'utiliser pour prédire un nombre d'amas visibles par un observateur, il est nécessaire de tenir compte du fait que seule une portion limitée de l'Univers lui est accessible. Deux raisons motivent ce constat. Premièrement, l'angle solide Ω couvert par des observations doit être pris en compte, puisque les amas situés en dehors de cette région ne pourront pas être détectés. Deuxièmement, il est nécessaire de calculer le volume effectivement accessible à un redshift donné. Étant donné la vitesse de déplacement finie de la lumière, celle qui nous parvient depuis un amas a été émise à une époque ancienne, à un redshift d'autant plus grand que l'amas est distant. Le volume accessible dans une tranche de redshifts $z_1 < z < z_2$ est donc celui

d'une coquille sphérique, de volume

$$V_c = \frac{4\pi}{3} [\chi^3(z_2) - \chi^3(z_1)], \quad (2.22)$$

où $\chi(z)$ est la distance comobile entre l'observateur et un point de redshift z , qui dépend des paramètres cosmologiques.

On définit alors l'abondance d'amas par unités de masse, de redshift et d'angle solide potentiellement accessibles par un observateur comme :

$$\frac{dN}{dM dz d\Omega} = \frac{dN}{dM} \frac{dV_c}{dz d\Omega}. \quad (2.23)$$

Enfin, la complétude du relevé traduit la probabilité de détecter un amas de galaxies à une masse M , un redshift z et une position \vec{r} donnée. Cette probabilité est nommée *fonction de sélection en masse* du relevé, $\hat{\chi}(z, M; \vec{r})$, et dépend à la fois de l'éventuelle anisotropie de la couverture du ciel réalisée, mais également de l'observable des amas considérée et de l'algorithme utilisé pour leur détection, comme discuté dans la section 2.3.2.

Le nombre d'amas potentiellement observables entre des redshifts z_1 et z_2 , entre des masses M_1 et M_2 , pour un relevé de complétude $\hat{\chi}$ couvrant un angle solide Ω , est alors donné par :

$$n(\vec{\vartheta}) = \int_{z_1}^{z_2} dz \int_{\Omega} d\Omega' \int_{M_1}^{M_2} dM \hat{\chi}(z, M; \Omega') \frac{dN}{dM} \frac{dV_c}{dz d\Omega'}. \quad (2.24)$$

Celui-ci dépend des paramètres cosmologiques $\vec{\vartheta}$ au travers de la fonction de masse dN/dM et du volume comobile V_c . Ainsi, un catalogue d'amas de masses et redshifts connus peut être organisé en différentes classes i de masse et redshift, et le nombre d'amas dans chaque classe N_i peut être comparé à celui attendu pour un jeu de paramètres cosmologiques $\vec{\vartheta}$. Les nombres d'amas dans chaque classe représentant un comptage discret, la comparaison se fait par le biais d'une fonction de vraisemblance poissonienne :

$$\ln \mathcal{L} = \sum_i \ln P(N_i | \vec{\vartheta}) = \sum_i [N_i \ln n_i(\vec{\vartheta}) - n_i(\vec{\vartheta}) - \ln N_i!]. \quad (2.25)$$

Cette fonction de vraisemblance peut être utilisée pour chercher le jeu de paramètres cosmologiques générant les comptages d'amas décrivant le mieux le jeu de données observé, ou encore injecté dans une procédure Monte Carlo à Chaînes de Markov (MCMC) pour trouver les distributions de probabilité permises pour les paramètres cosmologiques. De telles analyses seront détaillées dans le cadre de l'analyse des propriétés thermodynamiques du milieu intra-amas, au chapitre 5. L'exploitation cosmologique d'un relevé d'amas de galaxies requiert donc l'obtention d'un catalogue d'amas dont la masse et le redshift sont connus, qui peut être comparé à un modèle dépendant des paramètres cosmologiques à travers l'équation (2.25).

Un exemple d'analyse par comptage d'amas et de ses résultats est présenté en figure 2.7, illustrant la distribution d'amas en redshift dans l'échantillon SPT-SZ 2500 deg² [3] et les résultats cosmologiques issus de l'analyse [99]. On observe sur le panneau droit l'histogramme du nombre d'amas détectés par classes de redshift (en bleu) dans le relevé. La région verte montre les intervalles de confiance de l'ajustement de cet histogramme par un modèle de comptage d'amas (équation 2.24), en utilisant une fonction de vraisemblance similaire à l'équation (2.25). Les paramètres ajustés sont (entre autres) les paramètres cosmologiques Ω_m et σ_8 d'un modèle cosmologique standard Λ CDM. Le panneau droit montre les intervalles de confiance obtenus sur ces deux paramètres en bleu. Des études similaires ont

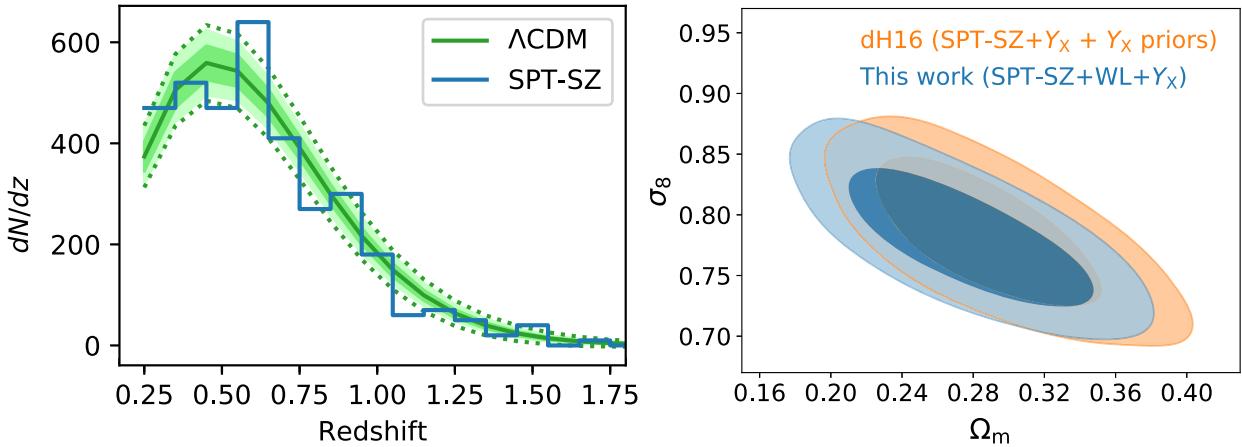


Figure 2.7 : **Gauche** : Comptage d'amas par intervalles de redshift dans le relevé SPT-SZ 2500 deg² (bleu) et résultats de l'ajustement du modèle de l'équation (2.24) (vert). **Droite** : Intervalles de confiance à 1 σ et 2 σ obtenus sur les paramètres (Ω_m , σ_8) par l'ajustement représenté dans le panneau gauche (bleu). L'ellipse orange montre les résultats de l'étude antérieure de [101] pour comparaison. Figures adaptées de [99].

pu être menées à partir des relevés réalisés par les instruments *Planck* [100] et *ACT* [2].

2.3.2 Relation d'échelle masse-observable

Nous avons vu dans la section précédente que les analyses cosmologiques nécessitaient l'obtention d'un catalogue d'amas de galaxies et la connaissance de leurs masses. Il s'agit d'une contrainte forte : en pratique, peu de relevés d'amas sont en mesure d'accéder aux valeurs des masses individuelles des amas qu'ils détectent. Il est donc nécessaire pour chaque relevé de quantifier le lien entre la masse des amas et l'une des observables directes du catalogue. Ce lien est nommé relation d'échelle masse-observable, et sa connaissance est cruciale afin de pouvoir mener des analyses cosmologiques précises et justes. Le chapitre 7 sera entièrement consacré à l'étude de cette relation d'échelle dans le cas de relevés SZ. Nous nous restreindrons donc ici aux éléments essentiels pour comprendre la nécessité de telles relations pour la cosmologie avec des amas.

Les relations d'échelle fournissent donc une estimation de la masse d'un amas à partir d'une grandeur observable d'un relevé. L'observable considérée doit donc être étroitement liée à la masse pour des contraintes précises : on qualifie alors l'observable de *proxy* de masse. Parmi les observables utilisées, on trouve la significativité de la détection des amas, c'est-à-dire la probabilité qu'une détection corresponde à un amas. Cette grandeur est un résultat direct de beaucoup d'algorithmes de détection d'amas dans toutes les longueurs d'onde, et est fréquemment utilisée comme *proxy* de masse dans les analyses cosmologiques¹¹. Plusieurs autres observables peuvent être considérées selon la longueur d'onde considérée. En X, la luminosité et la température du milieu intra-amas sont souvent utilisées (voir par exemple [102]). En optique, comme évoqué dans la section 2.2.1, la richesse des amas, définie par le nombre de galaxies membres, est fréquemment utilisée (voir par exemple [103]). En SZ, le paramètre de Compton intégré Y_Δ , quantifiant le contenu en énergie thermique dans le milieu intra-amas (cf équation 2.14), est fréquemment employé, pour des valeurs de contraste Δ dépendant de la couverture angulaire de l'instrument utilisé.

L'étalonnage de la relation d'échelle est effectué en étudiant le lien entre masse des amas et ob-

¹¹Par exemple dans les analyses des données SPT en SZ [99] utilisées à titre d'exemple dans la section 2.3.

servables pour un échantillon d'amas de masses connues. Un tel échantillon peut être dressé de deux façons. Dans le cas où un relevé offre une mesure directe de masse pour certains amas, la relation d'échelle peut être calculée à partir de ceux-ci : on parle alors d'auto-étalonnage. Par exemple, dans un relevé optique (cf. section 2.2.1), une fraction d'amas peut avoir des masses mesurées par lentillage gravitationnel ; la relation entre leur richesse et cette masse est alors calculée, et appliquée à tous les amas dont les galaxies d'arrière-plan ne permettent pas une mesure de masse (voir par exemple [104]). De même, les amas émettant fortement en X pourront permettre une mesure de masse par combinaison de leurs profils de densité et de température obtenu par spectroscopie (cf. section 2.2.2). Dans le cas où un relevé ne permet pas de mesure directe, une mesure de masse peut être obtenue par combinaison avec des données issues d'autres relevés, ou bien par un suivi dédié d'un échantillon d'amas. Les bénéfices apportés par de tels suivis seront discutés en 2.4.3. Dans les deux cas, il est important que l'échantillon d'amas dont les masses sont connues soit représentatif de la population sous-jacente, afin d'éviter de biaiser les analyses par l'étude d'une population d'amas aux propriétés particulières.

Les relations d'échelle sont le plus souvent modélisées comme une relation en loi de puissance entre l'observable et la masse. Alors, l'étude de la relation entre les logarithmes de l'observable et de la masse devient une relation affine :

$$\log Y = \alpha_{Y|M} + \beta_{Y|M} \log M. \quad (2.26)$$

Dans une approche probabiliste, la relation d'échelle peut être interprétée comme la probabilité pour un amas de masse M d'être détecté avec une valeur d'observable Y . Cette approche permet de tenir compte de la dispersion intrinsèque autour de la relation d'échelle : en effet, les différents phénomènes physiques ayant lieu au sein d'un amas peuvent affecter les observables sans modifier la masse. En modélisant cette dispersion comme gaussienne, la relation d'échelle peut s'écrire :

$$P(\log Y | \log M) = \mathcal{N}(\alpha_{Y|M} + \beta_{Y|M} \log M, \sigma_{\text{int}}^2) \quad (2.27)$$

où $\mathcal{N}(\mu, \nu)$ est la distribution de probabilité d'une loi normale de moyenne μ et de variance ν , et σ_{int} est la dispersion intrinsèque de la relation.

L'ajustement des relations d'échelle est complexe, et peut être faire appel à un grand nombre de méthodes selon les objectifs (voir par exemple la revue de [105]). Le chapitre 7 de cette thèse sera consacré à de tels ajustements dans le cadre d'un modèle bayésien hiérarchique. Une fois connue, la relation d'échelle peut être injectée dans l'équation (2.24) pour exprimer cette dernière en termes d'observables plutôt que de masse. Le nombre d'amas potentiellement observables entre des redshifts z_1 et z_2 , entre des valeurs d'observable Y_1 et Y_2 est alors donné par :

$$n(\vec{\vartheta}, \vec{\vartheta}') = \int_{z_1}^{z_2} dz \int_{\Omega} d\Omega' \int dM \int_{Y_1}^{Y_2} dY \chi(z, Y; \Omega') \frac{dN}{dM} \frac{dV_c}{dz d\Omega'}. \quad (2.28)$$

où $\chi(z, Y; \Omega)$ est la *fonction de sélection en observable* du relevé, liée à la fonction de sélection en masse par :

$$\hat{\chi}(z, M; \Omega) = \int P(Y|M) \chi(z, Y; \Omega) dY. \quad (2.29)$$

Le nombre d'amas attendu dépend donc à la fois des paramètres cosmologiques $\vec{\vartheta}$ et des paramètres de la relation d'échelle, $\vec{\vartheta}'$. La connaissance des valeurs des paramètres $\vec{\vartheta}' = (\alpha_{Y|M}, \beta_{Y|M}, \sigma_{\text{int}})$ décrivant le mieux les données est alors insuffisante, puisqu'elle ne permet pas de propager l'incertitude sur ces

grandeur au calcul des paramètres cosmologiques. Il est donc nécessaire de connaître les distributions de probabilité des paramètres de la relation d'échelle. Ces distributions peuvent être obtenues par l'ajustement de la relation par MCMC, qui sera discuté dans le chapitre 7, et utilisées comme *a priori* sur les paramètres dans l'ajustement simultané des paramètres cosmologiques et de ceux de la relation d'échelle. De telles études permettent de tenir compte des effets systématiques associés à l'étalonnage de masse des amas de galaxies dans les analyses cosmologiques (par exemple en SZ [99] ou en optique [103]).

2.4 État de l'art et perspectives

2.4.1 Relevés d'amas et résultats cosmologiques actuels

Le comptage d'amas de galaxies s'est imposé comme l'une des sondes cosmologiques majeures des dernières années grâce à l'augmentation du nombre d'amas détectés dans les relevés réalisés aux différentes longueurs d'onde discutées en 2.2.

En optique, les relevés du Sloan Digital Sky Survey (SDSS, [54, 106]), et plus récemment des Dark Energy Survey (DES, [103]) et Kilo-Degree Survey (KiDS, [107]), ont permis de construire des catalogues de plusieurs milliers d'amas de galaxies, jusqu'à des redshifts d'environ 0.6. Ces relevés ont utilisé les mesures de lentillage faible pour l'auto-étalonnage des masses des amas de ces catalogues (voir section 2.3.2).

Les catalogues d'amas de galaxies dressés grâce aux différents relevés réalisés par le satellite *ROSAT* (voir par exemple RASS [108] ou REFLEX [109]) ont laissé un héritage conséquent pour l'étude des propriétés des amas de galaxies en X, en particulier de leurs propriétés thermodynamiques et des relations d'échelles liant brillance de surface X et masse d'amas (voir par exemple [94, 110, 111]). Récemment, les relevés d'amas en X ont consisté en plusieurs relevés profonds concentrés sur de petites régions du ciel. En particulier, le relevé XXL a permis de construire un catalogue de 356 amas dans une région de 50 deg² jusqu'à des redshifts d'environ 1 [112].

Les différents relevés du CMB réalisés au cours de la dernière décennie ont permis la construction de grands catalogues d'amas grâce à l'effet SZ, notamment avec le satellite *Planck* [113], le South Pole Telescope (SPT, [4]) et l'Atacama Cosmology Telescope (ACT, [2]). Ce dernier constitue l'un des relevés d'amas de galaxies les plus grands et profonds existant à la date de l'écriture de cette thèse, avec plus de 4000 amas jusqu'à des redshifts d'environ 1.9 [2].

La distribution en masse et redshift des catalogues d'amas obtenus par l'effet SZ est représentée en figure 2.8. Celle-ci suit la forme du produit de la fonction de masse et du volume comobile, qui fournit une prédiction du nombre d'amas de masse et redshift donné accessible par un observateur¹². On voit que la plupart des amas détectés par ces relevés se situent à des redshifts inférieurs à 1.5, et à des masses supérieures à $2 \times 10^{14} M_{\odot}$. On y observe également les couvertures en masse et redshift de chacun des relevés. Le relevé réalisé avec le satellite *Planck* est caractérisé par une moins grande résolution angulaire que ses contreparties au sol, due à la difficulté technique liée à l'envoi d'un télescope aussi large que le SPT (10 m) ou l'ACT (6 m) dans l'espace. Par conséquent, il est sensible à de plus grandes échelles angulaires, ce qui le rend plus complet à bas redshift et haute masse, mais moins à haut redshift et faible masse. La profondeur du relevé influe également sur la possibilité de détecter des amas. Ainsi, les relevés ACT et SPT, plus profonds que le relevé *Planck*, sont plus complets pour les amas peu

¹²Comme expliqué dans la section 2.3 (page 40) (équation 2.23).

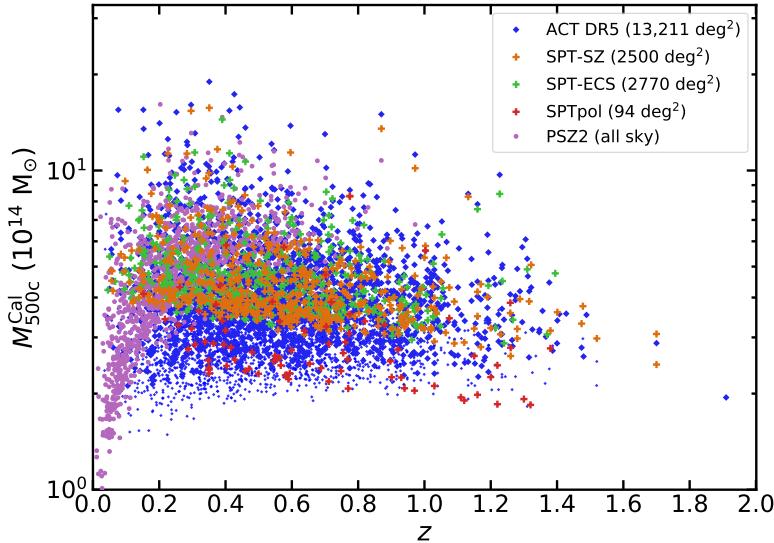


Figure 2.8 : Distribution dans le plan masse-redshift des amas détectés par les relevés ACT (bleu, [2]), SPT (orange [3], vert [4] et rouge [114]) et *Planck* (violet, [113]). Figure extraite de [2].

lumineux (de basse masse) que ce dernier. Enfin, le nombre d’amas détectés par un relevé dépend évidemment de sa couverture angulaire. Le relevé *Planck* couvre tout le ciel ($\sim 41\,000 \text{ deg}^2$), alors que les relevés ACT et SPT ne couvrent respectivement qu’un tiers et un huitième de cette surface. En conclusion, le catalogue d’amas issu du relevé *Planck* fournit une vision très complète de la distribution d’amas massifs et proches existant dans l’Univers observable, alors que les relevés ACT et SPT offrent une meilleure couverture de la distribution d’amas distants et peu massifs sur une portion plus restreinte du ciel.

Ces catalogues ont pu être utilisés pour contraindre les paramètres cosmologiques grâce à la distribution des amas en masse et en redshift. La figure 2.9 montre les contraintes apportées par chacun de ces relevés sur la combinaison de paramètres $S_8 \equiv \sigma_8 \sqrt{\Omega_m / 0.3}$, discutée en section 1.3.3 (page 22). Les mesures de ce paramètre par des analyses des anisotropies primaires du CMB et de la distribution des galaxies sont également présentées à titre de comparaison. Si les résultats apparaissent globalement en accord, à l’exception du comptage d’amas dans le Dark Energy Survey [103], les sondes basées sur l’Univers récent semblent favoriser des valeurs plus basses de S_8 que le CMB.

Ce constat, en parallèle avec la tension entre le CMB et les autres sondes dans les mesures du taux d’expansion de l’Univers (voir par exemple [38-40]), souligne la nécessité de l’étude des effets systématiques affectant l’inférence des paramètres cosmologiques. En effet, si une différence dans la mesure des paramètres cosmologiques par des sondes différentes peut être révélatrice de nouvelle physique (voir par exemple inhomogénéité de l’Univers [115], modèles d’énergie sombre interactive [116], neutrinos massifs [117] ou gravité modifiée [118]), elle peut également traduire une incompréhension des diverses incertitudes entrant en jeu lors des analyses. Il est donc crucial d’étudier les possibles effets systématiques affectant les analyses cosmologiques afin de pouvoir évaluer la significativité des tensions et la probabilité qu’elles soient dues à de nouveaux phénomènes.

2.4.2 Relevés futurs

Les décennies à venir promettent l’essor de la cosmologie avec des amas de galaxies, avec la réalisation d’un grand nombre de relevés du ciel, dans les longueurs d’onde de détection des amas, et avec

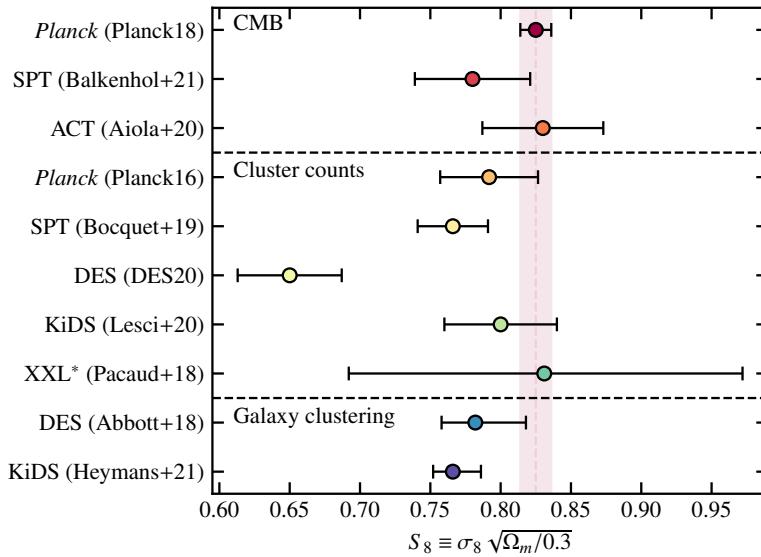


Figure 2.9 : Contrainte sur la combinaison de paramètres cosmologiques S_8 obtenues par différentes mesures du CMB (haut), de comptage d'amas (milieu), et de distribution de galaxies (bas). De haut en bas, les résultats sont issus de [13][119][120][100][99][103][107][102][121][122].

* : incertitudes estimées graphiquement.

une qualité de données sans précédent.

En optique, les relevés du Vera Rubin observatory et des satellites *Euclid* et *Roman* permettront la détection de centaines de milliers d'amas, avec des résolutions et des profondeurs permettant l'auto-étalonnage en masse à l'aide de lentillage faible jusqu'à haut redshift ($z \sim 2$, voir par exemple [9, 66]). En X, le relevé eROSITA anticipe la détection d'environ dix mille amas jusqu'à $z \sim 2$ [73]. Enfin, en millimétrique, le Simons Observatory prévoit également la détection par effet SZ de $\sim 10^5$ amas, jusqu'à des redshifts $z \sim 3$, pouvant être exploité par auto-étalonnage de masse grâce au lentillage du CMB [6, 48], et sera supplanté par CMB-S4 [7].

Ainsi, les analyses cosmologiques basées sur les amas de galaxies bénéficieront très prochainement d'une grande diminution des incertitudes statistiques. Dans ce cadre, les effets systématiques déjà importants deviendront dominants, et leur étude cruciale pour l'obtention de résultats compétitifs avec les autres sondes cosmologiques. En effet, toute méconnaissance des outils nécessaires aux analyses cosmologiques, comme les relations d'échelle masse-observable, la forme de la fonction de masse, ou le profil de pression moyen des amas¹³, aura un impact sur les paramètres cosmologiques qui pourra ne plus être négligeable devant les effets statistiques (voir par exemple [36, 37, 97]). Il sera alors important de maîtriser ces systématiques pour pouvoir tirer pleinement parti des jeux de données à venir.

2.4.3 Suivis dédiés d'échantillons d'amas

Toute observation astronomique doit faire face à un compromis entre qualité des observations d'objets individuels et nombre d'objets détectés. Dans le cadre des amas de galaxies, ce compromis permet de distinguer deux grandes catégories d'observations. D'une part, les relevés d'amas sont basés sur l'observation homogène d'une portion du ciel, puis sur la détection des objets présents dans les observations. Ils permettent l'établissement de grands catalogues nécessaires aux analyses cosmologiques

¹³discuté au chapitre suivant

comme le comptage d'amas, mais les données qui en sont issues ne permettent en général pas l'étude détaillée des propriétés physiques des amas. Par exemple, dans le cas de l'effet SZ, le relevé *Planck* [100] a été réalisé avec une basse résolution angulaire de l'ordre de 7 arcmin. Celle-ci ne permettant pas de résoudre les amas lointains, l'étude des propriétés thermodynamiques des amas individuels avec *Planck* est limitée aux objets proches. Or, la connaissance de ces propriétés est nécessaire aux analyses cosmologiques ; aussi, une évolution de celles-ci avec le redshift, indéetectable avec les données *Planck* seules, aurait des conséquences sur les analyses cosmologiques basées sur ce catalogue (voir par exemple [97]).

L'incapacité à mener des études détaillées des propriétés des amas à partir des relevés peut être compensée par un autre type d'observations : les suivis dédiés d'amas. Ceux-ci permettent d'utiliser un instrument dédié à l'observation d'un plus petit nombre d'amas, mais fournissent des informations plus détaillées sur leurs propriétés. Dans l'exemple de l'effet SZ, des instruments à haute résolution angulaire, comme les caméras NIKA2 [11, 123] ou MUSTANG2 [124], permettent de cartographier l'effet SZ en direction d'amas lointains et d'étudier leurs propriétés thermodynamiques. L'exemple du grand programme SZ de NIKA2 [125], au sein duquel s'inscrit le travail effectué au cours de cette thèse, sera détaillé dans le chapitre suivant, en section 3.4 (page 63). Si l'échantillon suivi est représentatif de la population d'amas dans l'Univers, les propriétés observées des amas, comme la relation masse-observable, peuvent être généralisées à un relevé ne fournissant pas une information aussi détaillée. De tels échantillons permettent donc l'étalonnage des outils nécessaires à la cosmologie sur un nombre restreint d'amas, avec une grande qualité de données, permettant l'étude des effets systématiques affectant les analyses cosmologiques basées sur les relevés d'amas. Nous reviendrons sur ce point au chapitre 7, dans lequel nous présenterons l'étalonnage de la relation d'échelle masse-observable pour le grand programme SZ de NIKA2.

2.4.4 Conclusions

Dans ce chapitre, nous avons discuté des propriétés des amas de galaxies et détaillé l'utilisation de leur distribution dans l'Univers comme sonde cosmologique. Au cours des dernières années, les différences entre les mesures des paramètres cosmologiques effectuées au travers de plusieurs sondes ont mis en évidence la nécessité de la maîtrise des effets systématiques propres aux analyses de chaque sonde. Dans ce contexte, les amas de galaxies s'imposent comme une opportunité hors pair de maximiser le retour scientifique de chaque relevé du ciel. En effet, de par la nature multilongueur d'onde de ces objets, la plupart des relevés du ciel sont en mesure de détecter des amas de galaxies : des observations du CMB, pouvant détecter des amas par effet SZ, aux relevés de galaxies, pouvant identifier les amas aux surdensités dans la distribution de ces dernières. Ainsi, des relevés originellement prévus pour une sonde cosmologique peuvent utiliser les amas pour des résultats indépendants de leur sonde primaire, augmentant le pouvoir contraignant de leurs données sur les paramètres cosmologiques.

Nous avons vu que les relevés du futur permettraient une réduction drastique des incertitudes statistiques associées aux comptage d'amas. La cosmologie avec des amas de galaxies entre donc dans une ère dans laquelle l'étude des effets systématiques sera cruciale à l'exploitation de relevés d'amas. À l'heure actuelle, les effets systématiques dominants en cosmologie avec des amas proviennent pour la plupart de la méconnaissance des propriétés des amas et de leur évolution avec le redshift. Dans ce contexte, la mise en place de suivis d'échantillons d'amas a le potentiel de fournir des études détaillées des effets systématiques liés à la détection et l'exploitation cosmologique des amas.

Les amas de galaxies promettent donc de jouer un rôle majeur dans la cosmologie des années à

venir, et l'étude détaillée de leurs propriétés permettra d'augmenter leur potentiel en tant que sonde cosmologique.

CHAPITRE 3

NIKA2 et son Grand Programme SZ

3.1	Le télescope de 30 mètres de l'IRAM	50
3.1.1	Localisation et conditions météorologiques	50
3.1.2	Caractéristiques instrumentales	51
3.2	La caméra NIKA2	52
3.2.1	Détecteurs à inductance cinétique	52
3.2.2	Éléments clés de la caméra NIKA2	54
3.3	Observations avec NIKA2	56
3.3.1	Balayage du ciel et données ordonnées en temps	56
3.3.2	Reconstruction du plan focal	57
3.3.3	Pointage et focalisation	57
3.3.4	Performances instrumentales de NIKA2	59
3.4	Le grand programme SZ de NIKA2	63
3.4.1	La caméra NIKA2 : un instrument idéal pour la mesure de l'effet SZ	63
3.4.2	Échantillon du grand programme SZ	65
3.4.3	Objectifs scientifiques du grand programme SZ	67
3.5	La base de données du grand programme SZ de NIKA2	72
3.5.1	Contexte	72
3.5.2	Contenu et interface de la base de données	73
3.5.3	Développements futurs	74

Nous avons vu au chapitre précédent que les amas de galaxies pouvaient être utilisés comme une sonde cosmologique. Nous avons également présenté les outils nécessaires pour pouvoir mener de telles études, et l'intérêt des suivis dédiés d'échantillons d'amas dans l'étalonnage de ces outils.

C'est dans ce contexte que s'inscrit le grand programme SZ de NIKA2, au sein duquel cette thèse s'est déroulée. NIKA2 est une caméra permettant d'observer le ciel dans le domaine millimétrique avec une résolution angulaire bien meilleure que celle des instruments utilisés pour construire des

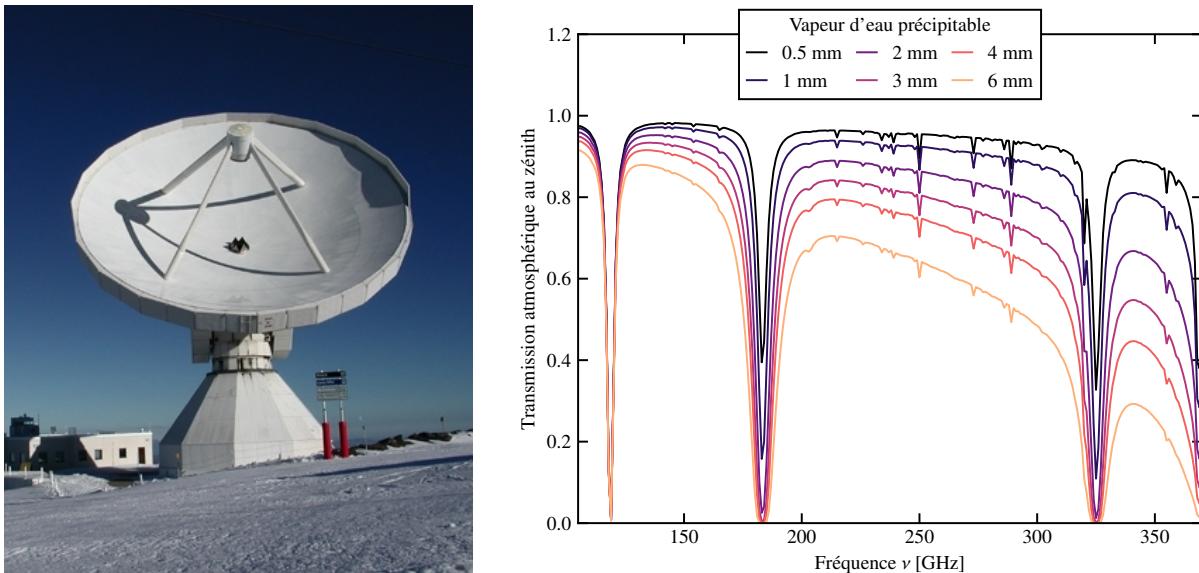


Figure 3.1 : **Gauche** : Photo du télescope de 30 mètres de l'IRAM. Crédit : IRAM. **Droite** : spectre de transmission atmosphérique en fonction de la quantité de vapeur d'eau précipitable, d'après le modèle de Pardo *et al.* [127].

catalogues d'amas. L'instrument a été construit à Grenoble par la collaboration NIKA2, et choisie par l'Institut de Radioastronomie Millimétrique (IRAM) pour être un instrument résident de son télescope de 30 mètres. En retour de la construction de NIKA2, la collaboration s'est vu attribuer un total de 1300 heures d'observations avec l'instrument. Le grand programme SZ est l'un des programmes bénéficiant de ce temps garanti.

Dans ce chapitre, nous décrivons la caméra NIKA2 et le grand programme SZ. Nous décrirons tout d'abord quelques éléments caractéristiques du télescope de 30 mètres de l'IRAM et de la caméra NIKA2. Nous détaillerons ensuite l'utilisation de l'instrument, ses caractéristiques, puis le grand programme SZ, de son utilisation des capacités de NIKA2 à ses objectifs scientifiques. La dernière section sera consacrée à la base de données du grand programme SZ, dont la création fait partie de ce travail de thèse.

3.1 Le télescope de 30 mètres de l'IRAM

3.1.1 Localisation et conditions météorologiques

Le télescope de 30 mètres de l'Institut de Radioastronomie Millimétrique (IRAM, figure 3.1) est installé dans la chaîne de montagnes de la Sierra Nevada, proche de Grenade, dans le sud de l'Espagne [126]. Il est situé à une altitude de près de 2900 mètres au dessus du niveau de la mer, proche du sommet Pico Veleta ; ses coordonnées sont $03^{\circ}23'58.1''\text{W}$, $37^{\circ}04'05.6''\text{N}$.

Comme nous le verrons au chapitre suivant, l'une des sources de bruit principales dans les observations millimétriques est l'atmosphère. Le bruit atmosphérique est principalement dû à son humidité. Le site de Pico Veleta a donc été choisi pour son altitude, diminuant l'épaisseur de l'atmosphère traversée par les photons avant d'atteindre le télescope, mais également pour ses conditions météorologiques, en particulier son faible taux d'humidité. Le spectre de transmission de l'atmosphère, calculé à l'aide du modèle de Pardo *et al.* [127], est représenté pour différentes valeurs de vapeur d'eau précipitable, correspondant à la quantité d'eau présente dans l'atmosphère. On peut voir que pour des fréquences

entre 100 et 300 GHz, la transmission est modulée par un continuum, dont la transmission diminue avec l'humidité, et par la présence de raies d'absorption (du dioxygène à ~ 118 GHz et de l'eau à ~ 185 et 325 GHz). Ces raies délimitent les gammes de longueurs d'onde observables depuis le sol, nommées fenêtres d'observation, centrées autour de 90, 150, 250, et 350 GHz.

En moyenne, les valeurs typiques de vapeur d'eau précipitable au télescope de 30 mètres sont de 7 mm en été, et de 3 mm en hiver. En effet, la température de l'atmosphère est plus élevée l'été, augmentant la valeur de pression de vapeur saturante et donc la quantité d'eau qu'elle peut contenir. Le télescope suit l'évolution de la quantité d'eau en temps réel en mesurant l'opacité atmosphérique au zénith τ_ν , grâce à un instrument résident dit taumètre, définie de façon à ce que la transmission de l'atmosphère T_ν à une fréquence ν soit

$$T_\nu = \exp(-\tau_\nu / \sin(el)), \quad (3.1)$$

où el est l'élévation de l'objet observé.

L'opacité usuellement mesurée est τ_{225} , à une fréquence de 225 GHz, et vaut en moyenne 0.2 en hiver et 0.5 en été au télescope de 30 mètres.

La partie du ciel observable par le télescope est limitée au nord de la sphère céleste. En effet, les objets de déclinaison inférieure à -20° en coordonnées équatoriales n'y atteignent jamais des élévations supérieures à 30° . À de telles élévations, la couche d'atmosphère traversée par les rayons incidents est épaisse, et le poids du télescope commence à nuire à sa précision de pointage et affecter la forme de son lobe. Les déclinaisons inférieures à -20° sont donc considérées comme inaccessibles pour le télescope. Cela empêche l'observation d'amas découverts dans certaines régions des relevés réalisés dans l'hémisphère sud. Comme le montre la figure 3.2, une conséquence notable est l'impossibilité d'utiliser le télescope de 30 mètres pour le suivi d'amas des catalogues SPT [3, 4], de même qu'environ la moitié du catalogue ACT [2]. C'est la raison pour laquelle les amas de galaxies suivis par le grand programme SZ de NIKA2¹ sont principalement issus du catalogue *Planck*, couvrant tout le ciel, et de quelques uns des amas les plus au nord du catalogue ACT.

3.1.2 Caractéristiques instrumentales

Le télescope est constitué d'un miroir primaire paraboloïde de 30 mètres de diamètre, composé de panneaux d'aluminium et de polyuréthane [126]. Afin de minimiser les variations de température induites par le lever et le coucher du soleil, qui entraînent une déformation du miroir primaire, ce dernier est activement thermalisé [128] et peint de manière à réfléchir le rayonnement infrarouge. Le miroir secondaire de 2 mètres de diamètre est maintenu à une distance de 10.5 mètres du primaire, égale à la distance focale de ce dernier. L'ensemble est couplé à une cabine Nasmyth et installé sur une monture alt-azimutale, pour une masse totale d'environ 800 tonnes. Le télescope est présenté en figure 3.1.

Les rayons lumineux arrivent donc sur le miroir primaire du télescope, sont réfléchis par celui-ci vers le miroir secondaire, qui à son tour les redirige vers le vertex au centre du miroir primaire par lequel les faisceaux pénètrent dans la cabine. La réponse angulaire du télescope, nommée *lobe instrumental*, décrit la distribution du signal optique attendu à la sortie de ce vertex. Dans le cas d'un signal monochromatique issu d'une source ponctuelle, cette réponse angulaire est donnée par la tâche d'Airy du miroir primaire à travers l'ouverture du vertex. La largeur de celle-ci, déterminée par la largeur

¹Voir section 3.4 (page 63)

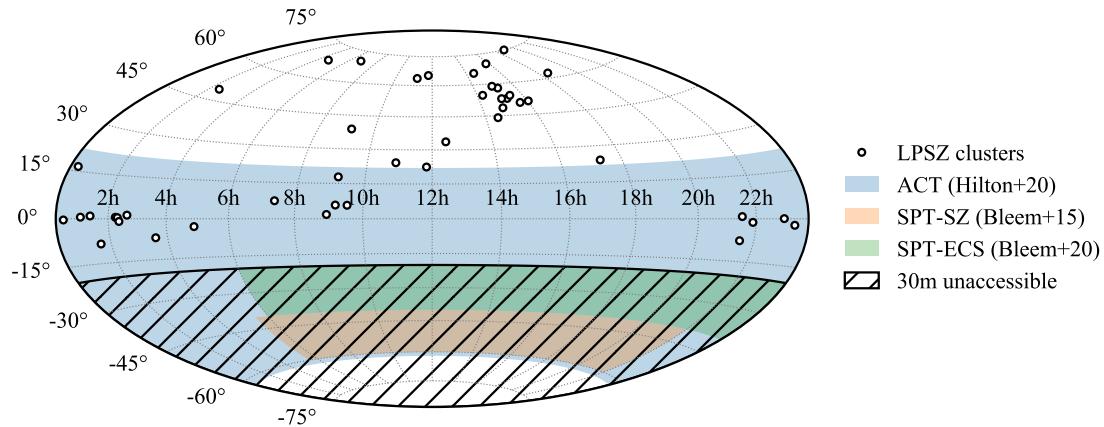


Figure 3.2 : Couvertures spatiales des relevés ACT (bleu, [2]) et SPT (orange et vert [3, 4]) en coordonnées équatoriales. Les axes horizontaux et verticaux représentent respectivement l'ascension droite et la déclinaison. La région hachurée marque les déclinaisons inaccessibles par le télescope de 30 mètres de l'IRAM. Les points représentent les coordonnées des amas du grand programme SZ de NIKA2 (voir section 3.4 page 63).

à mi-hauteur du pic principal, vaut $\theta = \arcsin(1.029 \times \lambda/D) = 14.1$ et 8.2 arcsec pour des rayonnements monochromatiques de fréquences 150 et 260 GHz respectivement. En pratique, le rayonnement incident est polychromatique, et ses multiples réflexions induisent un élargissement du lobe. C'est la raison pour laquelle les différents instruments récepteurs dans la cabine du télescope ont des résolutions angulaires effectives moins bonnes que ces valeurs.

3.2 La caméra NIKA2

La caméra NIKA2 (*New IRAM KID arrays 2*) constitue l'instrument principal utilisé au cours de cette thèse. Elle a été conçue par la collaboration internationale NIKA2, centrée à Grenoble, dans l'objectif de répondre à un appel d'offre de l'IRAM pour un nouvel instrument résident au télescope de 30 mètres. Dans cette section, nous présentons brièvement l'instrument et son fonctionnement. Plus de détails peuvent être trouvés dans [11, 123].

3.2.1 DéTECTEURS à inDUCTANCE CINÉTIQUE

Les détecteurs utilisés par NIKA2 sont des détecteurs à inductance cinétique (KID, pour *kinetic inductance detectors*, [129, 130]). Leur utilisation est l'une des spécificités de NIKA2, puisque la plupart des instruments millimétriques utilisent des détecteurs bolométriques.

Les KID sont des détecteurs supraconducteurs. Ils sont constitués d'un matériau caractérisé par une température critique T_c en dessous de laquelle les électrons sont séparés en deux populations : d'une part, des électrons appariés en *paires de Cooper*, et d'autre part des électrons libres, nommés *quasi-particules*. La densité de paires de Cooper n_s est alors donnée en fonction de la température T par [131] :

$$\frac{n_s}{n} = 1 - \left(\frac{T}{T_c} \right)^4, \quad (3.2)$$

où n est la densité totale des porteurs de charges, quasi-particules et paires de Cooper.

Les paires de Cooper forment des condensats de Bose-Einstein, pouvant être brisée par une énergie

dite de *gap*, notée 2Δ , et pouvant être calculée dans le cadre de la théorie Bardeen-Cooper-Schrieffer [132, 133] comme :

$$2\Delta \simeq 3.53 k_B T_c. \quad (3.3)$$

D'après l'équation (3.2), pour des températures $T \ll T_c$, la majorité des charges est donc portée par des paires de Cooper ($n_s \simeq n$). Alors, des photons incidents d'énergie supérieure au *gap* 2Δ pourront briser ces paires et changer la distribution des porteurs de charges entre paires de Cooper et quasi-particules.

Lorsqu'un supraconducteur est traversé par un courant électrique, la conduction peut être assurée par chacun des porteurs de charges. Dans le cas d'un courant alternatif cependant, les paires de Cooper sont sujettes à une inertie, qui freine leur mouvement, et s'oppose donc à la conduction. Ce phénomène crée une inductance qualifiée de cinétique L_k dans le matériau, dont l'impédance s'écrit alors $Z = R_{qp} + i\omega L_k$, où R_{qp} est la résistance liée aux quasi-particules, et ω la pulsation du courant alternatif. Il apparaît naturellement qu'une variation de densité de paires de Cooper, par brisure de certaines d'entre elles par des photons incidents, créera une variation d'inductance cinétique δL_k . Celle-ci est inversement proportionnelle à la variation de densité de paires de Cooper, et donc proportionnelle à la variation de puissance optique incidente (voir par exemple [129] pour une démonstration).

Le fonctionnement des KID est basé sur la mesure de variation de fréquence de résonance induite par un changement de puissance optique reçue. Les détecteurs de NIKA2 sont des détecteurs à inductance de type LEKID (*Lumped Element KID*, [130]). Chaque KID est un circuit résonant de type RLC, constitué d'une partie capacitive C , d'un méandre inductif servant d'absorbeur de flux, d'inductance géométrique L_g s'ajoutant à l'inductance cinétique L_k , et d'une résistance R . La fréquence de résonance d'un KID est donc donnée par :

$$f = \frac{1}{2\pi\sqrt{(L_g + L_k)C}}. \quad (3.4)$$

Ainsi, une variation de puissance optique incidente sur le détecteur va induire un changement de la proportion de porteurs de charges sous forme de paires de Cooper, ce qui créera une variation de l'inductance cinétique du supraconducteur, qui aura à son tour pour conséquence une modification de la fréquence de résonance du détecteur. La dérivation de l'équation (3.4) par rapport à l'inductance cinétique L_k permet d'exprimer cette variation de fréquence de résonance par rapport au changement d'inductance cinétique :

$$\delta f = \frac{\partial f}{\partial L_k} \delta L_k = -2\pi^2 C f^3 \delta L_k. \quad (3.5)$$

Comme nous l'avons vu précédemment, le changement d'inductance cinétique est inversement proportionnel à la variation de puissance optique incidente ; ainsi,

$$\delta f \propto \delta L_k \propto -\delta P_{\text{opt}}. \quad (3.6)$$

La mesure de cette variation de fréquence de résonance permet donc de remonter à la variation de puissance optique l'ayant causée. Afin de mesurer δf , les détecteurs sont couplés à des lignes de transmission, parcourues par un signal alternatif. Celles-ci sont couplées inductivement aux méandres des KID : le signal transmis est donc affecté par la présence de chaque détecteur. La variation en phase et en amplitude du signal de la ligne de transmission due à la présence d'un KID est liée à la fréquence de résonance de ce dernier. Ainsi, la mesure de la phase et de l'amplitude du signal de sortie de la ligne de lecture permet de reconstruire la variation de fréquence de résonance des détecteurs présents

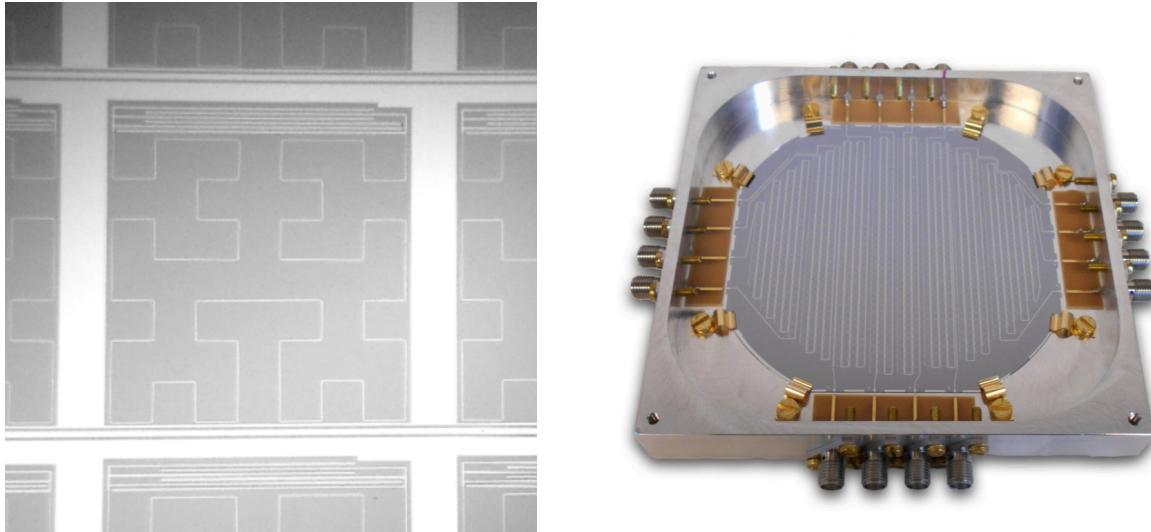


Figure 3.3 : **Gauche** : Image d'un détecteur KID de NIKA, de taille $2.3 \text{ mm} \times 2.3 \text{ mm}$. Le méandre inductif du détecteur couvre la plupart de sa surface, tandis que la partie capacitive C est située dans la partie haut du détecteur. **Droite** : Matrice de KID dans la bande à 260 GHz de NIKA2, comportant 1140 KID et 8 lignes de transmission. Figures extraites de [123].

le long de cette ligne, et *in fine* à la variation de puissance lumineuse.

Comme le montre l'équation (3.4), la fréquence de résonance d'un KID est liée à la capacité de son condenseur C . Ainsi, la fréquence de résonance de chaque détecteur peut être ajustée par contrôle des propriétés géométriques du condenseur. Cette caractéristique fait des KID des détecteurs intrinsèquement adaptés au multiplexage : un grand nombre de détecteurs peuvent avoir des fréquences de résonance différentes et être montés sur une même ligne de lecture. Par conséquent, la quantité de câbles électroniques nécessaires à la transmission des signaux d'une matrice de détecteurs est fortement diminuée par rapport aux bolomètres, nécessitant une ligne de lecture par détecteur. Au vu des contraintes cryogéniques importantes pour les détecteurs, une augmentation du nombre de détecteurs lisibles par une ligne de lecture permet l'augmentation du nombre de détecteurs sans pour autant augmenter le nombre de câbles devant sortir du cryostat. Cette qualité fait des KID des détecteurs particulièrement compétitifs.

Un des KID de la caméra NIKA et une matrice de détecteurs installée dans NIKA2 sont présentés en figure 3.3. Le panneau de gauche montre les différentes parties du détecteur : le condensateur (en haut), et le méandre inductif, occupant la majorité de la surface du détecteur et dans lequel les photons sont absorbés et peuvent briser des paires de Cooper. On voit également la ligne de transmission utilisée pour lire le détecteur qui en longe le côté inférieur. Sur le panneau droit, on voit une des matrices de 1140 KID de NIKA2. Les détecteurs individuels ne sont pas distinguables, mais on peut voir les huit lignes de transmission utilisées pour la lecture des 1140 détecteurs.

3.2.2 Éléments clés de la caméra NIKA2

Les KID de NIKA2 sont constitués d'une couche mince d'aluminium d'environ 18 nm d'épaisseur. La température critique de l'aluminium étant de $T_c = 1.19 \text{ K}$, l'énergie du *gap* vaut $2\Delta \simeq 0.36 \text{ meV}$, correspondant à des photons à une fréquence de l'ordre de 90 GHz. De tels détecteurs sont donc tout à fait appropriés pour la mesure du fond diffus cosmologique et de l'effet SZ (cf. figure 2.3).

La caméra NIKA2 est dotée de trois matrices de KID. Deux matrices (A1 et A3), chacune composée de 1140 détecteurs, sont placées derrière une lame dichroïque transmettant une gamme de fréquence

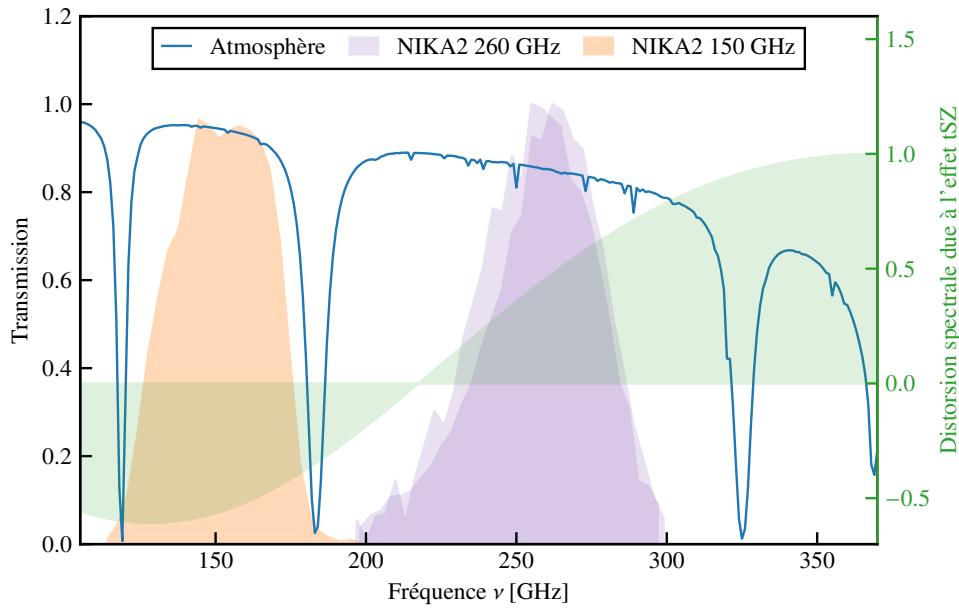


Figure 3.4 : Bandes passantes de NIKA2 à 260 GHz (violet) et 150 GHz (orange). La transmission atmosphérique calculée à partir du modèle de Pardo *et al.* [127] pour 2 mm de vapeur d'eau précipitable est également représentée en bleu. La distorsion spectrale due à l'effet tSZ est montrée en vert.

centrée à 260 GHz. La troisième (A2) est quant à elle dédiée à l'observation de photons de fréquence autour de 150 GHz. Les deux bandes passantes de NIKA2 sont donc adaptées aux fenêtres d'observation disponibles au vu de la transmission atmosphérique (cf. figure 3.1). Elles sont illustrées sur la figure 3.4. Il y apparaît que les bandes passantes de NIKA2 exploitent parfaitement les fenêtres spectrales disponibles pour des observations millimétriques depuis le sol, de même que la dépendance spectrale de l'effet Sunyaev-Zeldovich, avec une bande de chaque côté du zéro de la distorsion spectrale (cf. section 2.2.3). Ce point sera abordé plus en détail en 3.4 (page 63).

Comme nous l'avons discuté en 3.2.1, le fonctionnement des détecteurs à inductance cinétique requiert que ceux-ci soient refroidis à une température inférieure à leur température critique. Dans le cas de NIKA2, cette condition est assurée par un cryostat, maintenant les détecteurs à une température d'environ 150 mK, bien en deçà de la température critique de l'aluminium.

Un système optique dévie les rayons lumineux focalisés par le télescope (cf. section 3.1.2) vers les matrices de KID. Celui-ci est tout d'abord composé de quatre miroirs dans la cabine du télescope, dirigeant les rayons issus du miroir secondaire du télescope vers le cryostat de NIKA2. Une fois qu'ils ont pénétré le cryostat, les rayons lumineux sont réfléchis par une succession de miroirs refroidis, puis traversent un filtre dichroïque, séparant la lumière en deux. Celui-ci permet de diriger une partie des rayons incidents vers la matrice de KID à 150 GHz, et l'autre vers les matrices à 260 GHz. La présence des deux matrices à 260 GHz et d'un élément polariseur permet à NIKA2 d'être sensible à la polarisation des photons incidents à 260 GHz. Dans le travail développé au cours de cette thèse, cette capacité ne sera pas exploitée, et les données issues des deux matrices à 260 GHz sont combinées en une unique carte du ciel à cette fréquence.

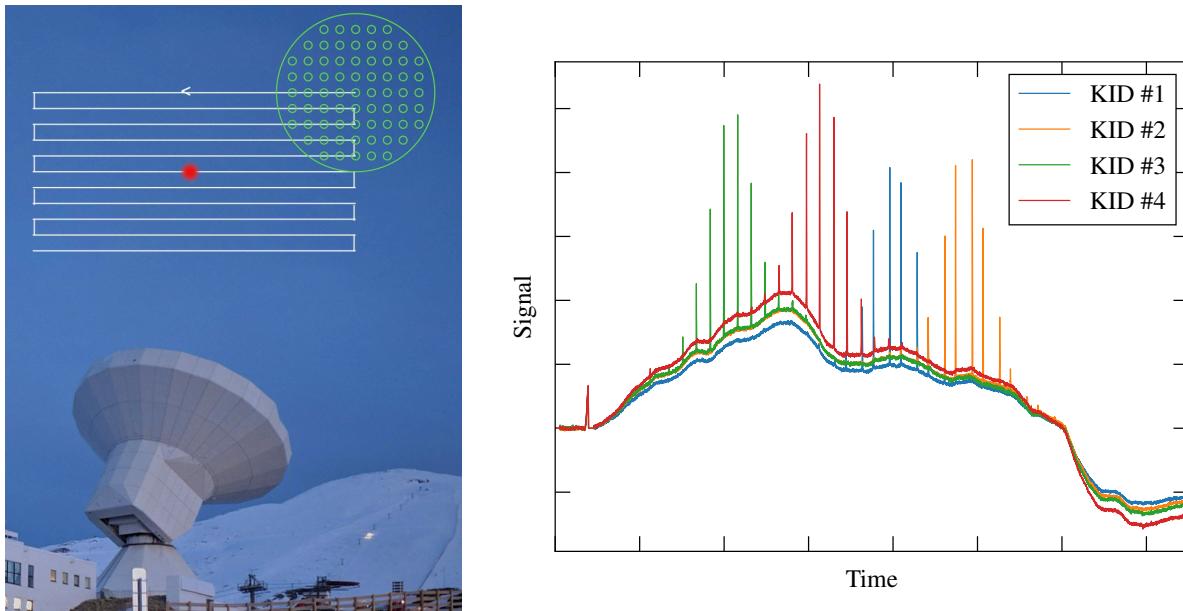


Figure 3.5 : **Gauche** : illustration du mouvement effectué par le télescope dans le ciel (en blanc) lors d'un scan OTF. La matrice de détecteurs est représentée en vert, et la source observée en rouge. **Droite** : représentation des données ordonnées en temps obtenues par scan d'Uranus pour trois KID. La détection de la source se manifeste par un pic de signal détecté.

3.3 Observations avec NIKA2

Les observations avec la caméra NIKA2 réalisées et étudiées au cours de cette thèse suivent une procédure précise, dont le but est la production de cartes fidèles du ciel. Dans cette section, nous détaillons cette procédure, ainsi que les performances de l'instrument. L'étape suivante, visant à isoler les signaux astrophysiques des différents contaminants, sera détaillée au chapitre suivant.

3.3.1 Balayage du ciel et données ordonnées en temps

Comme nous l'avons vu dans la section précédente, les KID permettent de mesurer une variation de puissance lumineuse. Ainsi, ils sont moins adaptés à la mesure de signaux stationnaires qu'à celle de signaux variant dans le temps. Pour pouvoir mesurer des signaux stationnaires, tels que les signaux astrophysiques étudiés au cours de cette thèse, il est donc nécessaire d'adopter une stratégie visant à moduler le signal d'intérêt dans le temps.

Dans le cas de NIKA2, cet objectif est accompli en balayant le ciel par des *scans*, dont le principe est le suivant. Le télescope se déplace suivant une trajectoire déterminée, de sorte à ce que la position sur le ciel d'un détecteur de référence soit connue à tout instant. Chacun des KID de NIKA2 enregistre alors la variation de la puissance lumineuse qu'il reçoit au cours du temps. Les données enregistrées sont par conséquent ordonnées en temps, et nommées TOI pour *Time Ordered Information*. Une fois le balayage terminé, une matrice de pointage est calculée, donnant la position de chaque KID sur le ciel au cours du temps à partir du pointage de référence du télescope et de la position relative de chaque KID par rapport à ce centre de référence (la mesure de celle-ci sera détaillée en 3.3.2). La matrice de pointage permet alors de projeter les TOIs de chaque KID en cartes du ciel. En plus de moduler temporellement le signal pour le rendre détectable par les KID, cette méthode permet d'exploiter la redondance du signal et les corrélations entre les différentes sources de bruit. Comme nous le verrons au chapitre suivant, cette redondance facilite la différenciation du signal et du bruit dans les données.

La trajectoire suivie par le télescope pour réaliser un scan est nommée stratégie de scan. Le choix de celle-ci est capital, puisqu'il a un grand impact sur la capacité à reconstruire le signal astrophysique (comme nous le verrons dans le chapitre suivant). Dans le cas des observations de l'effet SZ discutées au cours de cette thèse, la stratégie employée est basée sur la réalisation de scans OTF (pour *on the fly*) : une successions de *subscans* rectilignes parallèles en coordonnées équatoriales sont effectués, espacés d'un pas constant. Cette stratégie induit un mouvement du télescope en forme de serpentin, illustrée sur le panneau gauche de la figure 3.5. Les données ordonnées en temps obtenues pour un tel *scan* sur une source ponctuelle brillante (Uranus) sont représentées sur le panneau droit de la figure 3.5. Étant données les différentes positions des détecteurs dans le plan focal, ces derniers ne détectent pas le passage de la source au même moment, ce qui se manifeste par un décalage en temps du signal de la source. Les observations du grand programme SZ sont effectuées par groupes de quatre scans, à des angles de 0° , 45° , 90° , et 135° par rapport à l'axe des ascensions droites, afin de maximiser l'isotropie de la couverture du ciel. Ce point sera discuté plus en détail au chapitre suivant.

3.3.2 Reconstruction du plan focal

Le grand nombre de détecteurs présents sur une même ligne de lecture peut causer des difficultés à isoler les signaux provenant de chaque détecteur. En particulier, la lecture de la fréquence de résonance d'un détecteur ne permet pas de connaître sa position dans la matrice. De plus, la fréquence de résonance de certains détecteurs peut ne pas être trouvée, ou bien être trop proche de celle d'un détecteur voisin, auquel cas les signaux en provenance des deux KID sont incorrectement associés à un seul détecteur. Il est donc nécessaire, avant d'observer, de reconstruire le plan focal de la caméra en associant chaque fréquence de résonance à un KID, permettant de connaître sa position et d'identifier les détecteurs invalides.

Pour cela, une carte profonde d'une source brillante, par exemple une planète comme Uranus, est réalisée de façon à ce qu'une carte de la source puisse être construite pour chaque détecteur. Ces cartes, nommées *beammaps*, ont plusieurs vertus. Elles permettent d'une part, pour chaque détecteur, de connaître sa position dans la matrice par rapport à un KID de référence, ce qui permettra la reconstruction des observations du ciel détaillée par la suite. Elles permettent également l'identification des détecteurs en diaphonie, pour lesquels deux sources apparaissent sur une même carte, de même que les détecteurs hors résonance. Les gains des KID, quantifiant l'étalonnage relatif de chaque détecteur, sont également calculés à partir de ces cartes grâce au rapport entre le flux connu de la source et la variation de fréquence de résonance correspondante pour chaque détecteur. Enfin, ces cartes sont utilisées pour la mesure du lobe instrumental observé par chacun des détecteurs, discutée en 3.3.4 (page 59).

3.3.3 Pointage et focalisation

Avant d'observer des sources, il est nécessaire de s'assurer que le télescope est à même de livrer des observations fidèles du ciel. Il faut donc s'assurer que le KID de référence est correctement aligné avec l'axe optique du télescope, et que ce dernier est correctement focalisé.

Le pointage du télescope est vérifié en réalisant un mouvement d'aller-retour sur une source, tout d'abord en azimut, puis en élévation, formant une croix dans le plan du ciel. Des profils Gaussiens sont ajustés sur les TOIs d'un détecteur de référence au centre des matrices de NIKA2. La position de la source dans le ciel étant connue, il est alors possible d'associer des coordonnées célestes au détecteur de référence, et par extension aux autres KID des matrices de NIKA2. Un exemple de procédure de

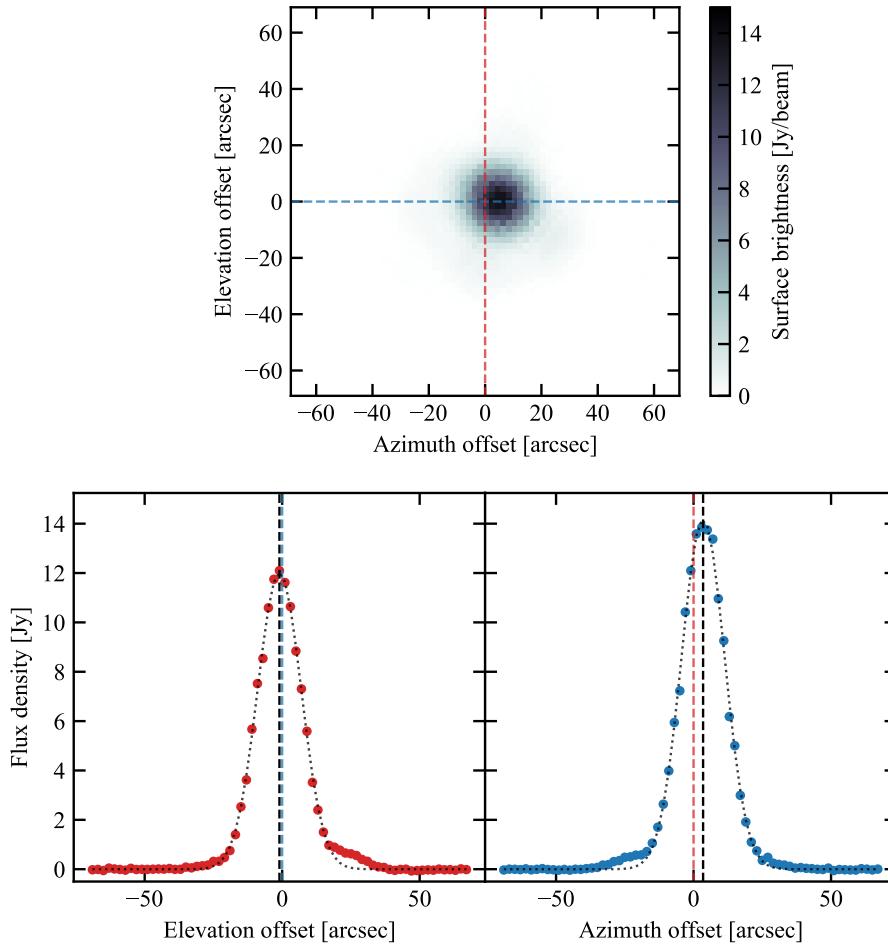


Figure 3.6 : Exemple de procédure de pointage. **Haut :** carte d’Uranus obtenue avec un scan en croix, suivant les directions des lignes bleue puis rouge, projetée en utilisant la matrice de pointage avant correction de pointage. **Bas :** profils de la planète le long de l’axe des élévations (*gauche*) et des azimuts (*droite*). L’ajustement des profils par une gaussienne (courbes pointillées noires) permet de mesurer le centre de la source dans les deux directions (lignes noires verticales). La différence entre la position de la source et celle du centre de la carte donne les corrections de pointage utilisées pour réaligner le KID de référence.

pointage est présenté en figure 3.6, illustrant la carte obtenue pour un pointage sur Uranus. L’ajustement des profils de la source le long des axes des azimuts et élévations mesurés par le détecteur de référence (bas de la figure) permet de faire correspondre les coordonnées de chacun des détecteurs dans le plan focal avec ses coordonnées célestes².

La focalisation du télescope quantifie l’accord entre la position du miroir secondaire et le point focal du miroir primaire. Lorsque le télescope est défocalisé, c’est-à-dire lorsque le miroir secondaire est éloigné du point focal du primaire, les rayons ne sont pas réfléchis correctement vers le système optique, induisant des déformations de l’image reconstruite avec NIKA2, et notamment un élargissement du lobe. Il faut alors positionner correctement le miroir secondaire. Afin de connaître la position optimale de celui-ci, un scan OTF est effectué sur une source brillante pour cinq distances différentes entre les miroirs primaire et secondaire, espacées de 0.4 mm. Les cartes obtenues à partir de ces ob-

²Notons que la procédure utilisée pour le pointage ajuste le profil de la source sur les TOI du KID de référence, alors que les profils représentés sur la figure 3.6 sont ceux de la carte déjà projetée.

servations sont créées et utilisées pour ajuster un modèle de lobe gaussien. La procédure est illustrée en figure 3.7 avec la matrice à 150 GHz de NIKA2. Cinq cartes d'Uranus sont produites (haut) et ajustées avec un modèle de lobe gaussien, permettant de mesurer la largeur du lobe et le flux de la source pour chacune des cartes. Ces résultats sont ajustés avec une loi parabolique, dont le maximum donne les valeurs de la distance primaire-secondaire maximisant le flux et minimisant la largeur du lobe. L'une ou l'autre de ces deux distances peut être choisie comme distance optimale pour la focalisation du télescope, selon la qualité de l'ajustement. On note que les résultats des deux estimateurs sont en général très proches – une différence de 0.03 mm est trouvée dans l'exemple de la figure 3.7.

Un pointage correct est important pour toutes les observations réalisées avec NIKA2, incluant les observations de l'effet SZ. Un mauvais pointage entraîne une connaissance imprécise de la position des détecteurs, et empêche donc la construction de cartographies fidèles du ciel. Une erreur systématique sur le pointage de toutes les observations d'une source entraîne une mauvaise reconstruction des coordonnées du signal dans le ciel, qui a pour conséquence l'impossibilité de comparer des cartes NIKA2 avec des données issues d'autres instruments. Comme nous le verrons dans les chapitres 5 et 6, de telles comparaisons sont au cœur des analyses du grand programme SZ traitées dans cette thèse, et sont nécessaires à l'étude des propriétés thermodynamiques des amas de galaxies et des contaminants astrophysiques. D'autre part, des imprécisions régulières (par exemple, variant d'un scan à l'autre) entraîne une dilution du signal, l'attribuant à des positions incorrectes, et donc un élargissement effectif du lobe de l'instrument. Comme nous le verrons en section 3.4 (page 63), la résolution angulaire de NIKA2 est l'un de ses principaux atouts pour la mesure de l'effet SZ. La précision du pointage pouvant être dégradée par les mouvements du télescope, celui-ci est vérifié régulièrement au cours des observations, environ une fois par heure. De même, comme le montre la figure 3.7, une mauvaise focalisation du télescope entraîne un élargissement du lobe, et donc une dégradation de la résolution angulaire. La focalisation variant principalement avec la géométrie du miroir primaire, qui se déforme avec les changements de température. C'est pour cette raison que les observations diurnes avec NIKA2 doivent éviter l'illumination du miroir primaire avec le soleil, et que la focalisation est vérifiée à intervalles réguliers (toutes les heures de jour, et toutes les deux à trois heures la nuit).

3.3.4 Performances instrumentales de NIKA2

Les caractéristiques de NIKA2 ont été mesurées par la collaboration du même nom au cours de sa phase de *commissioning*. Ces mesures sont décrites avec un grand niveau de détail par Perotto *et al.* [11], et nous n'en donneront que les résultats principaux nécessaires à la description du travail effectué dans cette thèse. La phase de *commissioning* de l'instrument a eu lieu entre Octobre 2015 et Avril 2017. Elle a permis de mesurer avec précision les caractéristiques de l'instrument, telles que son lobe et sa sensibilité. La mesure des performances de l'instrument a quant à elle eu lieu au cours de trois campagnes d'observations en Février et Octobre 2017 et Janvier 2018. Les résultats principaux de ces mesures sont consignés dans la table 3.1, et nous détaillons ici brièvement la signification des caractéristiques et la méthode de calcul.

Détecteurs valides

Le nombre de détecteurs valides est évalué à l'aide des *beammaps* décrites en 3.3.2. Sont considérés comme valides les détecteurs pour lesquels la mesure de la variation de fréquence de résonance lors du passage sur une source brillante permet de déduire une cartographie de cette source. En d'autre termes, les détecteurs invalides incluent les détecteurs bruités dont la projection des données en temps ne permet pas d'identifier une source, ainsi que les détecteurs en diaphonie. Il est à noter que l'invalidité

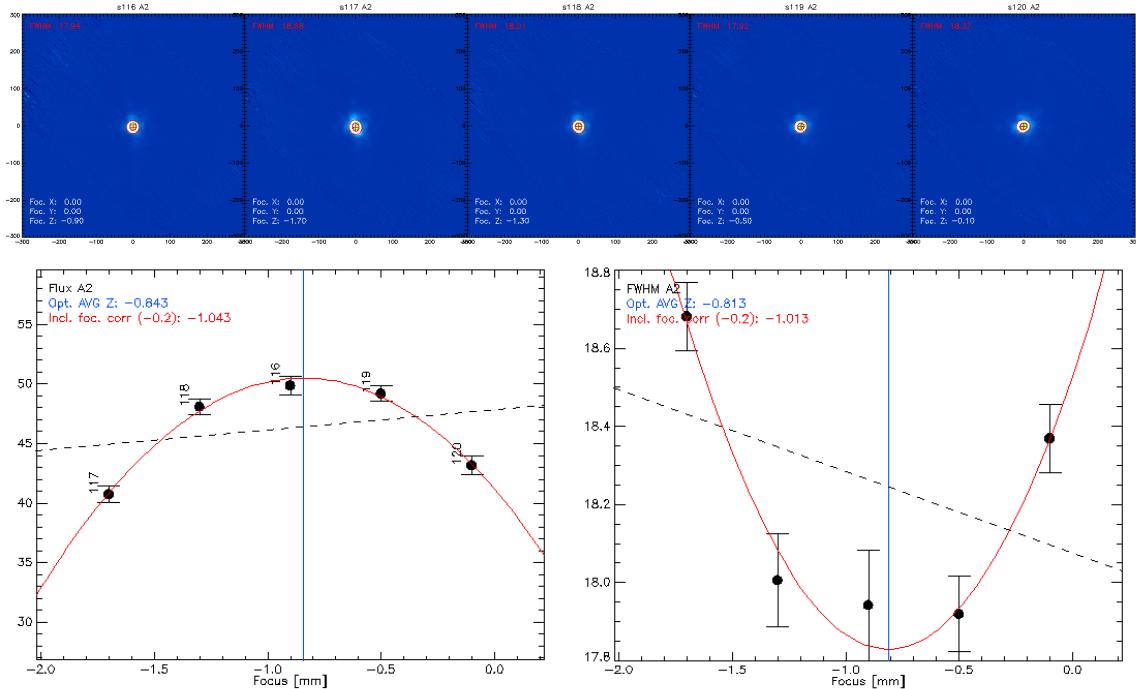


Figure 3.7 : Exemple de procédure de focalisation du télescope. **Haut** : cartes à 150 GHz d’Uranus obtenue pour cinq distances primaire-secondaire. **Bas** : Flux d’Uranus (*gauche*) et largeur à mi-hauteur (*droite*) mesurées dans chacune des cartes (noir). L’ajustement des points avec une loi parabolique (rouge) permet de mesurer la distance permettant une focalisation optimale du télescope (bleu).

d’un détecteur n’est pas définitive, et est liée à la mesure des fréquences de résonance par l’électronique de lecture et à la connaissance *a priori* de la position de ces fréquences de résonance dans le plan focal. Ainsi, un détecteur non-valide ne l’est pas définitivement, et peut redevenir valide. Ces valeurs sont de 84% et 90% à 260 et 150 GHz respectivement, montrant qu’une grande majorité des détecteurs peuvent être lus, et l’efficacité du multiplexage fréquentiel utilisé dans NIKA2.

Lobe instrumental de NIKA2

La connaissance de la réponse angulaire de l’instrument est cruciale à la construction et à l’étalonnage des cartes du ciel. Le lobe du couplage télescope-NIKA2 est une fonction complexe, dont la forme peut être modélisée de plusieurs façons. Nous renvoyons le lecteur vers la section 6 de [11] pour plus de détails. Le lobe principal, lié au pic central de la tache d’Airy du miroir primaire à travers l’ouverture du vertex, est modélisé comme une Gaussienne à deux dimensions, dont la largeur est ajustée sur les *beammaps* réalisées sur des sources ponctuelles brillantes lors des campagnes de *commissioning* de NIKA2. Le lobe complet, comprenant des composantes additionnelles dues par exemple aux déformations du miroir primaire, est mesuré à deux dimensions par coaddition de plusieurs *beammaps* afin d’obtenir un signal significatif. On peut alors définir l’efficacité du lobe principal comme le rapport entre l’angle solide qu’il couvre et l’angle solide du lobe total.

Le lobe instrumental est représenté sur la figure 3.8, extraite de [11]. On remarque sur le panneau gauche la structure complexe du lobe à 2D, incluant le lobe principal au centre, la figure de diffraction due à la structure quadrupolaire fixant le miroir secondaire au primaire (diagonales), et un anneau résultant de la diffraction par les jointures des panneaux pavant le miroir primaire. On voit sur le panneau droit que le lobe complet est bien modélisé comme une somme de trois gaussiennes, alors

Caractéristique	Matrices A1 et A3	Matrice A2
<i>Bandes passantes</i>		
Longueur d'onde de référence [mm]	1.15	2.00
Fréquence de référence [GHz]	260	150
<i>Nombre de détecteurs</i>		
Nombre de détecteur présents	1140×2	616
Fraction de détecteurs valides [%]	84	90
<i>Résolution angulaire et champ de vue</i>		
Largeur à mi-hauteur (FWHM) du lobe principal [arcsec]	11.1 ± 0.2	17.6 ± 0.1
Efficacité du lobe principal [%]	47 ± 3	64 ± 3
Champ de vue instantané [arcmin]	6.5	6.5
<i>Incertitudes de pointage et d'étalonnage</i>		
Erreurs de pointage [arcsec]	< 3	< 3
Incertitude sur l'étalonnage absolu [%]	5	5
Incertitude RMS sur l'étalonnage de sources ponctuelles [%]	5.7	3.0
<i>Sensibilité</i>		
NEFD [$\text{mJy} \cdot \text{s}^{1/2}$]	30 ± 3	9 ± 1
Mapping speed [$\text{arcmin}^2 \cdot \text{mJy}^{-2} \cdot \text{h}^{-1}$]	111 ± 11	1388 ± 174

Table 3.1 : Caractéristiques et performances instrumentales de NIKA2 mesurées par Perotto *et al.* [11] (voir texte pour des explications sur la signification des termes).

que l'ajustement avec une seule gaussienne ne permet d'englober qu'une portion de la densité de flux de la source, qui est quantifiée par l'efficacité du lobe principal.

Les largeurs et efficacités des lobes principaux de l'instrument mesurées au cours du *commissioning* de NIKA2 sont rapportés en table 3.1. Les largeurs à mi-hauteur des lobes principaux sont de 11.1" et 17.6" à 260 et 150 GHz, respectivement. Ces valeurs sont proches de celles calculées pour la tache de diffraction de rayons monochromatiques réfléchis par les miroirs primaires et secondaires, comme nous l'avons décrit en 3.1.2 (page 51). Comme nous le discuterons en 3.4.1, ces valeurs sont intéressantes car elles permettent de résoudre les amas de galaxies même lointains avec un grand niveau de détail. L'efficacité du lobe principal est de 47% et 64% à 260 et 150 GHz respectivement, montrant l'importance des structures secondaires du lobe dans les observations avec NIKA2.

Étalonnage

Comme nous l'avons vu précédemment, le système d'acquisition de données de NIKA2 mesure la variation de fréquence de résonance des KID dans le temps. Il est donc nécessaire d'associer l'amplitude de ces variations lors du passage sur une source astrophysique à une densité de flux. Pour cela, des sources de flux connu sont observées régulièrement au cours des campagnes d'observation avec NIKA2. Nous avons déjà vu en 3.3.2 que les *beammaps* permettaient une mesure des gains des détecteurs individuels, qualifiée d'étalement relatif. L'étalement absolu est réalisé à l'aide de *scans* courts réalisés sur des sources brillantes de flux connu. Il est alors possible d'utiliser ces *scans* pour calculer la relation entre variation de fréquence de résonance et variation de puissance lumineuse reçue, et donc

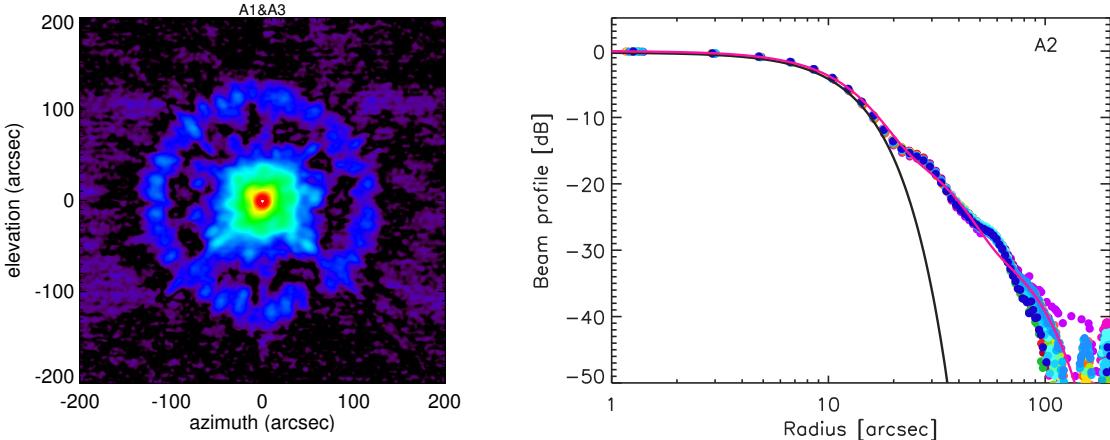


Figure 3.8 : **Gauche :** Lobe de NIKA2 à 260 GHz mesuré par empilement de cartes d’Uranus. **Droite :** Profils radiaux du lobe à 150 GHz mesurés pour plusieurs *beammaps* (points colorés). Le lobe principal et complet, ajusté avec un modèle à trois gaussiennes, sont respectivement représentés en noir et magenta. Figures extraites de [11].

d’étalonner les cartes en unités de densité de flux.

Sensibilité et *mapping speed*

La sensibilité de NIKA2 est quantifiée par sa NEFD (*Noise Equivalent Flux Density*). Celle-ci est définie comme l’erreur statistique à 1σ sur la densité de flux mesurée pour une observation d’une seconde d’une source ponctuelle. Elle est calculée pour une opacité nulle, c’est-à-dire une transmission atmosphérique égale à un (voir équation 3.1). Son évaluation pour NIKA2, décrite dans la section 10 de [11], est basée sur la dispersion du bruit résiduel dans les cartes de toutes les sources de densité de flux inférieures à 1 Jy observées pendant les campagnes de mesure des performances de l’instrument³. Les valeurs mesurées sont de 30 et 9 mJy · s^{1/2} à 260 et 150 GHz respectivement, faisant de NIKA2 une caméra à grande sensibilité.

La sensibilité de l’instrument peut également être définie par sa *mapping speed*. Celle-ci est définie comme la surface du ciel pouvant être cartographiée en une heure avec un niveau de bruit de 1 mJy, et est liée à la NEFD par :

$$M_s = \pi \left(\frac{d_{\text{FoV}}}{2} \right)^2 \frac{\eta}{\text{NEFD}^2}, \quad (3.7)$$

où η est la fraction de détecteurs valides, et d_{FoV} le diamètre du champ de vue instantané de l’instrument ($d_{\text{FoV}} = 6.5$ arcmin pour NIKA2).

En injectant les valeurs de NEFD mesurées au cours du *commissioning* de NIKA2, on trouve 111 et 1388 arcmin² · mJy⁻² · h⁻¹ à 260 et 150 GHz respectivement⁴. On peut alors calculer le temps nécessaire pour cartographier un objet de surface connue à un niveau de bruit choisi. Par exemple, pour un amas de masse $6 \times 10^{14} M_\odot$ à $z = 0.5$, la surface⁵ correspondant au rayon caractéristique R_{500} vaut $A = 33$ arcmin², et un niveau de bruit de $N = 0.1$ mJy à 150 GHz sur cette surface peut être atteint en un temps t :

$$t = \frac{A}{M_s \times N^2} = 2.4 \text{ h}. \quad (3.8)$$

³Voir section 4.3.3 (page 88) pour la définition des cartes de dispersion de bruit résiduel.

⁴Notons que tout comme la NEFD, cette *mapping speed* est calculée à opacité nulle.

⁵Voir figure 3.9 et discussion en section 3.4.1 (page 63).

Il est par conséquent possible de cartographier l'effet SZ en direction d'amas de galaxies en peu de temps avec NIKA2. Par exemple, nous verrons dans la section suivante le cas du grand programme SZ de NIKA2, qui réalise un suivi avec NIKA2 d'environ cinquante amas de galaxies, en passant environ deux à vingt heures d'observation par amas, pour un temps total de 300 heures. De telles durées sont courtes, en particulier en comparaison aux temps nécessaires aux observations X assez profondes pour pouvoir extraire des informations spectroscopiques. C'est la raison pour laquelle la grande *mapping speed* de NIKA2 permet de diminuer la nécessité d'observations profondes en X, en permettant la combinaison d'observations X peu profondes et d'observations SZ, comme nous le verrons par la suite.

3.4 Le grand programme SZ de NIKA2

Nous avons discuté au chapitre précédent (2.4.3) de l'intérêt pour la cosmologie des suivis dédiés d'amas de galaxies. Nous décrivons dans cette section l'un de ces programmes de suivi : le grand programme SZ de NIKA2 (ou LPSZ, [125]). Celui-ci consiste en un temps garanti de 300 heures d'observations NIKA2, parmi les 1300 heures offertes à la collaboration NIKA2 pour la construction de l'instrument.

3.4.1 La caméra NIKA2 : un instrument idéal pour la mesure de l'effet SZ

Les sections précédentes ont détaillé le fonctionnement de la caméra NIKA2 et ses performances. Celles-ci (voir table 3.1 et section 3.3.4) font de NIKA2 un instrument particulièrement adapté à la mesure de l'effet tSZ vers des amas de galaxies distants.

Un instrument multilongueur d'onde

La caméra NIKA2 cartographie le ciel simultanément dans deux bandes de fréquence, centrées autour de 150 et 260 GHz respectivement. Comme le montre la figure 3.4, ces bandes permettent de tirer parti de la dépendance spectrale de l'effet tSZ, en détectant un décrément en brillance de surface dans la bande à 150 GHz, et un incrément à 260 GHz. Cependant, plusieurs considérations empêchent la détection à 260 GHz. D'une part, NIKA2 est bien plus sensible à 150 GHz qu'à 260 GHz (la NEFD y est trois fois plus faible – voir table 3.1), exprimant la nécessité d'observer des densités de flux plus grandes (ou d'observer plus longtemps) pour atteindre le même niveau de rapport signal sur bruit. D'autre part, l'incrément en brillance de surface dû à l'effet tSZ est (en valeur absolue) plus faible dans les bandes passantes à 260 GHz que le décrément dans la bande à 150 GHz. En ajoutant le fait que la transmission atmosphérique est plus faible à 260 GHz qu'à 150 GHz pour une même teneur en eau (voir figures 3.4 et 3.1), en pratique, il est bien plus difficile de détecter un signal SZ significatif dans les bandes à 260 GHz que dans la bande à 150 GHz. Les observations à 260 GHz sont toutefois très utiles : le signal SZ y étant presque absent, elles permettent d'identifier les contaminants astrophysiques présents dans le champ des amas, et d'estimer leurs flux. Nous détaillerons dans les chapitres 5 et 6 l'importance de cette contamination et les méthodes pouvant être utilisées pour la traiter.

Haute résolution angulaire et grand champ de vue

Nous avons vu au chapitre précédent que les amas de galaxies avaient des masses caractéristiques $M_{500} \sim 10^{14} - 10^{15} M_\odot$. Des amas de telles masses ont des rayons caractéristiques R_{500} de l'ordre du

mégaparsec. Le diamètre angulaire de l'amas est définie comme le double de l'angle sous-tendu par ce rayon :

$$2\theta_{500} = 2 \times \arctan \left(\frac{R_{500}}{\mathcal{D}_A(z)} \right), \quad (3.9)$$

où $\mathcal{D}_A(z)$ est la distance diamètre angulaire jusqu'au redshift z . Dans l'hypothèse d'un Univers plat, celle-ci s'écrit en fonction du paramètre de Hubble $H(z)$ comme :

$$\mathcal{D}_A(z) = \frac{1}{1+z} \int_0^z \frac{c dz}{H(z)}. \quad (3.10)$$

Cette distance a la particularité de diminuer avec le redshift⁶ à $z \gtrsim 1.6$. Par conséquent, un objet de taille physique donnée apparaîtrait sous un angle plus petit à $z = 1.6$ qu'à $z = 0.1$, mais plus grand à $z = 2$ qu'à $z = 1.6$. Si on s'intéresse à un amas de galaxies de rayon caractéristique R_{500} , cette augmentation de la taille angulaire à haut redshift est compensée par la dépendance en redshift du rayon lui-même⁷, au travers de la densité critique de l'Univers (voir équation 2.1). La taille angulaire apparente d'un amas de masse M_{500} donnée varie donc très lentement avec le redshift (par un facteur ~ 2 entre $z = 0.5$ et $z = 2$). Par conséquent, une résolution angulaire modérée permet de résoudre tous les amas de l'Univers au delà d'une masse donnée.

Ce comportement est représenté pour des amas de différentes masses sur la figure 3.9. On voit que pour des amas de masses réalistes, $2\theta_{500}$ n'est résolu par *Planck* qu'à des redshifts inférieurs à ~ 0.3 . On voit également que les télescopes ACT et SPT, de résolutions comparables de l'ordre de la minute d'arc, permettent de résoudre les amas distants, mais avec un faible niveau de détail (voir discussion ci-après). En revanche, le lobe de NIKA2 permet de résoudre des tailles angulaires bien plus petites que $2\theta_{500}$ pour des amas jusqu'à haut redshift. De plus, le grand champ de vue de NIKA2 lui permet d'imager en instantané 6.5 arcmin, ce qui correspond à une taille plus grande que le diamètre caractéristique d'amas à $z \gtrsim 0.5$.

La résolution d'amas de galaxies est cruciale pour l'étude de leurs propriétés physiques. En effet, les observations réalisées par un instrument fournissent une image du ciel convoluée par son lobe. Par conséquent, la carte d'un amas de galaxies plus petit que la résolution de l'instrument qui l'observe est principalement une carte du lobe de l'instrument, qui ne permet pas une étude détaillée de la structure de l'amas. C'est pourquoi les cartes de l'effet SZ construites par la collaboration *Planck* [134], de résolution angulaire $\sim 10'$, permettent seulement l'étude détaillée d'amas proches et massifs (comme l'amas de Coma, $z = 0.023$, $M_{500} \simeq 7 \times 10^{14} M_\odot$ [135]).

De la même façon, résoudre un amas (c'est-à-dire l'observer avec un lobe plus petit que son rayon caractéristique) ne permet pas nécessairement un grand niveau de détail dans l'extraction de ses propriétés physiques. Pour des études précises, il est donc nécessaire que la taille angulaire de l'amas soit non seulement supérieure à la taille du lobe de l'instrument utilisé, mais aussi grande devant celle-ci. C'est pourquoi les instruments ayant une résolution de l'ordre de la minute d'arc, comme SPT et ACT, ne permettent pas l'étude détaillée d'amas distants : au delà d'un redshift de 0.5, les tailles d'amas $2\theta_{500}$ ne sont au maximum que cinq fois plus grands que les lobes de ces instruments. La cartographie détaillée de l'effet SZ dans les amas de galaxies nécessite donc l'utilisation d'un ins-

⁶En injectant l'expression du paramètre de Hubble de l'équation (1.10), et les valeurs des paramètres cosmologiques de la table 1.1, on peut tracer l'évolution de \mathcal{D}_A avec le redshift.

⁷On peut aussi définir le contraste de densité par rapport à la densité critique de l'Univers aujourd'hui. Dans ce cas, θ_{500} augmente avec le redshift à partir de $z \sim 1.6$. Le choix de définition du paramètre de contraste est purement conventionnel et n'a pas d'impact sur les analyses, tant qu'il est traité de manière cohérente tout au long de celles-ci.

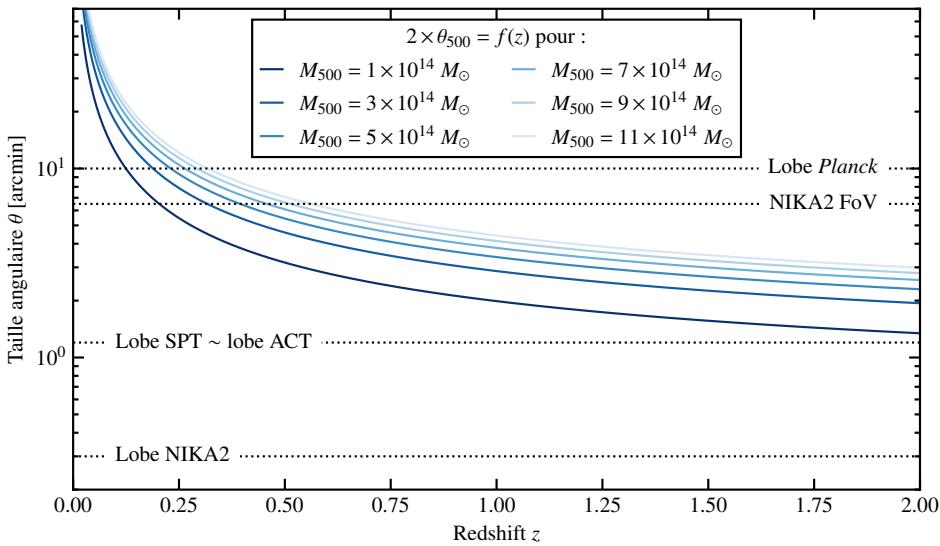


Figure 3.9 : Évolution du diamètre angulaire caractéristique $2\theta_{500}$ d’amas de galaxies de différentes masses avec le redshift. Les lobes de NIKA2 ($\sim 18''$ à 150 GHz), SPT ($\sim 1.2''$), ACT ($\sim 1.4''$), et *Planck* ($\sim 10''$) sont représentés à titre indicatif, de même que le champ de vue instantané (FoV pour *field of view*) de NIKA2 ($6.5''$).

trument à plus haute résolution angulaire. Comme le montre la figure 3.9, le lobe de NIKA2 à 150 GHz est inférieur par au moins un ordre de grandeur au diamètre angulaire $2\theta_{500}$ de tous les amas de masse $M_{500} > 3 \times 10^{14} M_\odot$, quel que soit leur redshift. NIKA2 est donc un instrument idéal pour les observations SZ à haute résolution des amas de galaxies distants de l’Univers.

Grande sensibilité

La cartographie de l’effet SZ en direction d’amas de galaxies requiert une grande sensibilité. En effet, la distorsion spectrale due à l’effet SZ est faible. Dans le cas d’observations avec NIKA2 à 150 GHz, la brillance de surface maximale en direction des amas est de l’ordre du mJy/beam ; à titre comparatif, la densité de flux d’Uranus à cette fréquence est de l’ordre de la dizaine de Jy/beam, de même que les fluctuations du bruit de l’atmosphère (ce point sera traité au chapitre suivant). Par conséquent, une grande sensibilité est nécessaire à la mesure de l’effet SZ dans un temps raisonnable. NIKA2 est capable de cartographier environ 1400 arcmin^2 avec un niveau de bruit de 1 mJy en une heure, ce qui permet d’atteindre le niveau de bruit requis pour la détection de l’effet SZ rapidement, comme discuté dans l’équation (3.8).

Ainsi, les caractéristiques instrumentales de NIKA2 en font un instrument de choix pour étudier avec un grand niveau de détail les amas de galaxies lointains, non-résolus par *Planck* et difficilement par SPT ou ACT.

3.4.2 Échantillon du grand programme SZ

Le temps alloué aux observations du LPSZ a été réparti sur un échantillon d’environ cinquante amas de galaxies. Cet échantillon est l’une des grandes forces du programme : il couvre une grande gamme de masses ($M_{500} \in [3, 11] \times 10^{14} M_\odot$) à haut redshift ($z \in [0.5, 0.9]$), avec des amas sélectionnés dans

des catalogues SZ. Il est représenté en figure 3.10.

Comme nous l'avons discuté au chapitre précédent (voir 2.2.3), les échantillons sélectionnés par effet SZ ont l'avantage d'être proches d'échantillons sélectionnés en masse. En effet, le paramètre de Compton intégré étant lié au contenu en énergie thermique de l'amas, il est étroitement lié à la masse (voir section 7.1.1, page 142). Ainsi, un échantillon construit avec pour critère une valeur de Y supérieure à un seuil choisi sélectionnera les amas sur un critère lié à leur masse, ce qui permet d'éviter de favoriser un type d'amas particulier. Un tel échantillon est alors dit représentatif, puisque la distribution des propriétés physiques des amas le constituant suit *a priori* la même distribution que les amas de l'Univers dans la gamme de masse considérée. À titre d'exemple, les amas sélectionnés selon leur pic de luminosité en X ne sont pas représentatifs : la brillance de surface X étant proportionnelle au carré de la densité d'électrons intégrée le long de la ligne de visée (équation 2.6), une telle sélection favorise les amas à cœur dense, qui présentent la caractéristique générale d'être plus relaxés (voir discussion sur les amas à cœur froid en 3.4.3).

La sélection de l'échantillon a donc été basée sur les catalogues SZ existant au début du programme, à savoir les catalogues *Planck* et ACT [5, 136]. Les amas ont été choisis parmi ceux présents dans les catalogues et visibles depuis le télescope de 30 mètres de l'IRAM (cf. 3.1.1 et figure 3.2), et de manière à couvrir la gamme de masse considérée de manière homogène. Pour cela, dix « boites » ont été définies, délimitant deux intervalles en redshift et cinq en masse. Cinq amas ont été tirés au hasard pour remplir chacune de ces boites, dans le catalogue ACT pour les deux boites de basse masse, et dans le catalogue *Planck* pour les autres. Dans les deux cas, les estimateurs de masse utilisés pour la répartition sont calculés à partir du signal SZ intégré mesuré par les relevés en utilisant la relation d'échelle de [94]. Sans cette procédure de division en boîtes du plan masse-redshift, les amas auraient été sélectionnés aléatoirement dans les catalogues *Planck* et ACT, et leur distribution en masse et en redshift aurait suivi la fonction de masse sous-jacente. À l'inverse, l'échantillon du grand programme SZ constitue une couverture homogène de la gamme de masse et de redshift considérée. Nous verrons en 7.5.2 (page 160) l'impact de cette sélection sur l'ajustement de la relation d'échelle masse-observable à partir des données du grand programme SZ.

Le suivi SZ à haute résolution d'un échantillon d'amas aussi nombreux, à des redshifts supérieurs à 0.5, et sur une si grande gamme de masse est sans précédent. Il permettra de comparer les propriétés d'amas de galaxies distants, mesurées en SZ avec une grande résolution angulaire, à celles mesurées sur des échantillons d'amas proches, mesurés en X ou en SZ avec une faible résolution.

Comme nous l'avons mentionné, le temps alloué au LPSZ est de 300 heures d'observation. La distribution de ces heures entre les différents amas de l'échantillon a été calculée de façon à obtenir un signal de qualité équivalente pour chacun des amas. Le temps alloué à un amas est ainsi calculé comme le temps d'observation nécessaire à l'obtention d'un rapport signal-sur-bruit du profil de brillance de surface SZ de 3σ au rayon caractéristique R_{500} . Ce calcul est basé sur plusieurs hypothèses. D'une part, il a été supposé que le profil de pression de chaque amas était donné par le profil de pression universel du milieu intra-amas d'Arnaud *et al.* [94], discuté en 2.2.4 (page 38). D'autre part, le calcul a été fait dans l'hypothèse où la masse M_{500} des amas de l'échantillon était égale à celle donnée par les catalogues *Planck* et ACT dans lesquels les amas sont choisis. Celles-ci ont été obtenues par l'utilisation d'une relation d'échelle liant l'observable des relevés SZ à la masse. À partir de ces deux hypothèses, un profil de paramètre de Compton a été calculé pour chaque amas. Le temps d'observation en a été déduit en suivant le calcul présenté en équation (3.8) comme le temps nécessaire pour atteindre un niveau de bruit trois fois inférieur à la valeur de ce profil à R_{500} . On note que tout écart aux hypothèses émises au cours de ce calcul, par exemple sur le profil de pression du milieu intra-amas ou

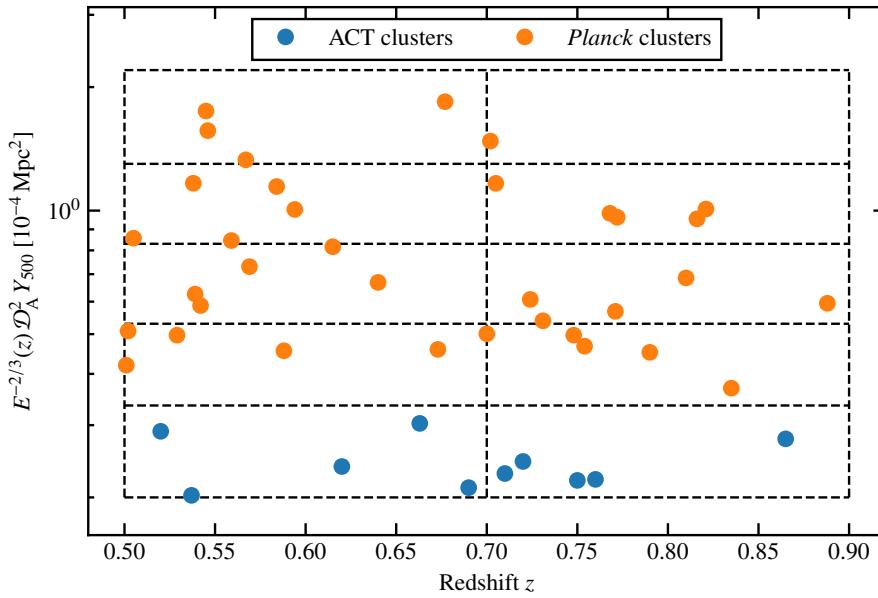


Figure 3.10 : Distribution des amas du grand programme SZ dans le plan redshift – paramètre de Compton intégré. Les points bleu et orange représentent respectivement les amas extraits des catalogues ACT [136] et *Planck* [5]. Les lignes noires pointillées représentent les limites des boîtes utilisées pour construire l'échantillon.

sur la relation d'échelle masse-observable, entraîne un changement dans la qualité du signal attendu. Ainsi, un amas dont la masse est mal estimée par la relation d'échelle, ou dont le profil de pression présente un écart au profil universel, peut présenter une qualité de données significativement différente de celle attendue. De plus, les temps d'observation ont été calculés avant la construction de NIKA2, et les performances réelles de l'instrument étaient encore inconnues. Par conséquent, il est possible que la qualité générale des observations NIKA2 des amas de l'échantillon soient moins bonne qu'attendu. Ce point nécessite une grande précaution, car pour étalonner les outils nécessaires à la cosmologie, l'échantillon doit être représentatif, et les observations des amas doivent avoir une qualité équivalente afin de ne pas donner un poids trop important à un type d'amas donné.

3.4.3 Objectifs scientifiques du grand programme SZ

Le grand programme SZ exploite les performances de NIKA2 et sa capacité à mesurer l'effet SZ à haute résolution dans le but de répondre à plusieurs questions de la cosmologie avec des amas de galaxies.

Études détaillées des amas de l'échantillon

Comme discuté en section 3.4.1, les performances et caractéristiques de NIKA2 sont en adéquation avec les besoins pour l'étude à haute résolution d'amas de galaxies. Par conséquent, les observations des amas du LPSZ avec NIKA2 permettent de les imager avec un grand niveau de détail. Les études des propriétés individuelles de chaque amas sont alors rendues possibles.

Les résultats de ces études seront rendus disponibles par la collaboration NIKA2 à la fin du programme. Certaines études individuelles le sont déjà, notamment les deux premières études individuelles du LPSZ : [137], et [138], menée au cours de cette thèse et présentée au chapitre 6. Les produits

finaux seront pour chacun des amas de l'échantillon du LPSZ :

- Les cartes NIKA2 de la région de l'amas, à 150 et 260 GHz, étalonnées et soustraites du bruit corrélé (voir chapitre 4) ;
- Le profil de pression de l'amas mesuré avec le logiciel PANC02, développé dans le cadre de cette thèse, et présenté au chapitre 5 ;
- Les profils d'autres propriétés thermodynamiques du milieu intra-amas obtenues par combinaison avec des données externes, notamment dans le domaine des rayons X ;
- Les valeurs des grandeurs caractéristiques intégrées de l'amas, R_{500} , Y_{500} , et M_{500} , qui pourront être utilisées par la communauté pour améliorer la mesure de la relation d'échelle masse-observable dans les analyses cosmologiques de catalogues d'amas de galaxies – voir section 2.3.2, et chapitre 7.

Toutes ces informations seront rendues publiques sous la forme d'un catalogue, et utilisables par la communauté pour combiner les informations issues de l'analyse des amas du LPSZ avec d'autres données futures, afin d'enrichir la connaissance de la physique des amas de galaxies grâce à l'échantillon du LPSZ.

Profil de pression moyen des amas de galaxies

L'un des objectifs principaux du LPSZ est la mesure du profil de pression moyen d'amas de galaxies à haut redshift. Le profil universel d'Arnaud *et al.* [94], obtenu par l'exploitation de l'échantillon REXCESS en X à $z < 0.2$, a été utilisé dans la construction de plusieurs catalogues d'amas en SZ, notamment par les collaborations ACT [2] et *Planck* [5]. Il est intéressant de noter que dans les deux cas, les catalogues sont construits à partir de détections des amas par effet SZ, et qu'ils comportent des amas à des redshifts bien plus grands que 0.2, redshift maximal du suivi REXCESS. Deux questions se posent alors :

- Le profil de pression moyen des amas pourrait-il être différent de la mesure d'Arnaud *et al.* ?
Une différence pourrait avoir plusieurs origines. D'une part, les mesures de profil de pression en X requièrent l'estimation de la température du milieu intra-amas par spectroscopie, qui présente plusieurs incertitudes systématiques complexes [69]. Il est donc possible que les profils de pression mesurés en X et en SZ ne soient pas strictement équivalents. D'autre part, le profil de pression moyen des amas pourrait évoluer avec le redshift. Dans ce cas, l'utilisation d'un profil de pression évalué à $z < 0.2$ serait inadaptée à la détection d'amas aussi lointains que $z = 2$ (dans le cas du relevé ACT), ou même jusqu'à $z = 1$ (dans le cas de *Planck*). Enfin, ce profil de pression pourrait ne pas être représentatif de tous les amas de l'Univers, à cause des effets de sélection caractéristiques des échantillons d'amas détectés en X discutés précédemment.
- Quel impact aurait une variation du profil de pression moyen sur les analyses cosmologiques basées sur l'effet SZ ?
Comme nous l'avons vu, la connaissance du profil de pression moyen des amas est nécessaire à l'extraction de catalogues d'amas à partir d'observations millimétriques. Une variation du profil – avec la méthode de détection ou avec le redshift – pourrait ainsi *a priori* modifier les catalogues issus des cartes de *Planck* et ACT, et donc les résultats des mesures de paramètres cosmologiques.

Cet impact a été quantifié dans le cas d'analyses du spectre de puissance de l'effet SZ par Ruppin *et al.* [97], en analysant la carte de paramètre de Compton issue des données de *Planck* [134] avec différents profils de pression moyens. Leurs résultats montrent qu'une modification de l'amplitude du profil de pression aussi faible que 15% peut affecter significativement les résultats des analyses cosmologiques.

L'objectif du LPSZ est de répondre à ces questions en mesurant le profil de pression moyen des amas de son échantillon. Plusieurs études récentes s'intéressent à de possibles variations du profil de pression moyen des amas avec le redshift, mais utilisent principalement des observations X (voir par exemple [139, 140]). L'utilisation de données SZ à haute résolution offrira des contraintes complémentaires, et créera l'opportunité de comparaisons de résultats multilongueur d'onde.

Combinaisons multilongueur d'onde et propriétés thermodynamiques

Comme nous l'avons vu au chapitre précédent, les observations de l'effet tSZ permettent de sonder la pression thermique des électrons du milieu intra-amas. Si celle-ci donne des informations importantes sur les amas, d'autres propriétés thermodynamiques revêtent également une grande importance pour l'étude de la physique de ces objets, et ne sont pas mesurables par l'effet tSZ seul.

C'est la raison pour laquelle le LPSZ comporte un suivi de tous ses amas à d'autres longueurs d'onde, complémentaires à l'effet SZ. Un suivi en X avec les satellites *XMM-Newton* et *Chandra* est notamment l'un des grands points forts du programme. La combinaison de ces observations avec celles effectuées avec NIKA2 permettra la première étude des propriétés thermodynamiques d'un échantillon d'amas à $z > 0.5$ avec des données X et SZ de qualité comparable. En effet, comme détaillé en section 2.2 (page 29), les observations X et SZ sont très complémentaires, puisqu'elles sondent respectivement la densité et la pression du milieu intra-amas qui peuvent être combinées pour calculer la masse de l'amas. Mais d'autres propriétés peuvent également être calculées par cette combinaison. Par exemple, le profil de température du milieu intra-amas, dans l'hypothèse d'un gaz parfait, s'écrit :

$$k_B T_e(r) = P_e(r)/n_e(r), \quad (3.11)$$

et permet d'étudier l'état dynamique de l'amas, pouvant révéler des mécanismes de fusion de plusieurs sous-amas (voir par exemple [141, 142]). Les mesures de température permettent également l'étude des corrections relativistes à l'effet SZ, qui dépendent de cette dernière, comme décrit en section 2.2.3 (page 32). On peut également définir le paramètre d'entropie, introduit par Voit *et al.* [143] comme :

$$K_e(r) = P_e(r) n_e(r)^{-5/3}. \quad (3.12)$$

Cette quantité, souvent nommée entropie du milieu intra-amas, est liée au caractère adiabatique de l'évolution de l'amas ; dans l'hypothèse où le milieu intra-amas est un gaz parfait monoatomique d'électrons, on peut écrire la loi de Laplace $PV^\gamma = \text{cste}$ comme :

$$P_e n_e^{-5/3} = \text{cste}, \quad (3.13)$$

et la combinaison avec l'équation (3.12) permet d'identifier K_e comme le terme constant. Ce paramètre d'entropie est donc intimement lié à l'entropie thermodynamique du milieu intra-amas : la formule de Sackur-Tetrode donne l'entropie par particule d'un gaz parfait s , qui peut s'exprimer en fonction de K_e comme :

$$s = k_B \ln K_e^{3/2} + \text{cste}. \quad (3.14)$$

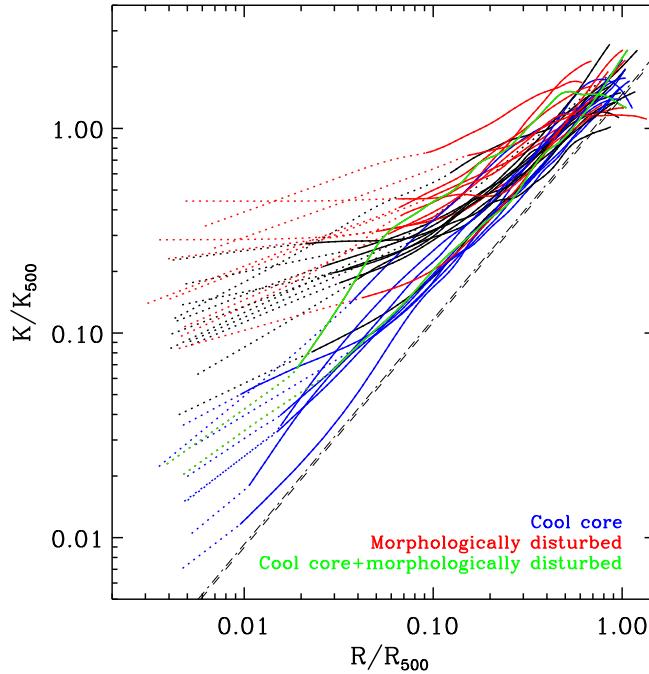


Figure 3.11 : Profils d’entropie mesurés à partir des observations *XMM-Newton* des amas de l’échantillon REXCESS. Les lignes pointillées noires marquent le profil en $r^{1.1}$ attendu pour une formation d’amas purement par effondrement gravitationnel. Les courbes bleues et rouges sont respectivement les profils d’amas classifiés comme *cool-core* et à cœur perturbé. Figure extraite de [111].

Le paramètre K_e peut donc être interprété comme une mesure en échelle exponentielle de l’entropie par particule s , ou encore comme une mesure du nombre d’états microscopiques Ω par comparaison de l’équation (3.14) et de l’équation de Boltzmann, $S = k_B \ln \Omega$.

Le paramètre d’entropie, simplement nommé entropie dans la suite, est d’un intérêt considérable pour l’étude de la physique des amas (voir [144] – en particulier la section IV – pour une revue). Il a historiquement été étudié dans des simulations d’amas de galaxies. Dans le cas où seuls les processus gravitationnels sont considérés pour la formation des amas, l’entropie dans les amas croît en loi de puissance avec le rayon. Ce comportement est observé dans les simulations à N corps considérant des particules de matière sombre seulement, donnant des profils d’entropie de forme $K_e(r) \propto r^{1.1}$ [144]. Dans les simulations plus complexes, de même que dans les amas réels, on observe un écart à ce comportement, en particulier dans le cœur des amas. Cet écart est dû aux phénomènes tels que des fusions de sous-structures, ou encore l’injection d’énergie par des noyaux actifs de galaxies ou des explosions de supernovas, principalement localisés dans le cœur des amas, et qui augmentent l’entropie du gaz. Il est illustré en figure 3.11 pour l’échantillon REXCESS, discuté précédemment. On voit que le profil d’entropie des amas suit bien $K_e(r) \propto r^{1.1}$ à grand rayon, mais que le centre des profils montre un écart significatif. Cet écart est d’autant plus fort que le cœur des amas est perturbé : les profils bleus, correspondant aux amas *cool-core*, sont plus en accord avec $K_e(r) \propto r^{1.1}$ que les courbes rouges, correspondant aux amas perturbés. L’entropie K_e représente donc un traceur de l’histoire thermique du milieu intra-amas, son écart au profil d’entropie purement gravitationnel donnant une forme de mémoire des processus non-gravitationnels ayant eu lieu au sein des amas.

Les combinaisons de mesures X et SZ permettront donc au LPSZ d’étudier en détail les propriétés thermodynamiques du milieu intra-amas jusqu’à un redshift de 0.9, pour un échantillon d’amas sélec-

tionnés en SZ. Outre leur importance directe pour la physique des amas, ces mesures pourront être utilisées pour classer les amas en fonction de leurs propriétés dynamiques. En particulier, on trouve dans l’Univers des amas à cœur froid (ou *cool core*), dont les propriétés centrales indique un état de relaxation, et les amas au cœur perturbé. Ces deux types d’amas ont des propriétés globalement différentes, comme décrit par exemple dans l’étude de Hudson *et al.* [145], qui étudie les différences entre les amas *cool-core* et perturbés sur le critère de 16 propriétés physiques différentes. En outre, les amas à cœur froid possèdent une entropie centrale bien plus faible que les amas perturbés, caractéristique d’injections d’énergie non-thermiques moins importantes. Ils ont également une densité centrale plus élevée (et sont ainsi plus facilement détectés en X), ainsi qu’un profil de pression bien plus piqué en leur centre. Nous avons discuté de l’importance de la mesure du profil de pression moyen des amas de galaxies pour les études cosmologiques ; une différence significative entre deux populations d’amas pourrait également avoir un impact sur de telles études. Les observations du LPSZ, en étudiant la température, l’entropie, et la pression du milieu intra-amas à l’aide d’observations résolues en X et en SZ, pourront distinguer ces populations, et les étudier séparément pour approfondir la connaissance de leurs propriétés thermodynamiques et de leur distribution dans les différentes populations.

De plus, comme nous l’avons discuté au chapitre précédent, les observations SZ permettent de réduire le temps nécessaire par rapport aux observations X permettant les mesures spectroscopiques. Traditionnellement, les études résolues d’amas faisaient appel à des observations X profondes, afin de pouvoir extraire un profil de température par spectroscopie, et ainsi de pouvoir déduire une mesure du profil de pression. Avec un profil de pression fourni par les mesures en SZ, la spectroscopie X n’est plus nécessaire pour remonter à la masse. Toutefois, certains des amas du LPSZ ont été suivis en X avec des observations profondes, permettant les études spectroscopiques. Pour de telles observations, la combinaison des observations SZ et X (sans spectroscopie) pourra être comparée aux mesures X seules, pour étudier de potentielles différences pouvant être dues aux incertitudes systématiques affectant les mesures de température en spectroscopie X (voir [69] pour une revue).

Enfin, des suivis en optique sont également en cours pour les amas du LPSZ, avec le Gran Telescopio Canarias à l’observatoire des Canaries. Ce suivi spectroscopique permettra d’étudier la distribution des galaxies des amas du LPSZ, ainsi que leurs masses par dispersion des vitesses (voir par exemple [146] pour une étude similaire de suivi des amas *Planck* avec le Gran Telescopio Canarias). De plus, certains des amas du LPSZ font partie d’échantillons d’amas suivis par d’autres instruments ; par exemple, l’échantillon CLASH [58], ayant réalisé un suivi d’amas avec le *Hubble Space Telescope* dont les données sont aujourd’hui publiques. Ces données permettront des études de la masse d’amas du LPSZ par des mesures optiques, en utilisant la dispersion des vitesses pour le suivi spectroscopique et le lentillage gravitationnel pour les données de HST. Ces estimateurs de masses, comme nous l’avons discuté en 2.2 (page 29), ne sont pas affectés par le biais hydrostatique, mais peuvent souffrir de biais différents [60, 62]. La comparaison des mesures de masse dans ces suivis avec celles obtenues par combinaison des observations SZ avec NIKA2 et X permettra donc d’étudier ce biais pour une partie de l’échantillon [147].

La relation d’échelle

Comme nous l’avons vu en 2.3.2, la connaissance du lien entre la masse d’un amas et son observable dans un relevé est cruciale pour l’exploitation cosmologique du relevé en question. Tout comme pour le profil de pression moyen du milieu intra-amas, l’étude de ce lien a aujourd’hui principalement été menée en utilisant des amas proches, et des masses souvent mesurées à l’aide uniquement d’observations X (voir par exemple [94, 148]). Les mêmes questions se posent alors, interrogeant sur de

possibles variations de la relation avec le redshift, avec l'utilisation de données SZ, et sur l'impact de telles variations sur les résultats cosmologiques.

Afin de pouvoir utiliser l'échantillon du LPSZ pour étudier de possibles évolutions de la relation d'échelle, il est nécessaire de connaître la masse des amas. Comme nous l'avons vu en 2.2.3, l'effet SZ seul ne permet pas de mesurer la masse d'un amas. Celle-ci peut cependant être estimée à partir de la combinaison d'observations SZ et X, comme discuté en 2.2.2. La mesure de la relation d'échelle fera donc appel à la combinaisons des observations SZ et X des amas de l'échantillon détailléé précédemment. En plus des mesures de masse, l'étude des propriétés thermodynamiques des amas et leur classification basée sur ces propriétés permettra des études systématiques détaillées sur l'impact des propriétés des amas sur la relation d'échelle. Par exemple, les amas perturbés pourraient suivre une relation d'échelle différente de celle suivie par les amas à cœur froid. Une étude intéressante serait la réduction de la dispersion autour de la relation d'échelle grâce aux propriétés thermodynamiques des amas, mesurées en détail à l'aide de la combinaison des observations avec NIKA2 et *XMM-Newton* ou *Chandra*. Jusqu'à présent, les relations d'échelles sont mesurées en supposant des relations en loi de puissance avec une dispersion intrinsèque, comme décrit en section 2.3.2 (page 42) et au chapitre 7. Une autre approche pourrait être la réduction de cette dispersion basée sur des propriétés observables des amas. De telles études sont menées pour mesurer les paramètres cosmologiques à l'aide des supernovas de type Ia, par exemple avec le modèle SALT2 ([149]). Dans de tels travaux (voir par exemple [150]), le module de distances de chacune des supernovas est corrigé à partir de propriétés observables de l'objet, permettant de réduire la dispersion autour du diagramme de Hubble. On parle alors des supernovas comme de chandelles standardisables, dont l'exploitation cosmologique repose sur une correction empirique basée sur la mesure des propriétés de chacun des objets. Des études similaires pourraient être réalisées avec les amas du grand programme SZ, grâce à des observables des amas telles que leurs propriétés thermodynamiques, pour une approche différente à la mesure de relations d'échelle.

Comme nous l'avons mentionné au chapitre précédent, la mesure de relations d'échelle fait intervenir beaucoup d'effets systématiques, tels que les effets de sélection dus au choix des amas de l'échantillon. Ces effets peuvent être pris en compte au prix d'une modélisation statistique complexe. L'étude de la relation d'échelle liant la masse au paramètre de Compton intégré au chapitre 7, en utilisant des simulations pour évaluer les différents effets systématiques affectant la mesure de cette relation dans le cadre du LPSZ.

3.5 La base de données du grand programme SZ de NIKA2

3.5.1 Contexte

Avec l'avancement du grand programme SZ de NIKA2, la quantité d'informations pertinentes pour le programme a augmenté rapidement. Il a donc été nécessaire de mettre en place un moyen permettant de centraliser ces informations, comportant des renseignements sur les observations NIKA2 ou sur l'existence de données externes. Un exemple est le besoin d'avoir accès aux coordonnées les plus récentes d'un amas. En effet, les amas du grand programme SZ sont en majorité issus du catalogue *Planck* [5], construit à partir d'observations dont la résolution est proche de la taille du champ de vue de NIKA2. Les coordonnées des amas dans ce catalogue sont donc imprécises ; en observant à ces coordonnées avec NIKA2, il est possible d'observer à côté de l'amas. Il est donc préférable d'utiliser les coordonnées d'amas obtenues avec des suivis à plus haute résolution, par exemple en X. Il est donc important de pouvoir accéder aux coordonnées d'un amas les plus précises possibles. Un autre

exemple d'utilisation est la possibilité d'avoir un outil permettant de connaître rapidement le nombre de *scans* ayant été effectués sur chaque amas, qui peut grandement faciliter la prise de décision sur la liste de sources choisies pour être observées au télescope. De même, savoir si la collaboration dispose de données X ou optiques associées à un amas permet de choisir les sources considérées pour une analyse multilongueur d'onde. Enfin, en sauvegardant les résultats des analyses de manière centralisée et accessible, la collaboration pourrait plus facilement garder une traçabilité de l'évolution des résultats avec les développements de logiciels de traitement des données, et choisir des amas de caractéristiques similaires pour une étude d'échantillon.

Il a donc été décidé de construire une base de données du grand programme SZ. Dans ce contexte, nous avons choisi de collaborer avec les développeurs de AMI (*Atlas Metadata Interface*, [151]), membres du service informatique du Laboratoire de Physique Subatomique et de Cosmologie. AMI propose un cadre permettant de développer une base de métadonnées, c'est-à-dire un regroupement d'informations sous forme de texte (les données de NIKA2 ou d'autres instruments ne sont pas centralisées). J'ai été nommé responsable scientifique de cette base de données, et ai pu collaborer avec les développeurs pour établir un cahier des charges, visant à concilier les besoins scientifiques et les possibilités offertes par AMI.

3.5.2 Contenu et interface de la base de données

Le cahier des charges défini pour la base de données comprend plusieurs exigences. D'une part, sa structure doit être assez souple pour stocker de l'information évoluant avec le temps, comme les coordonnées les plus précises existant pour chaque amas, l'existence de données externes, etc. D'autre part, la base de données doit pouvoir être utilisée à plusieurs étapes de la chaîne d'analyse du grand programme SZ :

- Pendant les observations au télescope, pour accéder aux coordonnées des amas, au temps d'observation nécessaire et déjà effectué, etc.
- Pendant l'analyse des données brutes, pour accéder à la liste des *scans* associés à un amas et à leurs propriétés.
- Pendant l'analyse des propriétés physiques des amas, pour accéder aux produits de l'analyse des données brutes (cartes SZ, etc.), aux propriétés des amas (redshift, masse estimée par *Planck*).

Chacune des étapes est associée à des besoins d'interactivité différents : une table avec interface graphique est utile pendant les observations et pour avoir des informations rapides sur les amas, mais un accès par ligne de commande peut être plus utile pour l'interface avec les logiciels d'analyse.

La base de données AMI-LPSZ est accessible sur demande pour les membres de la collaboration NIKA2⁸. Elle est composée de plusieurs tables, correspondant chacune à un type de données, et interconnectées. Les principales sont les suivantes :

- Une table CLUSTER, dont chaque élément est un amas. Les données stockées incluent le nom de chaque amas, ses coordonnées mesurées à partir de différents jeux de données (SZ, optique, X), sa masse et son redshift, le nombre de *scans* à réaliser pour compléter les observations de l'amas, etc. Un champ est également présent pour stocker l'information sur l'existence de données externes, par exemple X avec *Chandra* ou *XMM-Newton*, ou submillimétrique pour la contamination par des sources ponctuelles (voir section 4.5).

⁸<https://ami-nika2-lpsz.in2p3.fr/>

- une table SCAN qui regroupe chaque *scan* effectué par la caméra NIKA2 depuis sa mise en service. Les informations associées à chaque *scan* incluent la date et l'heure de l'observation, le nom de la source, le type de *scan*, l'opacité atmosphérique et l'élévation, ou encore les commentaires laissés par les observateurs au moment de sa réalisation. J'ai été en charge de l'écriture du programme pour interroger la base de données de l'IRAM dans laquelle ces informations sont consignées pour pouvoir les inclure dans AMI. Pour les *scans* dont le nom de source est l'un des amas de galaxies du grand programme SZ, un lien est créé entre le *scan* et l'entrée correspondant à l'amas dans la table CLUSTER, de sorte à ce qu'il soit possible d'extraire la liste des *scans* effectués pour chaque amas.
- Une table CAMPAIGN, dont chaque élément est une campagne d'observation avec NIKA2. Chaque élément de la table SCAN est ainsi associé à une campagne d'observations. Cela permet de conserver des informations sur chaque campagne d'observations, comme une météo particulièrement capricieuse ou un problème dans l'acquisition des données, afin de pouvoir éventuellement exclure des *scans* lors des analyses.
- Une table ANALYSIS, qui répertorie les résultats d'analyse des amas individuels. La structure de cette table et son contenu sont toujours en cours de discussion. À terme, le but est de garder une trace des analyses des données brutes de NIKA2 (chapitre 4) et des propriétés thermodynamiques (chapitre 5) dans AMI.

La base de données est accessible de deux façons. D'une part, l'interface web offre une interface graphique qui permet de chercher des entrées dans les différentes tables. Ces recherches peuvent être faites sur n'importe quel critère pouvant être construit à partir des différents champs présents dans les tables. Deux exemples sont présentés en figure 3.12 : une recherche des amas du programme dans l'intervalle de redshift $0.5 < z < 0.7$, et des *scans* de l'amas ACT-CL J0215.4+0030 effectués dans des conditions atmosphériques $0.1 < \tau_{225} < 0.2$. Dans les deux cas, AMI permet d'extraire la liste des objets – amas dans le premier cas, et *scans* dans le second – satisfaisant ces conditions, et leurs propriétés. Chaque élément de la liste peut également être inspecté individuellement.

La base de données peut également être interrogée en Python. Cette fonctionnalité est particulièrement intéressante pour interfacer les logiciels de traitement de données du grand programme SZ avec la base de données AMI-LPSZ. Par exemple, le traitement des données brutes de NIKA2 décrit au chapitre 4 nécessite de connaître la liste des *scans* associés à l'amas considéré. Cette liste est facilement accessible grâce à l'interface Python, ce qui simplifie grandement le processus. Il est également possible d'utiliser cette interface pour créer automatiquement les fichiers nécessaires pour les observations au télescope.

3.5.3 Développements futurs

La base de données AMI-LPSZ offre un grand nombre de fonctionnalités qui ont déjà énormément facilité la gestion du grand programme SZ. Certains aspects restent cependant à développer. L'un des objectifs serait de stocker dans AMI les résultats des différentes étapes de l'analyse des données de chaque amas, de l'analyse des données brutes décrite au chapitre 4 aux propriétés thermodynamiques traitées au chapitre 5. Pour l'instant, les résultats d'analyse ne sont pas stockés sur la base de données, car la structure des métadonnées à y sauvegarder n'a pas été définie. De plus, la gestion des données externes est pour l'instant sommaire. L'information sur ces données est consignée dans un champ correspondant à un dictionnaire clé-valeur pour chaque amas. Dans ce champ, chaque clé indique le nom d'un instrument (*XMM-Newton* ou *Chandra* pour les observations X, SPIRE, PACS ou FIRST pour les sources ponctuelles, etc), et la valeur associée est un booléen, indiquant l'existence ou non de

Left Side (Cluster Search):

Search criteria: Q2: z IN / ! IN [Min: 0.5, Max: 0.7] (Number of selected items: 25)

details	clusterId	shortName	fullName	pakoName	completion	ra
512	PSZ2G106	PSZ2-G106.15+25.75	PSZ2G106		0.00	284.1
539	PSZ2G111	PSZ2-G111.61-45.71	PSZ2-G111.6		100.00	4.6
346	PSZ2G081	PSZ2-G081.02+50.57	PSZ2G081		100.00	234.1
434	PSZ2G094	PSZ2-G094.56+51.03	PSZ2G094		100.00	227.1
477	PSZ2G099	PSZ2-G099.86+58.45	PSZ2-G099.8		100.00	213.1
651	PSZ2G133	PSZ2-G133.59+50.68	PSZ2G133		100.00	176.1

Right Side (Scan Search):

Search criteria: Q2: cluster:fullName IN / ! IN [PSZ2-G144.83+25.11, PSZ2-G201.50-27.31, PSZ2-G091.83+26.11, PSZ2-G228.16+75.20, ACT-CLJ0215.4+0030] and Q3: tau225 IN / ! IN [Min: 0.1, Max: 0.2] (Number of selected items: 67)

details	scanId	status	url	nbSubscans	startTime
20180117s257	VALID	null value	26	2018-01-17 19:33:56.000000	
20180117s258	VALID	null value	26	2018-01-17 19:40:39.000000	
20180117s259	VALID	null value	26	2018-01-17 19:46:22.000000	
20180117s260	VALID	null value	26	2018-01-17 19:52:10.000000	
20180117s261	VALID	null value	26	2018-01-17 19:57:52.000000	
20180117s262	VALID	null value	26	2018-01-17 20:03:32.000000	
20180117s263	VALID	null value	26	2018-01-17 20:09:13.000000	
20180117s264	VALID	null value	26	2018-01-17 20:15:01.000000	
20180117s265	VALID	null value	26	2018-01-17 20:20:42.000000	
20180117s266	VALID	null value	26	2018-01-17 20:26:29.000000	

Figure 3.12 : Deux exemples de requêtes dans l'interface graphique de AMI (haut) et leurs résultats (bas). **Gauche :** recherche des amas du grand programme SZ dont le redshift est compris entre 0.5 et 0.7. **Droite :** recherche des scans de l'amas ACT-CL J0215.4+0030 ayant été effectués pour une opacité atmosphérique τ_{225} entre 0.1 et 0.2.

données correspondant à chaque instrument. Un objectif de la base de donnée serait de compléter ces champs avec plus d'information, par exemple l'adresse à laquelle les données peuvent être récupérées. Cette base de données représente donc un outil déjà extrêmement utile pour le grand programme SZ de NIKA2, mais qui pourra le devenir encore plus avec une définition de la structure des données pertinentes à stocker pour les données externes et les résultats d'analyse.

CHAPITRE 4

Construction de cartes de l'effet SZ avec NIKA2

4.1	Composantes des données en temps	78
4.2	Soustraction du bruit corrélé	80
4.2.1	Estimation de mode commun	81
4.2.2	Masquage de source et analyse itérative	84
4.3	Des données en temps aux cartes du ciel	86
4.3.1	Projection et cartographie du signal d'intérêt	86
4.3.2	Cartes de couverture et de bruit	87
4.3.3	Coaddition de cartes	88
4.4	Performance de la décorrélation	89
4.4.1	Bruit corrélé résiduel	90
4.4.2	Filtrage du signal	91
4.5	Contamination par des sources ponctuelles	93
4.5.1	Données d'entrée	94
4.5.2	Mesures de coordonnées et flux à 260 GHz	95
4.5.3	Ajustement du spectre de chaque source	96
4.5.4	Inférence du flux à 150 GHz	97
4.6	Conclusions et perspectives	98

Dans le chapitre précédent, nous avons décrit l'instrument NIKA2, installé au télescope de 30 mètres de l'IRAM, et ses performances. Nous avons vu en quoi ses performances étaient parfaitement adaptées aux observations de l'effet SZ en direction d'amas de galaxies, et expliqué comment sont réalisées les observations avec NIKA2.

Dans ce chapitre, nous présentons la procédure employée pour produire des cartes de l'effet SZ avec NIKA2. Comme nous l'avons vu, les données brutes de l'instrument sont des données ordonnées en temps, ou TOI, donnant pour chaque détecteur de NIKA2 l'évolution de la puissance lumineuse

reçue au cours d'un mouvement de balayage du ciel. Nous détaillons les différentes composantes de ces TOI, et la méthode utilisée pour en isoler le signal d'intérêt, et le projeter en une carte du ciel. Nous présenterons enfin l'évaluation des performances de cette procédure au travers de deux critères : l'évaluation du bruit résiduel dans les données, et du filtrage subi par le signal d'intérêt. Ces critères sont également des outils utilisés pour tenir compte des systématiques que représentent le filtrage du signal et du bruit résiduel caractéristiques des cartes d'amas construites à partir des observations NIKA2, comme nous le verrons au chapitre 5. Nous présenterons également le logiciel PSTools, développé au cours de cette thèse, dont l'objectif est d'estimer l'amplitude de la contamination de l'effet SZ par des sources submillimétriques.

4.1 Composantes des données en temps

Au cours des observations avec NIKA2, le ciel est balayé suivant un motif donné pour enregistrer la variation de fréquence de résonance des KID de l'instrument, qui est ensuite convertie en variation de puissance lumineuse par la procédure d'étalonnage. Malheureusement, plusieurs composantes viennent s'ajouter à ces données en temps, en plus du signal astrophysique d'intérêt. La TOI d'un KID k à une fréquence ν est donc la somme de ces différentes composantes, et peut s'écrire :

$$\text{TOI}_k(t, \nu) = S_k(t, \nu) + \alpha_k(t)A_k(t, \nu) + \varepsilon_k(t)E_k(t) + G_k(t) + C_k(t) + N_k(t), \quad (4.1)$$

L'objectif de la réduction des données est d'isoler la première composante $S_k(t, \nu)$, représentant le signal astrophysique d'intérêt, des autres termes, qui sont des contaminants à ce dernier. Nous détaillons par la suite la nature de chacun des termes de l'équation (4.1).

Signal astrophysique d'intérêt : $S_k(t, \nu)$

L'objectif des observations avec NIKA2 est la production d'une cartographie aussi fidèle que possible du ciel dans les bandes passantes de l'instrument. Dans le cas des observations du grand programme SZ, les portions du ciel observées contiennent le signal SZ en provenance des amas de galaxies, mais aussi du signal provenant d'autres sources, par exemple des objets astrophysiques d'avant-plan (ou d'arrière-plan), des galaxies membres de l'amas, le fond diffus cosmologique, ou encore un fond créé par les sources non-résolues (nommé CIB pour *Cosmic Infrared Background*). L'ensemble de ces constituants forme une carte de la brillance de surface du ciel dans les bandes passantes de NIKA2, notée $M(x, y, \nu)$, où (x, y) représentent les coordonnées célestes. Lors du balayage de la région par le télescope, un détecteur k à une fréquence ν enregistre une variation de brillance de surface pouvant s'écrire :

$$S_k(t, \nu) = P_k(t, x, y) \times M(x, y, \nu), \quad (4.2)$$

où $P_k(t, x, y)$ est la matrice de pointage, décrivant la position du détecteur sur le ciel à un instant t :

$$P_k(t, x, y) = \begin{cases} 1 & \text{si le KID } k \text{ passe par la position } (x, y) \text{ au temps } t, \\ 0 & \text{sinon.} \end{cases} \quad (4.3)$$

Nous verrons en 4.3 (page 86) que c'est cette matrice qui est utilisée pour la projection des TOI en une carte après soustraction des autres composantes.

Bruit atmosphérique : $A_k(t, \nu)$

Comme nous l'avons mentionné en section 3.1.1 (page 50), l'atmosphère est l'un des contaminants principaux des observations millimétriques. En effet, l'atmosphère est caractérisée non seulement par une absorption qui atténue le signal reçu (voir figure 3.1), mais également par une émission aux longueurs d'onde millimétriques. Dans le domaine millimétrique, cette émission est bien modélisée par une émission de corps gris, et suit :

$$I_{\text{atm.}}(\nu) \propto \nu^2 T_{\text{atm.}} [1 - \exp(-\tau_\nu / \sin(el))], \quad (4.4)$$

où $T_{\text{atm.}}$ est la température de l'atmosphère, τ_ν l'opacité au zénith à la fréquence ν , détaillée dans l'équation (3.1), et el l'élévation des observations.

Pour une température donnée, l'émission de l'atmosphère dépend donc de son opacité τ_ν . Celle-ci, comme nous l'avons détaillé en 3.1.1 (page 50), est dominée par la présence de dioxygène et d'eau dans l'atmosphère. Le dioxygène étant réparti de façon homogène dans l'atmosphère, son émission thermique ne varie pas spatialement (à élévation constante). En revanche, la distribution d'eau n'est pas homogène, du fait notamment de la présence de nuages. Son émission thermique n'est donc pas homogène sur le ciel. Cette anisotropie est la source principale de la variation temporelle de signal enregistrée par NIKA2 au cours du balayage du ciel par le télescope, même dans les meilleures conditions météorologiques.

Dans le cadre d'observations avec le télescope de 30 mètres de l'IRAM, l'émission atmosphérique vue par les détecteurs à un instant donné est la même. En effet, la distance de Fraunhofer pour une ouverture $D = 30$ m vaut $2D^2/\lambda \sim 1000$ km aux fréquences d'observation de NIKA2, soit bien au-delà de l'atmosphère. Celle-ci peut alors être intégrée en champ proche¹, et on peut l'approximer au premier ordre comme homogène sur le ciel à un temps donné. Cette propriété permet d'approximer que l'émission atmosphérique se manifeste de la même façon pour tous les détecteurs dans les TOI obtenues au cours d'un *scan* : on a alors $A_k(t, \nu) = A(t, \nu) \forall k$. Le coefficient $\alpha_k(t)$ de l'équation (4.1) quantifie alors la réponse de chaque détecteur au bruit atmosphérique.

Bruit électronique : $E_k(t, \nu)$

Nous avons vu dans le chapitre 3 que les KID de NIKA2 étaient multiplexés, c'est-à-dire qu'un grand nombre de détecteurs sont lus simultanément par une même ligne de lecture. L'électronique utilisée pour la lecture de ces signaux induit une composante de bruit non-négligeable. Celle-ci est, par construction, corrélée pour les détecteurs lus par une même boîte électronique. Cette propriété de corrélation est utilisée pour isoler la composante de bruit électronique et la soustraire aux données en temps des détecteurs. Si le bruit atmosphérique est, comme nous l'avons vu dans la section précédente, corrélé entre tous les détecteurs, le bruit électronique ne l'est que pour les détecteurs lus par une même boîte électronique. Une composante en sous-bande électronique est également présente, corrélée pour les détecteurs couplés à une même ligne de lecture. Nous verrons en 4.2.1 (page 81) qu'il est ainsi possible de traiter les TOI des détecteurs par groupe de KID afin d'isoler au mieux le signal astrophysique d'intérêt du bruit. Le coefficient $\varepsilon_k(t)$ de l'équation (4.1) représente la réponse de chaque KID k à ce bruit électronique.

¹Voir <https://www.cv.nrao.edu/~sransom/web/Ch3.html>, section 3.2, ainsi que le chapitre 9 de [152] pour plus de détails.

Glitches et bruit du cryostat : $G_k(t)$ et $C_k(t)$

Les observations avec NIKA2 sont soumises à d'autres types de contaminants. Les *glitches* $G_k(t)$ sont des événements ponctuels dus à l'impact des rayonnements cosmiques sur les détecteurs. Ceux-ci sont facilement identifiés dans les données en temps, dans lesquelles ils se manifestent comme des pics de Dirac isolés correspondant à des grandes valeurs de flux.

Les KID sont également soumis à un bruit supplémentaire provenant des vibrations du système cryogénique, $C_k(t)$. Celles-ci ont une signature spectrale très spécifique, se manifestant par une oscillation sinusoïdale dans les données en temps. Leur contribution au TOI est mesurée dans les transformées de Fourier de ces dernières, dans lesquelles elles apparaissent comme des pics de Dirac isolés. Elles peuvent donc être soustraite des données.

Ces deux types d'évènement, de nature très différente, ont des contributions aux données en temps des détecteurs également différentes. Cependant, ces contributions peuvent être identifiées par recherche de pics dans les TOI ou dans leur transformée de Fourier. Elles peuvent ensuite en être soustraite avant le début de la procédure d'estimation du bruit corrélé.

Bruit intrinsèque du détecteur : $N_k(t)$

La dernière composante de l'équation (4.1) est le bruit intrinsèque des détecteurs $N_k(t)$. Ce bruit est dominé par les brisures spontanées de paires de Cooper et le bruit de photons. Il est par nature aléatoire et propre à chaque détecteur, et n'est donc pas corrélé entre les différents KID de NIKA2. Par conséquent, sa contribution aux données en temps ne peut pas être évaluée, ni donc être soustraite : il s'agit donc de la composante limitant la sensibilité de l'instrument. Ce bruit n'étant pas corrélé entre les différents détecteurs, sa contribution aux cartes du ciel construites à partir des observations NIKA2 est bien décrite par un bruit blanc non-corrélé spatialement. Par conséquent, une procédure de soustraction des contaminants parfaite peut aboutir à une carte du ciel dans laquelle le bruit est entièrement décorrélé, justifiant les hypothèses de bruit blanc dans les cartes qui seront discutées par la suite.

4.2 Soustraction du bruit corrélé

Comme nous l'avons vu dans la section précédente, les données en temps enregistrées par les détecteurs sont la somme du signal astrophysique et de différents types de bruits. Nous avons vu comment le bruit décorrélé lié aux rayonnements cosmiques et au fonctionnement du cryostat pouvait être retiré des données, grâce à la connaissance des propriétés de ces contributions. Néanmoins, les contributions les plus importantes de bruit, c'est-à-dire l'émission atmosphérique et le bruit électronique, restent présentes dans les données. Celles-ci engendrent des variations de brillance de surface de plusieurs Jy au cours d'un *scan*. À titre de comparaison, la brillance de surface au pic d'un amas de galaxies du grand programme SZ de NIKA2 est de l'ordre du mJy, soit plusieurs ordres de grandeurs plus faible que le bruit. Il est donc crucial de supprimer ces contributions dans les données afin de cartographier le signal d'intérêt. Cette procédure se nomme décorrélation, puisque son objectif est de soustraire les composantes de bruit corrélées entre les KID pour ne laisser que le bruit blanc intrinsèque à chaque détecteur.

4.2.1 Estimation de mode commun

Le niveau de bruit présent dans les données peut être estimé grâce à la corrélation entre les différents détecteurs. Celle-ci permet d'assimiler le bruit corrélé à un mode commun aux TOI des différents KID de NIKA2. Ce mode commun peut être calculé de plusieurs façons, dont la description progressive permet de comprendre le fonctionnement de l'algorithme de décorrélation et ses performances.

Mode commun par matrice de KID

Comme nous l'avons vu, le signal atmosphérique dans les TOI, qui représente la source majeure de bruit, est le même pour tous les KID d'une matrice de NIKA2. C'est précisément cette similarité qui peut être exploitée pour retirer la contamination atmosphérique des données. En effet, le signal d'intérêt, par construction, apparaît comme une composante différente d'un détecteur à un autre, puisque chaque KID voit à un instant donné un point différent du ciel. Ainsi, la stratégie d'observation par balayage du ciel fait que chaque KID passe sur la source à un instant différent ; le signal astrophysique apparaît donc dans les TOI décalé en temps. À l'inverse, on peut assimiler l'atmosphère à une source de bruit commune à chacun des détecteurs, et la calculer comme la médiane des TOI de chaque KID considéré². L'utilisation de la médiane permet ici de limiter le biais induit par le signal dans l'estimation de la contamination atmosphérique, qui sera discuté plus en détail en section 4.2.2.

On modélise donc le bruit corrélé dans les observations, dominé par l'atmosphère, comme un mode commun dans les TOI de tous les détecteurs d'une matrice. Afin de calculer ce mode commun, les TOI de chacun des N détecteurs considérés sont tout d'abord inter-étalonnées sur l'émission atmosphérique, estimée comme la médiane des TOI, en calculant pour chaque KID k un gain g_k d'après :

$$\text{TOI}_k(t, \nu) = g_k \times \text{Med}[\text{TOI}_{k'}(t, \nu)]_{k'}. \quad (4.5)$$

On peut alors calculer le mode commun aux détecteurs comme la moyenne des TOI de tous les KID, pondérée par les gains de ceux-ci :

$$\text{CM}(t, \nu) = \frac{1}{N} \sum_{k=1}^N \frac{1}{g_k} \text{TOI}_k(t, \nu). \quad (4.6)$$

Ce mode commun (CM) peut être soustrait aux données en temps de chaque KID comme un estimateur du bruit corrélé. Cependant, comme nous l'avons vu dans l'équation (4.1), chaque détecteur a une réponse différente aux différentes composantes de bruit corrélé. Afin de tenir compte de ces différentes réponses, on peut calculer un modèle de bruit comme une fonction affine du mode commun pour chaque KID k :

$$B_k(t, \nu) = a_k \text{CM}(t, \nu) + b_k, \quad (4.7)$$

où a_k et b_k sont calculés par régression linéaire pour chaque TOI.

En soustrayant cette estimateur de bruit propre à chaque détecteur à ses données en temps, on obtient des données nettoyées d'un mode commun qui, dans l'hypothèse où le bruit est entièrement donné par l'émission atmosphérique, sont dominées par le signal d'intérêt.

Cette procédure est représentée en figure 4.1. Celle-ci montre les données en temps de 20 KID de la matrice à 150 GHz de NIKA2 lors d'une observation d'Uranus. Le panneau gauche montre les données brutes, où on voit le signal astrophysique comme des pics décalés en temps, et le bruit atmosphérique

²On ne considère que les détecteurs valides ; voir section 3.3.2 (page 57).

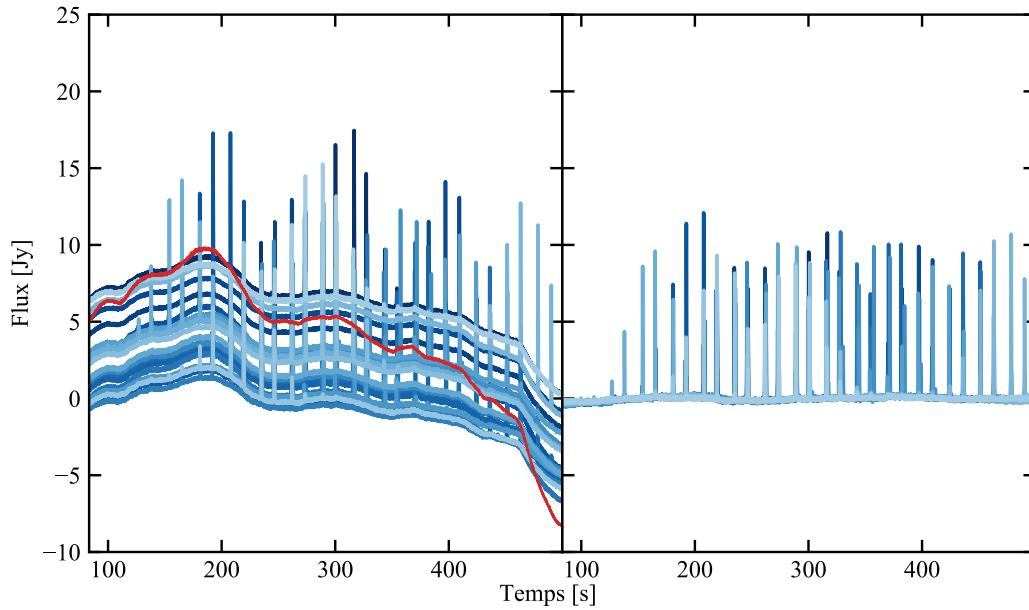


Figure 4.1 : **Gauche** : Données en temps pour 20 détecteurs de la matrice à 150 GHz de NIKA2 lors d'une observation d'Uranus (bleu). Le passage de chaque KID sur la planète est visible par les pics d'émission décalés en temps. Le mode commun estimé sur les TOI de la matrice entière est représenté en rouge.

comme le comportement à grande échelle de tous les détecteurs. Le mode commun estimé sur les TOI de toute la matrice de KID est représenté en rouge. Sur le panneau droit, on voit les TOI après soustraction de l'estimateur de bruit calculé à partir du mode commun (équation 4.7). On voit que le comportement à grande échelle du bruit est supprimé, et que seuls restent le signal astrophysique et un bruit mineur, visiblement décorrélaté entre les détecteurs.

Mode commun par boîte électronique

Pour la décorrélation illustrée en figure 4.1, nous avons émis l'hypothèse que le bruit corrélé pouvait être assimilé à un mode commun à tous les KID d'une matrice. Cette hypothèse repose sur le fait que l'atmosphère est le contaminant dominant dans les données en temps des KID, et que celle-ci est la même pour tous les détecteurs. Cependant, comme le montre l'équation (4.1), une autre composante de bruit corrélé est présente dans les données : le bruit dû à l'électronique de lecture. Celui-ci, par construction, ne présente pas de corrélation entre tous les détecteurs de la matrice, mais seulement entre ceux lus par une même boîte électronique (voir section 3.2.1, page 52).

Ainsi, à l'issue de la soustraction d'un mode commun atmosphérique, il reste dans les données une composante de bruit corrélé électronique. Les corrélations entre les KID lors de l'observation d'Uranus présentée en figure 4.1 sont illustrées en figure 4.2. Sur la gauche sont représentées les corrélations avant soustraction d'un mode commun : tous les détecteurs sont très corrélés entre eux, avec des coefficients de corrélation supérieurs à 99%. Sur la droite, on voit les corrélations résiduelles entre les KID après soustraction d'un mode commun calculé sur les TOI de toute la matrice à 150 GHz de NIKA2. On voit que, si les corrélations sont fortement diminuées – atteignant des valeurs absolues inférieures à 20% –, des structures persistent. En particulier, on remarque deux types de corrélations résiduelles. Les carrés qui apparaissent corrélés, délimités par des lignes pointillées noires, correspondent aux

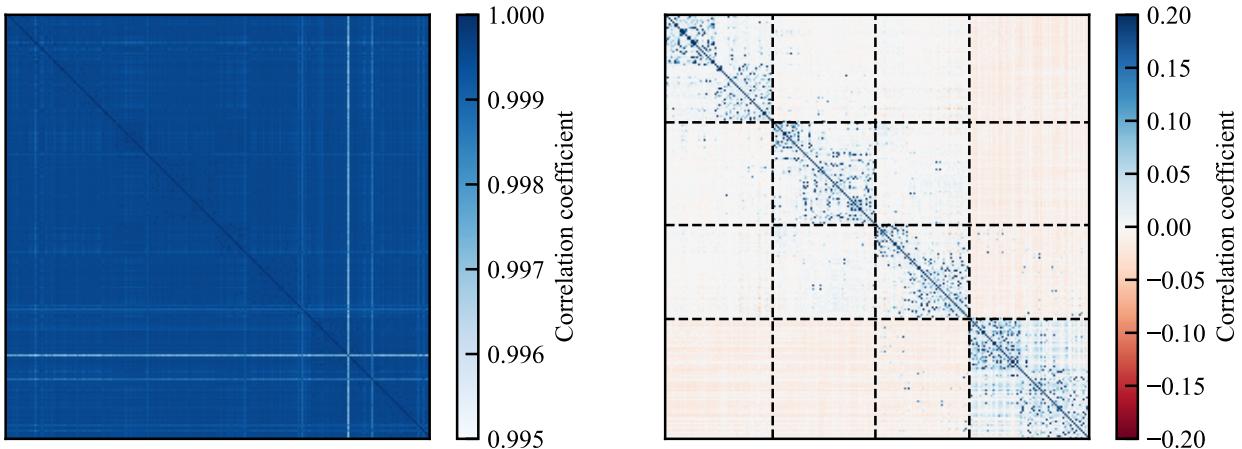


Figure 4.2 : Matrice de corrélation entre les TOI des détecteurs de la matrice NIKA2 à 150 GHz lors d'une observation d'Uranus. **Gauche** : Corrélations des données brutes, avant soustraction d'un mode commun. **Droite** : Corrélations après soustraction d'un mode commun calculé à partir de tous les détecteurs de la matrice. Les lignes noires pointillées délimitent les détecteurs appartenant à chacune des quatre boîtes de lecture de NIKA2.

KID lus par une même boîte électronique. Les structures carrées plus petites correspondent quant à elles aux détecteurs couplés à la même ligne de lecture. Ces corrélations montrent l'imperfection de la décorrélation par soustraction d'un seul mode commun pour tous les KID d'une matrice.

Le bruit électronique est en général plus faible que le bruit atmosphérique, bien que les deux composantes puissent avoir une contribution comparable dans le cas d'observations avec d'excellentes conditions météorologiques. Dans le cas des observations d'Uranus utilisées à titre illustratif en figures 4.1 et 4.2, le bruit électronique est plusieurs ordres de grandeur plus faible que le signal de la planète, d'une dizaine de Jy. Cependant, pour les observations de l'effet SZ, donc la brillance de surface est au maximum de l'ordre du mJy, le bruit électronique n'est plus négligeable. Il est donc important, au même titre que pour l'atmosphère, de le retirer des TOI.

Pour ce faire, la procédure d'estimation d'un mode commun détaillée précédemment peut être utilisée, avec une modification notable. Plutôt que de calculer un mode commun pour tous les détecteurs d'une matrice, la même opération peut être réalisée par groupe de KID, en particulier en les regroupant selon la boîte électronique par laquelle ils sont lus. Les modes communs résultants contiennent donc à la fois le bruit atmosphérique et le bruit électronique. D'après l'équation (4.1), les composantes des TOI restant après soustraction de ce mode commun ne contiennent donc *a priori* plus que le signal d'intérêt et le bruit intrinsèque aux détecteurs.

Mode commun par groupe de détecteurs corrélés

La décorrélation par ligne de lecture permet donc de tenir compte de deux corrélations différentes entre les KID : celle entre tous les détecteurs d'une matrice, engendrée par le bruit atmosphérique, et celle entre les détecteurs lus par une même électronique de lecture. D'autres corrélations peuvent être présentes dans les données, mais celles-ci sont difficiles à quantifier : on parle alors de corrélations résiduelles. On voit notamment dans le panneau droit de la figure 4.2 des corrélations entre KID couplés à une même ligne de lecture. Celles-ci se manifestent par des blocs carrés dans la matrice de corrélation des détecteurs, environ deux fois plus petits que les blocs de KID associés à la même boîte électronique.

Afin d'éliminer au mieux ces corrections résiduelles, on peut procéder à un calcul de mode commun entre détecteurs corrélés. Pour cela, la matrice de corrélation entre les TOI de tous les détecteurs est calculée. Les détecteurs sont ensuite groupés par bloc de détecteurs les plus corrélés, indépendamment de toute autre considération (par exemple, de la ligne de lecture à laquelle ils sont couplés). Un mode commun est calculé pour chacun de ces groupes de KID corrélés, et soustrait de la même façon que pour les méthodes à mode commun par matrice ou par boîte électronique.

Cette méthode est à l'heure actuelle la méthode privilégiée au sein de la collaboration NIKA2, et en particulier pour les observations du grand programme SZ, pour deux raisons principales. Tout d'abord, comme l'a montré la progression au cours de cette section, elle représente l'aboutissement de la réflexion sur l'exploitation des corrélations entre détecteurs pour retirer le bruit corrélé des données en temps. De plus, au cours de nombreux tests de différentes méthodes sur des simulations comme sur des données réelles, c'est celle qui a livré les meilleures performances, à la fois en termes de bruit résiduel et de filtrage du signal. Le filtrage et le bruit résiduel, utilisés à la fois comme métriques pour juger de la qualité du processus de décorrélation et comme outils pour tenir compte des caractéristiques de ce processus lors de l'analyse des cartes, seront discutés en section 4.4 (page 89). Cette méthode de décorrélation est nommée *common mode one block* dans le pipeline d'analyse des données de la collaboration NIKA2, et décrite dans [11] sous le nom de *most correlated pixels*.

4.2.2 Masquage de source et analyse itérative

L'estimation du bruit par calcul d'un mode commun permet donc d'obtenir des données en temps dominées par le signal pour chaque détecteur. Cependant, il faut noter que cette opération est imparfaite : le mode commun des TOI est affecté par la présence de signal dans ces dernières. Si cet effet est atténué par l'utilisation de la médiane des TOI dans l'estimation des modes communs, il n'est pas complètement annulé. En particulier, si une proportion significative des détecteurs utilisés dans l'estimation du mode commun passe sur la source au même instant, le mode commun à cet instant contiendra une partie du signal d'intérêt, qui sera soustraite au données et donc perdue dans les cartes. Le mode commun est donc biaisé par la présence de signal. Ce biais est d'autant plus important que la source est étendue, puisque la probabilité que deux détecteurs la voient en même temps s'en retrouve augmentée ; il est donc particulièrement important dans le cas de l'effet SZ, et *a priori* peu présent pour les sources ponctuelles. De plus, la méthode d'estimation du mode commun par bloc de détecteurs corrélés augmente par construction l'ampleur de ce biais. En effet, si la corrélation des TOI pour construire les blocs de données en temps est calculée sans faire la différence entre signal et bruit, des détecteurs voyant simultanément un signal similaire seront corrélés non seulement à cause du bruit, mais également à cause du signal. Ils auront donc une plus grande corrélation que deux détecteurs voyant la source à des instants différents, et seront groupés pour l'estimation du mode commun.

C'est pourquoi il est important de traiter différemment le signal du bruit au cours de l'estimation du mode commun. La méthode employée est l'utilisation d'un masque, permettant de ne pas tenir compte des régions contenant du signal dans l'estimation du mode commun. Ce masque est défini comme la région du ciel contenant du signal astrophysique :

$$M(x, y) = \begin{cases} 0 & \text{si du signal est attendu aux coordonnées } (x, y), \\ 1 & \text{sinon.} \end{cases} \quad (4.8)$$

En combinant ce masque avec la matrice de pointage de chaque KID, définie par l'équation (4.3), on

peut calculer un masque $M_k(t)$ pour les données en temps de chaque KID,

$$M_k(t) = P_k(t, x, y) \times M(x, y) = \begin{cases} 0 & \text{si du signal est attendu pour le KID } k \text{ au temps } t, \\ 1 & \text{sinon.} \end{cases} \quad (4.9)$$

Le calcul de mode commun décrit précédemment peut donc être réalisé en considérant uniquement les régions des TOI dans lesquelles le masque est non-nul. Cela permet d'éviter au mode commun d'être biaisé par le signal d'intérêt, puisque les régions contenant du signal ne seront pas considérées dans le calcul du mode commun.

Le calcul du masque requiert donc une connaissance préalable de la région contenant du signal d'intérêt. Cette opération est délicate, puisque la distinction des contributions du signal et du bruit est l'objectif même de la procédure de décorrélation. On ne connaît donc pas la forme du signal au début de la procédure de décorrélation. Dans le cadre des observations d'amas de galaxies avec NIKA2, le masque est défini de façon itérative. Une première estimation du masque est réalisée *a priori* en considérant un disque, centré sur les coordonnées de l'amas, et de rayon suffisamment grand pour inclure tout le signal significatif provenant de l'amas. La décorrélation par calcul d'un mode commun par blocs de détecteurs corrélés est réalisée dans un premier temps en utilisant ce masque, donnant une première estimation de la distribution du signal astrophysique d'intérêt dans les données en temps de chaque détecteur. Cette procédure est appliquée pour tous les *scans* réalisés pour l'observation d'un amas. Les données en temps obtenues, dominées par le signal d'intérêt, sont projetées en cartes du ciel correspondant à chaque *scan*, qui sont ensuite coadditionnées pour obtenir une unique carte de la région de l'amas. Les procédures de projection et de coaddition seront détaillées en section 4.3 (page 86). La carte obtenue est utilisée pour déterminer une nouvelle réalisation du masque. Pour cela, une carte de signal-sur-bruit est calculée comme le rapport entre les cartes de signal astrophysique et de bruit obtenues par la procédure de coaddition. La région du ciel dans laquelle on détecte un signal astrophysique significatif est utilisée pour mettre à jour le masque $M(x, y)$:

$$M(x, y) = \begin{cases} 0 & \text{si le signal-sur-bruit est supérieur à } 3\sigma \text{ aux coordonnées } (x, y), \\ 1 & \text{sinon.} \end{cases} \quad (4.10)$$

La décorrélation est alors répétée avec ce nouveau masque, qui représente une estimation plus fiable de la distribution de signal astrophysique dans la région cartographiée que le masque initial. La procédure est alors répétée plusieurs fois, en affinant à chaque itération la connaissance du masque grâce à l'itération précédente. La décorrélation converge alors vers une carte finale, et après quelques (~ 5) itérations, le masque n'évolue plus significativement.

Une décorrélation itérative est illustrée en figure 4.3 pour l'amas de galaxies PSZ2-G183.90+42.99. On voit en haut de la figure, de gauche à droite, l'évolution de la carte de l'amas au fil des itérations. À partir de la troisième («Iter 2»), la carte est stable, et n'évolue plus significativement. Le panneau central montre la différence entre les cartes de deux itérations successives. On remarque en effet que cette différence converge vers zéro, avec des différences inférieures à 0.1 mJy/beam entre la dernière itération et l'avant-dernière. Enfin, le panneau bas montre l'évolution du masque utilisé. On voit que celui-ci mime la morphologie circulaire de l'amas, et que sa forme converge, avec des différences de moins en moins importantes avec les itérations.

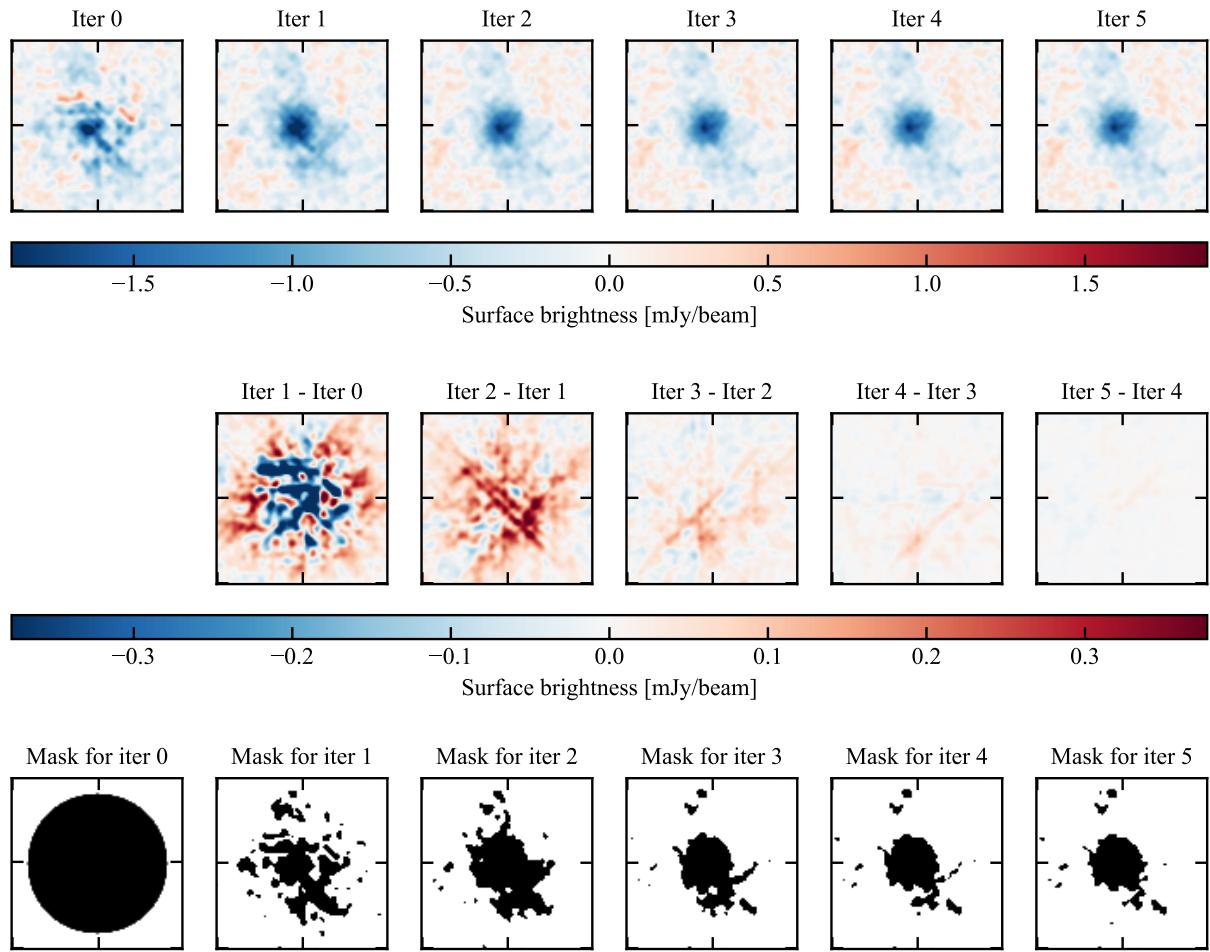


Figure 4.3 : Évolution (de gauche à droite) des résultats avec les itérations de la décorrélation pour les observations NIKA2 de l’amas PSZ2-G183.90+42.99. **Haut** : Carte de brillance de surface obtenue pour chaque itération. **Centre** Différence entre la carte de brillance de surface obtenue pour une itération et pour l’itération précédente. **Bas** : Masque utilisé pour l’évaluation du mode commun pour chaque itération.

4.3 Des données en temps aux cartes du ciel

Nous avons vu au cours des sections précédentes comment la procédure de décorrélation permettait d’obtenir des données en temps nettoyées du bruit corrélé, et dominées par le signal d’intérêt. Il reste donc à projeter ces données pour constituer des cartes du ciel tel qu’observé avec NIKA2.

4.3.1 Projection et cartographie du signal d’intérêt

La projection utilisée par le *pipeline* de la collaboration NIKA2 est une projection en *nearest grid point*. Pour cela, une carte de pixels carrés en coordonnées équatoriales et tout d’abord définie. La taille des pixels est un paramètre de l’analyse, qui peut être choisi. La taille optimale dépend à la fois de la vitesse de balayage du ciel au cours des observations et de la fréquence d’échantillonnage à laquelle les fréquences des KID sont lues. En effet, la combinaison de ces deux grandeurs fixe la distance sur le ciel séparant deux échantillons, c’est-à-dire la distance parcourue par le télescope entre deux mesures de fréquence de résonance. Pour les observations du grand programme SZ de NIKA2, la fréquence d’échantillonnage est $f = 24$ Hz, et la vitesse de déplacement du télescope de $v = 40$ arcsec/s. Le rapport v/f donne donc la distance angulaire séparant deux échantillons consécutifs, qui est de 1.67 arcsec.

La taille de pixels est donc choisie comme 3 arcsec, permettant à deux échantillons consécutifs de ne jamais être séparés de plus d'un pixel. Ce critère assure que chaque pixel de la carte soit utilisé pour la projection d'au moins un échantillon.

Une fois la grille de pixels définie, la matrice de pointage $P_k(t, x, y)$ associée à chaque KID est calculée. La position de chaque détecteur par rapport à un détecteur de référence est connue à chaque instant, comme nous l'avons vu en 3.3.2 (page 57). De plus, la position du détecteur de référence dans le ciel étant également connue grâce au pointage du télescope, comme nous l'avons vu en 3.3.3 (page 57). On connaît donc, pour chaque mesure de flux réalisée par un KID, la position dans le ciel à laquelle elle a eu lieu. On peut alors associer chaque mesure de flux au pixel de la carte le plus proche de la position à laquelle elle a été effectuée.

La carte de brillance de surface résultante $I(x, y)$ est obtenue en calculant la moyenne pondérée des mesures associées à chaque pixel :

$$I(x, y) = \frac{\sum_{k,t} P_k(t, x, y) w_k \text{TOI}'_k(t)}{\sum_{k,t} P_k(t, x, y) w_k}, \quad (4.11)$$

où $\text{TOI}'_k(t) = \text{TOI}_k(t) - B_k(t)$ représente les TOI des détecteurs après soustraction de l'estimation de bruit, c'est-à-dire après la décorrélation.

Le poids w_k associé à chaque détecteur k est calculé comme l'inverse de la variance de la TOI du détecteur en dehors du masque M_k (équation 4.9). On a donc :

$$w_k = \frac{1}{\text{Var}[\text{TOI}_k(t)]_{t \setminus M_k(t) \neq 0}}, \quad (4.12)$$

où $\text{Var}[\dots]_{t \setminus M_k(t) \neq 0}$ représente la variance sur les échantillons t hors du masque, c'est-à-dire la variance hors-source des TOI.

Cette pondération permet de donner moins d'importance aux détecteurs dont les données sont toujours fortement bruitées à l'issue de la décorrélation.

4.3.2 Cartes de couverture et de bruit

Afin de quantifier la couverture des différentes parties du ciel pendant un *scan*, ainsi que le bruit attendu dans les cartes projetées pour ce *scan*, on peut également estimer une carte de couverture et de variance du bruit. Ces deux cartes sont des sorties naturelles de la procédure de projection. La carte de couverture, aussi appelée carte de nombre de coups ou de *hits*, représente une cartographie du nombre d'échantillons utilisés pour la projection. En d'autre termes, il s'agit, pour chaque pixel de la carte, du nombre d'échantillons y ayant été associés par la projection en *nearest grid point*. Cette carte est obtenue en sommant les matrices de pointages de tous les KID pour tous les échantillons :

$$N_{\text{hits}}(x, y) = \sum_{k,t} P_k(t, x, y). \quad (4.13)$$

Pour un pixel donné, plus ce nombre est élevé, plus le nombre d'échantillons utilisés pour la projection est grand, diminuant ainsi l'incertitude statistique associée au signal en ce point de la carte.

La carte de variance est également un produit de la projection. En effet, nous avons vu que les échantillons de chaque détecteur étaient pondérés par l'inverse de la variance de sa TOI en dehors de la source. On peut alors définir une carte de poids, en projetant les poids de chaque détecteur de

la même façon que les TOI elles-mêmes sont projetées ; ou bien, de façon équivalente, projeter la variance associée à chaque détecteur. On obtient alors une carte, de pixelisation identique à la carte du signal, représentant la variance des détecteurs utilisés pour la projection du signal dans chaque pixel. La racine carrée de cette carte donne alors une carte d'écart-type ou de déviation standard, fournissant une estimation de l'amplitude des fluctuations de bruit, et donc de l'incertitude sur le signal astrophysique en tout point de la carte.

Des exemples de cartes de couverture et de déviation standard de bruit pour un *scan* réalisé au cours des observations de l'amas ACT-CL J0215.4+0030 dans le grand programme SZ sont présentées dans le panneau haut de la figure 4.4. On observe la forme rectangulaire des *scans*, caractéristique des observations du grand programme SZ (discutées au paragraphe suivant). On note également que la carte de couverture montre un nombre plus important de coups dans le centre de la carte qu'en périphérie. Cette inhomogénéité de couverture s'explique par le fait que plus de détecteurs passent dans les régions du centre de la carte, pour des raisons géométriques, alors que sa périphérie n'est visitée que par un plus faible nombre de KID. Pour les mêmes raisons, la déviation standard du bruit est plus élevée dans les régions externes de la carte, entraînant des observations de meilleure qualité au centre. Une analyse détaillée des propriétés de cet amas sera présentée au chapitre 6.

4.3.3 Coaddition de cartes

Tout comme la carte de brillance de surface associée à un *scan* est calculée par une moyenne pondérée des signaux reçus par les KID, la combinaison de plusieurs *scans* permet de construire une carte de brillance de surface en calculant la moyenne pondérée des cartes de ces derniers. La carte de brillance résultant de la combinaison de n *scans*, donc les cartes de brillance de surface individuelles $I_s(x, y)$ sont calculées par l'équation (4.11), est donnée par :

$$I_{\text{tot}}(x, y) = \frac{\sum_{s=1}^n w_s(x, y) I_s(x, y)}{\sum_{s=1}^n w_s(x, y)}, \quad (4.14)$$

où les poids $w_s(x, y)$ sont définis pour chaque carte à partir de la carte de couverture $N_{\text{hits}}(x, y)$ et de la variance de la carte de brillance de surface en dehors du masque de la source :

$$w_s(x, y) = \frac{N_{\text{hits},s}(x, y)}{V[I_s(x, y) \times \sqrt{N_{\text{hits},s}(x, y)}]_{(x,y) \setminus M(x,y) \neq 0}}. \quad (4.15)$$

Tout comme la pondération par la variance des TOI nous a permis de donner moins d'importance aux détecteurs bruités, cette pondération permet de donner moins d'importance dans la carte finale aux *scans* dans lesquels le bruit est toujours important après la décorrélation. Le terme $\sqrt{N_{\text{hits},s}(x, y)}$ au dénominateur permet de normaliser la carte au temps passé par pixel.

Nous avons vu au chapitre précédent que la stratégie de *scan* utilisée pour les observations du grand programme SZ de NIKA2 était basée sur une succession de groupes de quatre *scans*, réalisés à des angles de 0° , 45° , 90° , et 135° par rapport à l'axe des ascensions droites. La procédure de coaddition est donc particulièrement importante dans ce contexte. Le panneau bas de la figure 4.4 montre les cartes de couverture et de déviation standard du bruit obtenues par coaddition de quatre *scans* de l'amas ACT-CL J0215.4+0030, dont celui présenté dans le panneau haut de la figure. Les observations faites sur le panneau haut (voir section précédente) restent valides : le centre de la carte bénéficie d'une meilleure couverture, et donc d'une déviation standard de bruit plus faible. On voit également que la

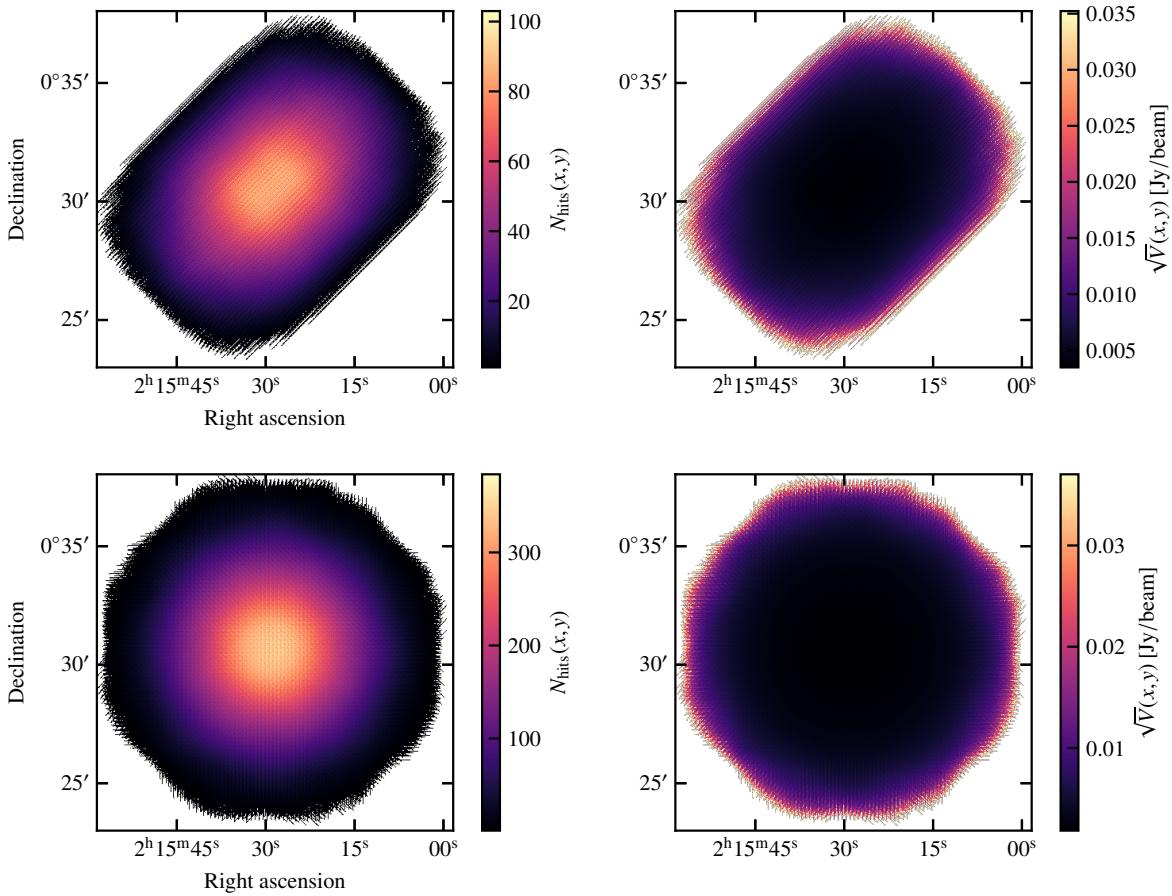


Figure 4.4 : **Haut :** carte de couverture (gauche) et de déviation standard du bruit (droite) à 150 GHz pour un *scan* du grand programme SZ de NIKA2. **Bas :** Mêmes cartes pour la combinaison de quatre *scans* dans quatre orientations différentes, incluant le *scan* du panneau haut.

combinaison de *scans* dans quatre angles différents a eu pour effet d'augmenter la couverture dans la carte (puisque'on considère quatre *scans* au lieu d'un seul), mais également de la rendre plus isotrope. De même, le bruit dans la carte coadditionnée est plus bas, et reste bas dans une grande région à symétrie circulaire. C'est là l'objectif de l'utilisation de *scans* dans plusieurs directions mise en place pour les observations du grand programme SZ de NIKA2, en plus de l'optimisation du filtrage, que nous discuterons en section 4.4.2 (page 91).

4.4 Performance de la décorrélation

Quelle que soit la méthode choisie pour la procédure de décorrélation, celle-ci est toujours imparfaite. Nous avons vu que la décorrélation avait deux objectifs, qui sont de retirer un maximum de bruit corrélé des données en temps sans porter atteinte au signal astrophysique d'intérêt. Par conséquent, on quantifie la performance de la procédure au travers de critères évaluant l'atteinte de ces deux objectifs : le bruit corrélé résiduel, et le filtrage du signal. Ces deux quantités seront ensuite utilisées au cours de l'analyse des données de NIKA2 afin de tenir compte des caractéristiques de la décorrélation.

4.4.1 Bruit corrélé résiduel

À l'issue de la procédure de décorrélation, une partie du bruit reste inévitablement présente dans les données en temps. Nous avons vu précédemment qu'une partie de ce bruit était le bruit blanc intrinsèque à chaque détecteur, ne pouvant par nature pas être soustrait des données en temps. En réalité, une partie du bruit corrélé dû à l'atmosphère et à l'électronique reste également présent. La présence de ce bruit corrélé résiduel est donc une indication de l'imperfection de la méthode de décorrélation. Puisque ce bruit reste dans les données en temps après la décorrélation, il est projeté dans les cartes de brillance de surface au même titre que le signal astrophysique. De plus, puisqu'il est corrélé, il peut faire apparaître des structures de bruit dans les cartes, pouvant être faussement interprétées comme du signal d'intérêt. Il est donc important de connaître le comportement de ce bruit en vue de la mesure des propriétés des amas de galaxies à partir des observations avec NIKA2.

L'information pertinente sur le bruit corrélé résiduel est présente dans son spectre de puissance. L'évaluation de celui-ci ne peut pas avoir lieu directement dans les données NNIKA2, puisque celles-ci contiennent également du signal d'intérêt. Cependant, il est possible d'utiliser les cartes de demi-différences entre les *scans*. Pour cela, on réalise la coaddition des n cartes des différents *scans* d'après :

$$I'(x, y) = \frac{\sum_{s=1}^n w_s(x, y) I_s(x, y) \times (-1)^s}{\sum_{s=1}^n w_s(x, y)}, \quad (4.16)$$

où les poids w_s sont les mêmes que pour la coaddition des *scans*, calculés par l'équation (4.15). Ainsi, la carte d'un *scan* sur deux se voit multipliée par (-1) dans la coaddition. Puisque le signal astrophysique est *a priori* la seule composante des cartes constante d'un *scan* à l'autre, il est complètement annulé³ dans la carte $I'(x, y)$. En revanche, le bruit corrélé pouvant indifféremment être positif ou négatif, sa structure n'est pas modifiée au cours de l'opération. On obtient ainsi une carte, vide de tout signal d'intérêt, contenant seulement du bruit dont les propriétés sont les mêmes que celles du bruit présent dans la carte de signal obtenue par coaddition.

Le spectre de puissance du bruit résiduel est calculé sur la carte de demi-différences, normalisée par la carte de déviation standard du bruit :

$$P_{\text{bruit}}(k) = \left| F \left(\frac{I'(x, y)}{\sqrt{V(x, y)}} \right) \right|^2, \quad (4.17)$$

où F représente l'opération de transformée de Fourier.

Cette normalisation permet de tenir compte de la plus grande amplitude de bruit attendue dans les bords des cartes, du fait de la couverture plus faible dans ces régions. La forme du spectre de puissance nous renseigne sur la corrélation du bruit résiduel. Si ce dernier est blanc, c'est-à-dire ne présente pas de corrélations résiduelles, son spectre de puissance doit être plat. À l'inverse, le spectre de puissance d'un bruit fortement corrélé à grande échelle se manifeste comme une fonction en loi de puissance, $P(k) \propto k^{-\alpha}$. Pour la décorrélation *common mode one block* réalisée sur les amas du grand programme SZ, c'est ce dernier comportement qui est observé. Nous verrons par exemple au chapitre 6 que le bruit résiduel dans les observations de l'amas ACT-CL J0215.4+0030 est très corrélé, comme en atteste le spectre de puissance présenté en figure 6.2.

³On note la nécessité d'un nombre pair de *scans* dans cette procédure, sans quoi le signal ne peut complètement s'annuler. De même, il est nécessaire d'ordonner les *scans* par angle lors de la coaddition, afin de ne pas systématiquement ajouter ou soustraire les cartes obtenues à un angle donné, ce qui crée des structures de bruit différentes.

C'est pourquoi la présence de bruit corrélé doit être prise en compte au cours des exploitations des observations d'amas de galaxies par NIKA2. Nous verrons au chapitre suivant que cette prise en compte est possible au travers de la matrice de covariance du bruit résiduel. Celle-ci peut être calculée à partir du spectre de puissance comme suit :

1. Un grand nombre ($N \sim 10^5$) de cartes de bruit blanc est généré. Dans chacune des cartes, chaque pixel est tiré aléatoirement dans une distribution gaussienne de moyenne 0 et de variance 1.
2. Chacune de ces cartes est convoluée par le spectre de puissance du bruit calculé par l'équation (4.17). Les cartes résultantes sont des réalisations de bruit normalisées suivant le spectre de puissance d'entrée.
3. Chacune des cartes est ensuite multipliée par la carte de déviation standard du bruit $\sqrt{V}(x, y)$. Les cartes résultantes sont des réalisations de bruit résiduel dont l'amplitude et les corrélations sont les mêmes que dans la carte de demi-différences. Chaque carte, de dimension $(n \times n)$, est déroulée en un vecteur de dimension $(n^2 \times 1)$.
4. La matrice de covariance des N vecteurs est calculée.

Pour une carte d'entrée de dimensions $(n \times n)$ pixels, le résultat est une matrice de taille $(n^2 \times n^2)$, dans laquelle chaque élément (i, j) représente la covariance du bruit dans les pixels i et j de la carte. Nous verrons au chapitre suivant que c'est cette matrice qui est utilisée dans la fonction de vraisemblance multivariée de l'ajustement des données NIKA2 par un modèle de propriétés physiques du milieu intra-amas, afin de tenir compte de la corrélation du bruit résiduel. Elle sera utilisée au travers d'une fonction de vraisemblance gaussienne multivariée, discutée en section 5.2.1 (page 106).

La procédure d'estimation du bruit résiduel et de sa matrice de covariance est présentée en figure 4.5. Le cas choisi pour l'illustration est la décorrélation *common mode one block* des observations de l'amas PSZ2-G183.90+42.99, dont le processus itératif est représenté en figure 4.3. Le spectre de puissance du bruit résiduel (haut droit de la figure) n'est pas plat, indiquant un bruit résiduel corrélé à grande échelle. Par conséquent, la matrice de covariance du bruit (bas gauche) n'est pas diagonale. Dans le cas où le bruit résiduel est proche d'un bruit blanc, les corrélations du bruit entre deux pixels différents de la carte sont faibles, et les éléments extra-diagonaux de la matrice de covariance sont négligeables. Ce comportement traduirait une décorrélation parfaite. De plus, comme nous le verrons au chapitre suivant, ce comportement est idéal, car il permet de ne pas tenir compte de la matrice de covariance du bruit lors de l'ajustement des propriétés des amas à partir des observations NIKA2, ce qui accélère grandement le processus.

4.4.2 Filtrage du signal

Si la décorrélation ne permet pas de retirer complètement le bruit corrélé des données en temps, elle est aussi imparfaite du point de vue de la conservation du signal. En effet, nous avons vu en section 4.2.2 que le masque utilisé pour éviter le biais dû au signal dans l'estimation du mode commun était défini comme la région du ciel dans laquelle le rapport signal sur bruit était supérieur à 3σ . Par construction, tout signal moins significatif n'est pas protégé par ce masque, et peut être inclus dans le mode commun. Il est alors soustrait comme du bruit, ce qui induit un filtrage du signal d'intérêt. Par ailleurs, nous avons vu que le masque était défini de façon itérative, en commençant avec un masque circulaire assez grand pour englober le signal. Il ne s'agit toutefois pas simplement de définir

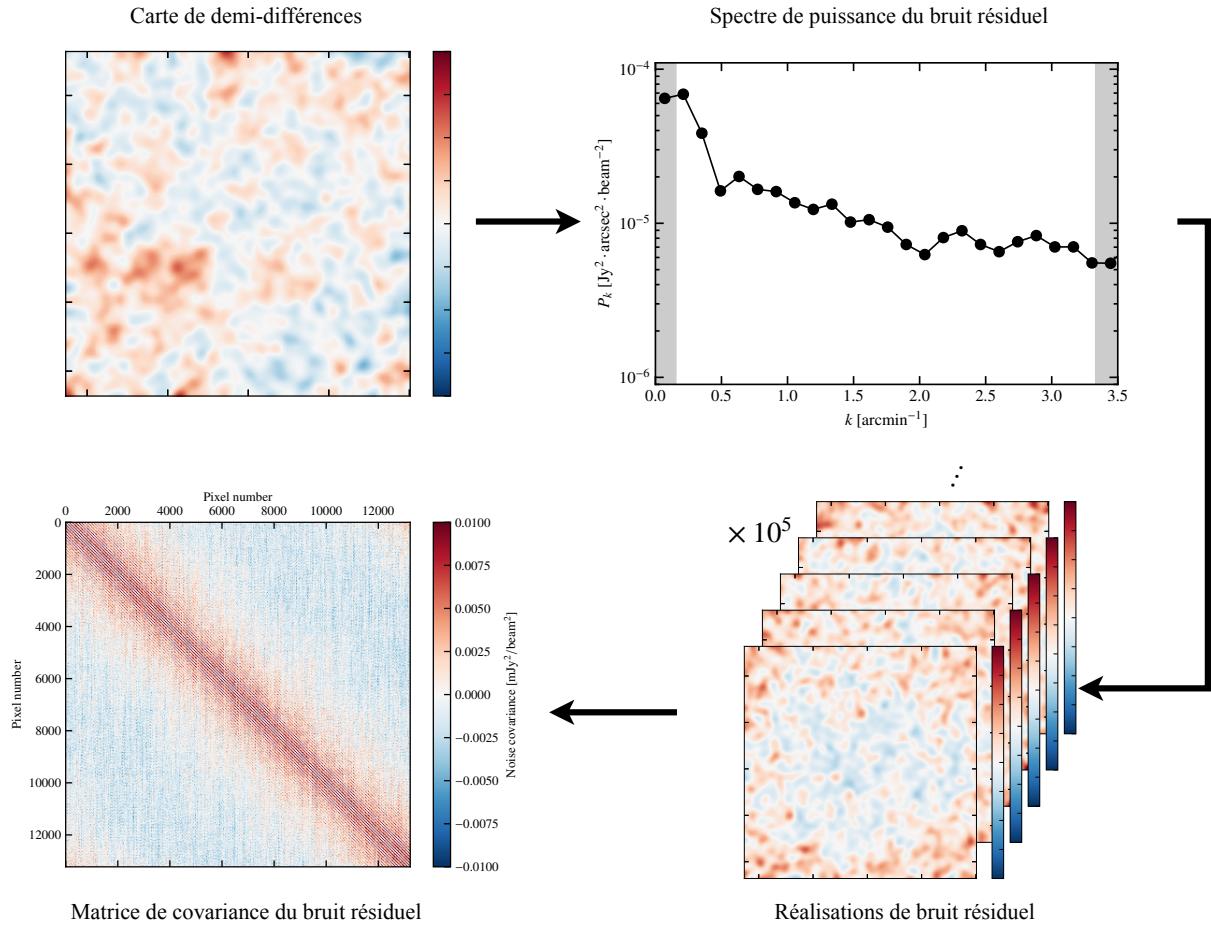


Figure 4.5 : Illustration du calcul de la matrice de covariance du bruit résiduel pour la décorrélation *common mode one block* de l’amas PSZ2-G183.90+42.99 (voir figure 4.3). La carte de demi-différences (*haut gauche*) est utilisée pour calculer le spectre de puissance du bruit résiduel (*haut droit*), à son tour utilisé pour générer des réalisations de bruit corrélé de même spectre de puissance (*bas droit*). La covariance de ces réalisations donne la matrice de covariance pixel-à-pixel du bruit résiduel (*bas gauche*).

un masque le plus grand possible pour masquer une grande portion de la carte, puisqu’une couverture du ciel minime entraînera un mode commun estimé sur peu de détecteurs, et donc potentiellement une mauvaise estimation de bruit. Ces défauts sont donc d’autant plus importants que la source observée est spatialement étendue. Les amas de galaxies observés dans le cadre du grand programme SZ de NIKA2 étant grands devant le lobe de l’instrument (voir section 3.4), ils sont affectés par ce filtrage. Il est donc nécessaire de connaître les propriétés de ce filtrage, afin de pouvoir faire le lien entre le signal SZ réel associé à l’amas et le signal détecté, faute de quoi les propriétés physiques mesurées seront biaisées.

Ce filtrage est estimé par la fonction de transfert de l’analyse. Celle-ci est évaluée par la mesure du filtrage subi par un signal connu lorsque celui-ci est soumis à la procédure de décorrélation. Pour l’analyse des données associées à un amas, une fois la décorrélation terminée, une carte de l’effet SZ dû à l’amas est simulée par intégration d’un profil de pression universel le long de la ligne de visée. Cette carte est projetée sur la même grille pixélisée que la carte de brillance de surface NIKA2. Une réalisation de bruit blanc, d’amplitude largement inférieure au signal SZ simulé, est ajoutée à la carte afin d’éviter la présence de zéros dans la carte et dans son spectre de puissance. L’importante de ce bruit sera discutée par la suite. La carte simulée obtenue S_{in} est ensuite convertie en données en temps

pour chaque détecteur grâce à la matrice de pointage de ces derniers. Les données en temps ainsi obtenues, entièrement dominées par le signal simulé, sont ajoutées aux données réelles. Cette somme forme donc des TOI pour chaque détecteur contenant les différents contaminants caractéristiques des observations avec NIKA2, le signal astrophysique d'intérêt, et le signal simulé.

Les données en temps ainsi obtenues sont soumises à la même procédure de décorrélation que les données NIKA2 ne contenant pas le signal simulé. La procédure itérative n'est pas nécessaire, puisqu'elle vise seulement à définir le masque le plus adapté à la protection du signal d'intérêt. Ainsi, la procédure de décorrélation est réalisée sur les TOI contenant le signal simulé en considérant le masque utilisé pour la dernière itération de la décorrélation des données brutes. De cette façon, le traitement subi par les données contenant le signal simulé est identique à celui subi par les données brutes. On peut donc supposer que le filtrage subi par le signal simulé est identique à celui subi par le signal astrophysique d'intérêt.

La carte de signal simulé reconstruit $S_{\text{out.}}$ est ensuite calculé comme la différence entre la carte issue de la décorrélation des données en temps contenant le signal simulé et celle issue de la décorrélation des données brutes (ne contenant pas de signal simulé). La fonction de transfert de l'analyse $\text{TF}(k)$, quantifiant le filtrage subi par le signal simulé, est calculée comme le rapport entre les spectres de puissance des cartes de signal simulé avant et après la décorrélation :

$$\text{TF}(k) = \frac{|F(S_{\text{out.}})|^2}{|F(S_{\text{in.}})|^2} \quad (4.18)$$

Puisque le processus subi par le signal simulé est identique à celui subi par le signal astrophysique d'intérêt, cette fonction de transfert représente un bon estimateur du filtrage subi par ce dernier au cours de la décorrélation ne contenant pas de signal simulé. L'intérêt de l'ajout de bruit blanc dans la carte de signal simulé $S_{\text{in.}}$ apparaît clairement dans l'équation (4.18). Si cette étape n'était pas réalisée, il existe un risque que le spectre de puissance de la carte présente des zéros. Ce dernier étant au dénominateur de l'équation (4.18), la fonction de transfert peut diverger.

Le processus d'estimation de la fonction de transfert pour la décorrélation des observations de l'amas PSZ2-G183.90+42.99 est schématisé en figure 4.6. La fonction de transfert résultante, représentée à droite de la figure, est caractéristique de la décorrélation *common mode one block* des observations NIKA2 d'amas de galaxies. Elle est caractérisée par un filtrage des grandes échelles angulaires ($k < 0.5 \text{ arcmin}^{-1}$), résultant en une suppression quasi-totale du signal aux échelles plus grandes que le champ de vue instantané de NIKA2 (6.5 arcmin), représentées par la région grisée à gauche de la figure. En revanche, les échelles angulaires plus petites sont bien conservées, avec un filtrage proche de 1. Nous verrons dans le chapitre suivant que cette fonction de transfert est utilisée lors de la mesure des propriétés thermodynamiques du milieu intra-amas à partir des observations NIKA2 des amas du grand programme SZ, afin de tenir compte du filtrage induit par la décorrélation.

4.5 Contamination par des sources ponctuelles

Nous avons vu au chapitre 3 que l'effet SZ pouvait être mesuré avec NIKA2 par le décrément en brillance de surface dans la carte à 150 GHz. Ce signal peut être affecté par des sources ponctuelles, de flux positif qui compense partiellement ou totalement le décrément. Afin d'obtenir des contraintes non-biaisées sur le signal SZ d'un amas – et donc sur ses propriétés thermodynamiques – il est donc nécessaire de traiter cette contamination en estimant le flux des sources. Celui-ci ne peut par construction

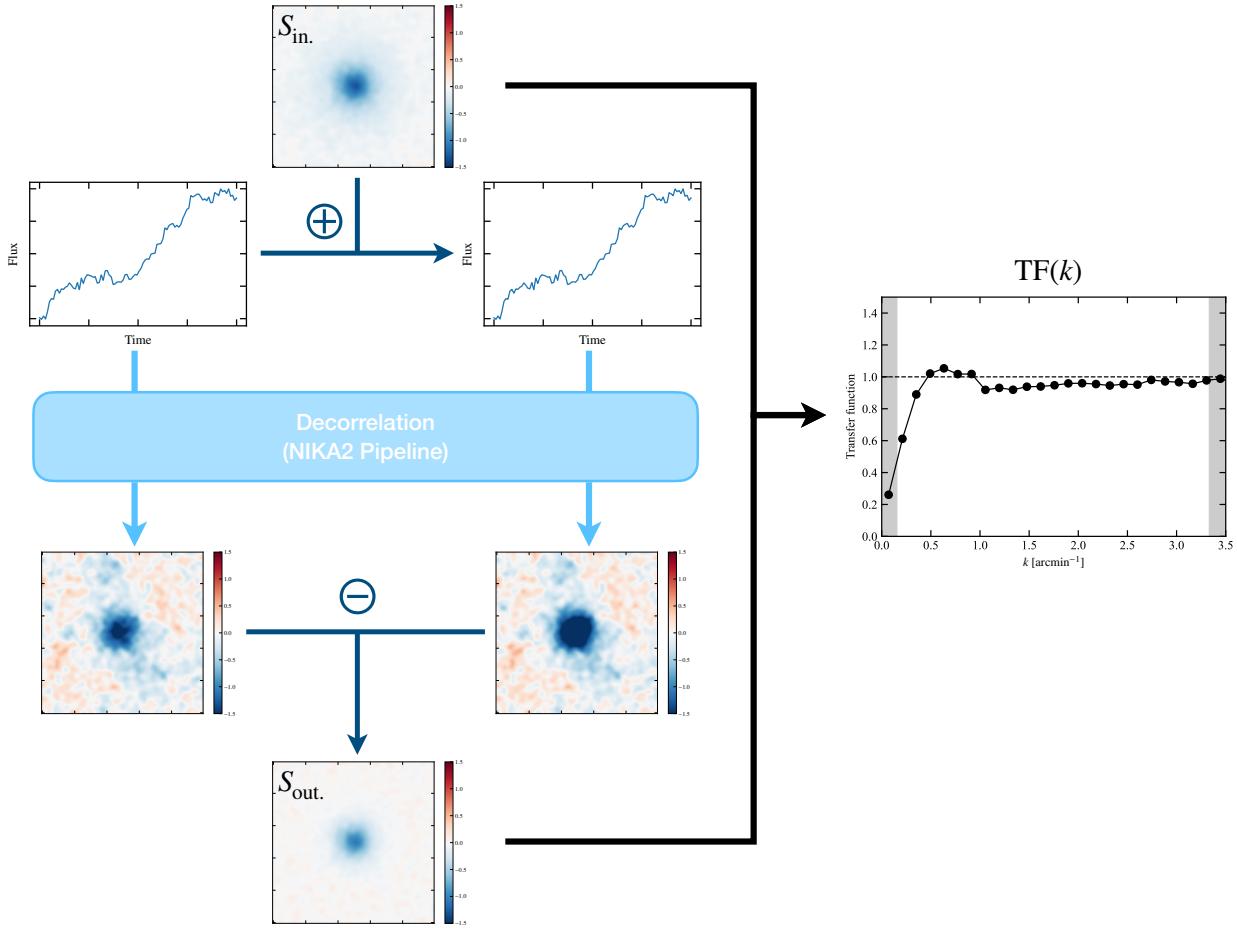


Figure 4.6 : Illustration du calcul de la fonction de transfert pour la décorrélation *common mode one block* de l’amas PSZ2-G183.90+42.99 (voir figure 4.3). Le rapport entre les spectres de puissance d’une carte simulée d’entrée ($S_{\text{in.}}$, haut) et de la carte après décorrélation ($S_{\text{out.}}$, bas) donne la fonction de transfert de la décorrélation.

pas être mesuré directement dans la carte NIKA2 à 150 GHz, puisqu’il est dégénéré avec le décrément SZ inconnu. Il est donc nécessaire de s’appuyer sur des données externes.

Un estimateur du flux d’une source est donné par l’évolution de son émission avec la fréquence d’observation, qui est quantifiée par son spectre (SED, pour *Spectral Energy Distribution*). Le logiciel PSTools a été développé dans le cadre de cette thèse, et permet d’ajuster ce flux pour des sources sub-millimétriques à partir d’un catalogue de sources et des données NIKA2. Ce logiciel a été documenté [153] et mis à disposition de la collaboration NIKA2, devenant l’outil principal d’estimation de la contamination des cartes SZ obtenues avec NIKA2 par des sources sub-millimétriques. Cette section décrit la procédure employée et implémentée dans PSTools ; son utilisation dans le cadre de l’analyse des propriétés de l’amas ACT-CL J0215.4+0030 sera présentée au chapitre 6.

4.5.1 Données d’entrée

Les sources ponctuelles contaminant le signal SZ peuvent être des sources d’avant-plan, d’arrière-plan, ou même des galaxies membres de l’amas. Dans tous les cas, ces sources peuvent être classifiées en deux catégories. D’une part, elles peuvent être des galaxies poussiéreuses émettant dans le domaine sub-millimétrique (SMG pour *submillimeter galaxies*), avec un flux croissant avec la fréquence dans le

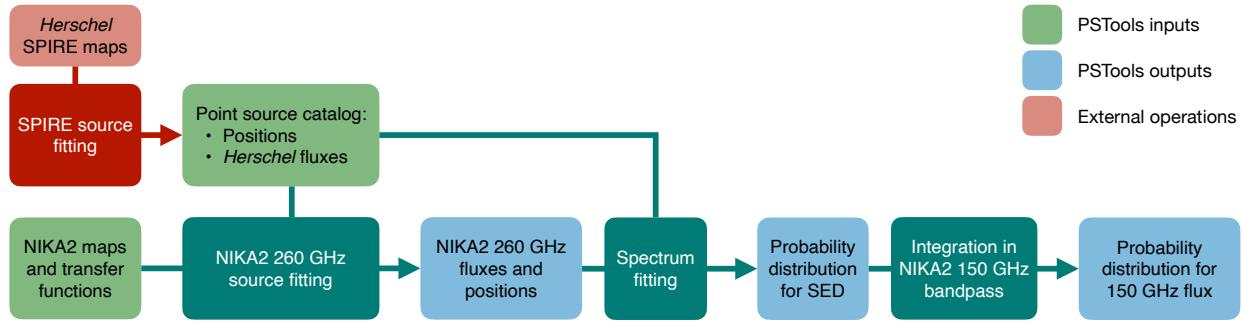


Figure 4.7 : Schéma illustrant le fonctionnement de `PSTools`. Les données d'entrée sont représentées en vert, et les données de sortie en bleu. La partie rouge représente les opérations externes au logiciel, c'est-à-dire la détection des sources dans les cartes SPIRE.

domaine couvert par les observations NIKA2. Elles peuvent également être des sources radio, présentant une émission plus forte à basse fréquence⁴. Les SMG sont plus simples à identifier avec NIKA2, puisque leur flux est plus élevé dans la bande à 260 GHz, où le signal SZ n'est en général pas détectable.

Afin de mesurer le spectre d'une source, la connaissance de son flux dans une large gamme de fréquences est nécessaire. `PSTools` permet de traiter la contamination due à des sources submillimétriques détectées par l'instrument SPIRE à bord du satellite *Herschel* [154]. Nous utilisons la combinaison des mesures de flux dans chacune des bandes de SPIRE (250, 350, et 500 μm , soit 1200, 860, et 600 GHz, respectivement) et dans la bande à 260 GHz de NIKA2, dans laquelle le signal SZ est négligeable devant le bruit et le flux des sources ponctuelles. Il est également possible d'utiliser les flux des sources à plus haute fréquence dans les bandes passantes de l'instrument PACS à bord de *Herschel* [155], lorsque ceux-ci sont publics. Par la suite, nous considérerons le cas où seules des données SPIRE sont disponibles, cas bien plus fréquent pour les amas du grand programme SZ.

Un schéma du fonctionnement de `PSTools` est présenté en figure 4.7. Les données d'entrée sont donc un catalogue de sources avec leurs positions et flux dans les trois bandes de SPIRE d'une part, et les données NIKA2 à 260 GHz d'autre part, incluant la carte de signal, de bruit, et la fonction de transfert associée. Dans la suite, nous présentons les trois grandes étapes de l'analyse, représentées en bleu-vert sur la figure 4.7, et les illustrons pour des sources considérées dans l'analyse de l'amas ACT-CL J0215.4+0030, présentée au chapitre 6.

4.5.2 Mesures de coordonnées et flux à 260 GHz

Le flux de chaque source à 260 GHz et sa position sont mesurés dans la carte NIKA2 à cette fréquence. Pour cela, un modèle de lobe de NIKA2 est ajusté dans une portion carrée de la carte, centrée sur la position de détection de chaque source dans SPIRE. L'évolution de la mesure du flux des sources avec la taille de cette portion a été étudiée, montrant de meilleurs résultats pour des tailles d'environ une minute d'arc. Le lobe est modélisé par une Gaussienne de largeur à mi-hauteur fixée à $\text{FWHM} = 12.5''$, en accord avec le modèle utilisé pour l'étalonnage des cartes [11], avec une amplitude et une position libres, ainsi qu'un éventuel niveau zéro. L'ajustement est réalisé à l'aide de la librairie Python `iminuit` [156], et l'amplitude de la Gaussienne ajustée livre une estimation du flux de la

⁴Notons que cette distinction entre galaxie poussiéreuse et source radio est une simplification, le spectre d'émission de la poussière présentant une composante d'émission radio.

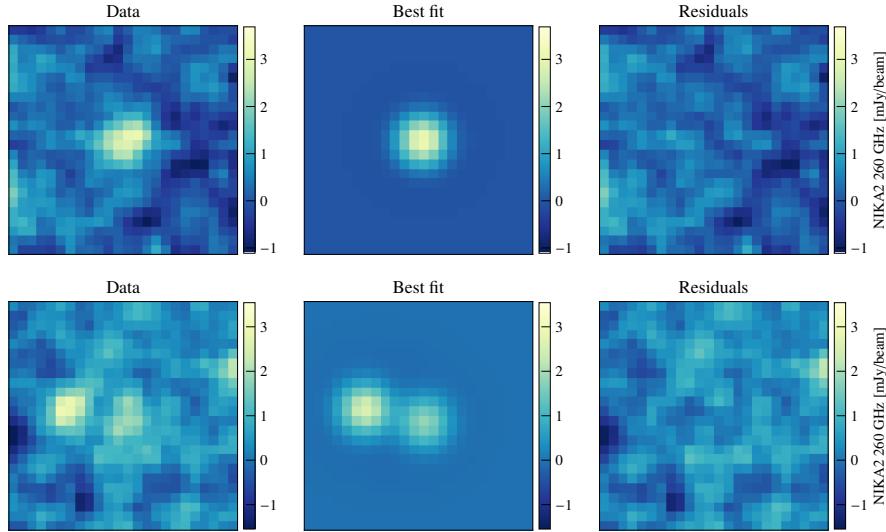


Figure 4.8 : Résultats d'un ajustement de source à 260 GHz pour une source simple (*haut*) et pour deux sources proches traitées comme double (*bas*) dans le champ de l'amas ACT-CL J0215.4+0030. De gauche à droite, les cartes représentent les données NIKA2 à 260 GHz, le modèle correspondant au meilleur ajustement, et la différence entre les deux cartes.

source à 260 GHz. Lorsque deux sources sont proches, elles sont ajustées simultanément dans la même portion de carte, afin d'éviter une surestimation des flux. Un exemple d'ajustement est présenté en figure 4.8 pour une source simple et un ensemble de deux sources proches, montrant les données, le meilleur ajustement, et la différence entre les deux cartes. On voit que cette différence ne présente pas de structure au niveau des sources, indiquant un ajustement correct.

4.5.3 Ajustement du spectre de chaque source

Le flux à 260 GHz de chaque source est ensuite combiné avec les trois flux SPIRE, permettant une couverture spectrale allant de 260 à 1200 GHz. La SED de chaque source est ensuite ajustée sur ces quatre flux, en utilisant un modèle d'émission de corps noir modifié :

$$F(\nu) = A_0 \left(\frac{\nu}{\nu_0} \right)^\beta B_\nu(T), \quad (4.19)$$

où $B_\nu(T)$ est le spectre d'un corps noir de température T , A_0 est l'amplitude de la SED à une fréquence de référence $\nu_0 = 500$ GHz, β est l'indice spectral de la poussière, et T sa température effective, dégénérée avec le redshift de la source z comme $T = T_{\text{dust}}/(1+z)$.

L'ajustement de chaque SED est effectué en utilisant un échantillonnage Monte Carlo à chaînes de Markov (MCMC) et la librairie Python `emcee` [157]. Le principe de fonctionnement du MCMC sera présenté au chapitre 5. Nous considérons une fonction de vraisemblance gaussienne. Lorsque les flux des sources ponctuelles dans les bandes de PACS ne sont pas disponibles, la couverture spectrale est relativement étroite, et ne permet pas de lever la dégénérescence entre les trois paramètres du modèle de corps noir modifié [158-160]. Il n'est donc pas possible de laisser ces paramètres complètement libres, avec des *a priori* non-informatifs. La normalisation de la SED, A_0 , est donc ajustée linéairement sur les données et fixée pour chaque ajustement. Un *a priori* sur l'indice spectral de chaque source est imposé par une distribution gaussienne $\beta \sim \mathcal{N}(2, 0.5)$, tel que suggéré par les mesures de SED de grands nombres de galaxies (voir par exemple [159]). Enfin, un *a priori* uniforme est utilisé pour la

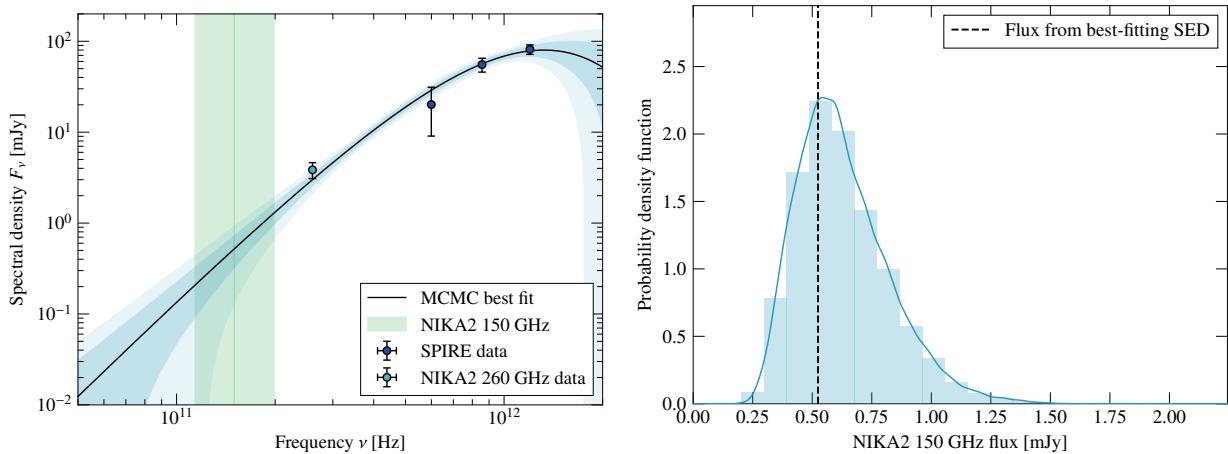


Figure 4.9 : **Gauche** : Résultat de l’ajustement de la SED de la source SMG1 dans le champ de l’amas ACT-CL J0215.4+0030, présenté au chapitre 6. Les points bleu foncé représentent des flux SPIRE. Les points bleu clair représentent le flux mesuré dans la carte NIKA2 à 260 GHz. La ligne bleue montre la SED ajustant le mieux les données, et les enveloppes l’entourant montrent les intervalles de confiance à 1σ et 2σ . La région verte correspond à la bande passante à 150 GHz de NIKA2. **Droite** : Distribution de probabilité du flux de SMG1 dans la bande passante à 150 GHz de NIKA2. L’histogramme représente la distribution des flux calculés grâce au MCMC (voir texte). La courbe bleue montre la densité de probabilité obtenue en appliquant une estimation par noyaux à cette distribution. La ligne verticale noire montre le flux obtenu grâce à la SED ajustant le mieux les données.

température effective de chaque source, $0 < T < 50$ K.

À chaque étape du MCMC, une correction de couleur est appliquée aux flux mesurés afin de tenir compte de la différence entre le spectre de chaque source et celui de référence utilisé pour l’étalonnage des cartes. Pour les flux SPIRE, la correction de couleur est interpolée à partir de mesures existantes⁵; pour le flux NIKA2, nous utilisons les corrections décrites dans [11].

À titre illustratif, la partie gauche de la figure 4.9 montre le résultat de cet ajustement pour l’une des sources proches de l’amas ACT-CL J0215.4+0030. Les flux utilisés pour contraindre le modèle sont représentés par les points bleus, et le modèle de SED résultant par la ligne bleue et les enveloppes d’erreur l’entourant.

4.5.4 Inférence du flux à 150 GHz

L’ajustement de SED produit pour chaque source un échantillonnage de la distribution de probabilité dans l’espace des paramètres (β, T). Chaque échantillon de cette distribution est alors utilisé pour calculer une SED grâce à l’équation (4.19). Chacune de ces SED est extrapolée pour calculer une valeur de flux à 150 GHz, et une correction de couleur est appliquée pour en déduire une valeur de flux attendue dans la bande passante de NIKA2. La distribution de probabilité associée à ce flux est calculée à partir de ces mesures en utilisant une estimation par noyau (KDE pour *Kernel Density Estimation*). La combinaison de cette mesure avec l’ajustement des sources à 260 GHz nous donne accès, pour chaque source, à une connaissance précise de sa position et de la distribution de probabilité de l’amplitude de sa contamination de l’effet SZ à 150 GHz. Le résultat de ce calcul est illustré dans la partie droite de la figure 4.9.

Le produit principal de PSTools est donc la distribution de probabilité du flux de chaque source

⁵http://herschel.esac.esa.int/Docs/SPIRE/html/spire_om.html, §5.2.7

dans la bande à 150 GHz de NIKA2. Il existe plusieurs façons d'utiliser cette information pour traiter la contamination due à chaque source. La plus simple consiste à soustraire un modèle de source à la carte NIKA2, avec un flux donné par le flux le plus probable issu de l'ajustement de SED, correspondant à la valeur maximisant la distribution de probabilités obtenue par PSTools. Cette approche ne permet toutefois pas de tirer pleinement profit de l'information disponible, par exemple l'incertitude sur ce flux. Nous verrons au chapitre suivant (section 5.1.3) que les flux des sources peuvent être traités comme des paramètres de nuisance dans la modélisation de la carte de l'effet SZ d'un amas. Cette approche permet de propager l'incertitude sur les flux des sources aux propriétés physiques de l'amas en utilisant la distribution de probabilités issue de PSTools comme *a priori* dans l'analyse.

4.6 Conclusions et perspectives

Ce chapitre a présenté le traitement des données brutes de NIKA2 pour obtenir des cartes de l'effet SZ. Nous avons vu en détail le principe de la décorrélation, c'est-à-dire l'extraction du signal astrophysique dans les observations NIKA2 dominées par différentes sources de bruit. Le calcul de la matrice de covariance du bruit résiduel dans les cartes et du filtrage du signal a également été présenté. Nous avons également présenté le logiciel PSTools, qui permet d'estimer l'amplitude de la contamination de l'effet SZ par des sources ponctuelles sub-millimétriques en combinant les observations NIKA2 et *Herschel*. Nous verrons dans le chapitre 5 comment ces éléments sont inclus dans l'analyse des propriétés physique d'un amas pour tenir compte des effets systématiques qu'ils engendrent.

Comme nous l'avons discuté en 4.4, la décorrélation est un processus imparfait, qui n'est pas en mesure de distinguer parfaitement le bruit du signal d'intérêt. Sa performance peut se juger sur deux quantités mesurables : le spectre de puissance du bruit résiduel, et le filtrage du signal. Comme nous le verrons au chapitre suivant, ces quantités servent à la fois de critères pour quantifier la performance de la décorrélation et d'outils pour tenir compte de son imperfection dans l'analyse des cartes NIKA2. En pratique, le choix d'une méthode est basé sur un compromis entre filtrage et bruit résiduel. En effet, les méthodes dites agressives, qui permettent d'atteindre un bruit résiduel très faiblement corrélé, sont en général caractérisées par un fort filtrage du signal d'intérêt. À l'inverse, les méthodes préservant le mieux le signal sont en général moins performantes pour la soustraction du bruit corrélé. Sur la base de ce compromis, la méthode *common mode one block* est celle retenue pour l'analyse des observations du grand programme SZ de NIKA2. Elle permet d'obtenir un filtrage relativement faible du signal aux échelles angulaires d'intérêt pour l'étude des amas de galaxies du grand programme SZ, au prix d'un bruit résiduel très corrélé.

Au sein du grand programme SZ de NIKA2, la décorrélation est réalisée à l'aide du *pipeline* de la collaboration NIKA2, fruit d'un effort de la collaboration pour obtenir un outil permettant de passer des données en temps aux cartes du ciel. J'ai développé un logiciel basé sur ce *pipeline* pour réaliser ce traitement dans le cas des observations du grand programme SZ. Ce logiciel permet d'interroger la base de données AMI du grand programme SZ, décrite en section 3.5 (page 72), pour obtenir la liste de *scans* associés à un amas, réaliser la décorrélation de ces *scans* et les coadditionner, puis de calculer la matrice de covariance du bruit résiduel et la fonction de transfert de la décorrélation. Il a été mis à disposition de la collaboration, et est aujourd'hui systématiquement utilisé pour le traitement des données brutes des amas du grand programme SZ.

La décorrélation telle que présentée dans ce chapitre représente donc la procédure standard pour la décorrélation des données NIKA2 du grand programme SZ à l'heure de l'écriture de cette thèse. L'amélioration de la procédure est toutefois un sujet très actif au sein de la collaboration NIKA2.

Parmi les axes d'amélioration en cours d'étude, de nouvelles méthodes de décorrélation sont en cours de développement et de test. On notera en particulier le développement de méthodes novatrices, basées sur une décorrélation itérative sans masque. Dans ces méthodes, à une itération i , plutôt que de masquer la région contenant du signal significatif dans la carte de l'itération $i - 1$ dans le calcul du mode commun, la carte de l'itération $i - 1$, supposée dominée par le signal d'intérêt, est soustraite aux données brutes. L'intérêt est ici d'éviter de devoir estimer des modes communs sur trop peu d'échantillons. Pour des sources étendues comme des amas de galaxies, masquer la région dans laquelle le signal est présent peut laisser peu de détecteurs pour l'estimation du mode commun, résultant en une mauvaise estimation du bruit. Le processus itératif permet alors de converger vers une carte contenant le signal d'intérêt en ayant mieux retiré le bruit. Ces méthodes ont mené à plusieurs tests au cours de cette thèse, dont les résultats sont encore très préliminaires. Elles montrent une performance remarquable dans la soustraction du bruit corrélé, mais un filtrage important du signal astrophysique, qui est encore mal compris. C'est pourquoi *common mode one block* reste la méthode choisie pour les observations d'amas de galaxies avec NIKA2.

Un autre axe en cours de développement, intimement lié avec le développement de nouvelles méthodes, est le calcul de la fonction de transfert de la décorrélation. La procédure présentée en 4.4.2 se base sur plusieurs hypothèses. L'une de ces hypothèses a été discutée : nous avons supposé que le filtrage subi par le signal simulé au cours de notre analyse était identique à celui subi par le signal astrophysique d'intérêt. En réalité, ces deux filtrages ne sont vraisemblablement pas exactement identiques. La décorrélation de deux jeux de données, contenant le même bruit, mais dont l'un contient environ deux fois plus de signal que l'autre, peut être différente. En particulier, les régions en périphérie du masque, dans lesquelles on trouve du signal significatif à $\lesssim 3\sigma$ dans l'analyse des données brutes de NIKA2, contiendront du signal plus significatif dans les données contenant le signal simulé. Il est donc possible que l'estimateur de fonction de transfert présenté en 4.4.2 ne soit pas complètement juste. Afin de quantifier cet effet, un travail est en cours pour créer des TOI entièrement synthétiques, contenant un bruit corrélé réaliste et du signal astrophysique pour chaque détecteur. L'objectif est d'appliquer la procédure de décorrélation à ces TOI simulées pour évaluer la différence entre le filtrage réellement subi par un signal simulé, et l'estimateur de fonction de transfert présenté précédemment.

Une autre hypothèse implicite dans l'équation (4.18) est celle de l'isotropie du filtrage, c'est-à-dire de la symétrie circulaire du rapport entre les transformées de Fourier des cartes d'entrée et de sortie. Cette hypothèse est justifiée par la stratégie de *scan* du grand programme SZ de NIKA2, discutée en 4.3.3, qui permet une couverture très isotrope de la région cartographiée en utilisant des *scans* à différents angles. Cependant, certaines directions peuvent rester privilégiées dans le filtrage, notamment les directions des *scans*. Une étude est donc en cours pour quantifier l'écart à la symétrie circulaire du filtrage subi par le signal d'intérêt lors de la décorrélation. Les résultats en sont très prometteurs, montrant l'intérêt de l'utilisation de fonctions de transfert à deux dimensions pour les analyses du grand programme SZ [161].

La décorrélation reste donc un sujet d'étude actif, et les prochaines années permettront des avancées pour maximiser la qualité des cartes de l'effet SZ construites grâce aux observations NIKA2.

CHAPITRE 5

Pipeline pour la mesure des propriétés thermodynamiques du milieu intra-amas

5.1	Modélisation du signal SZ	103
5.1.1	Profil de pression et effet Sunyaev-Zeldovich	103
5.1.2	Modèle de carte de brillance de surface SZ	104
5.1.3	Contamination par des sources ponctuelles	105
5.1.4	Signal SZ intégré	106
5.2	Ajustement par méthode Monte Carlo à Chaînes de Markov	106
5.2.1	Fonction de vraisemblance	106
5.2.2	Distribution <i>a priori</i> des paramètres	107
5.2.3	Échantillonnage de la distribution de probabilités	107
5.2.4	Évaluation des propriétés thermodynamiques	110
5.3	Validation de PANCO sur des amas simulés	112
5.3.1	Simulations d'amas simples	112
5.3.2	Optimisation de la distribution radiale	118
5.3.3	La simulation MUSIC et l'échantillon jumeau du grand programme SZ	120
5.3.4	Ajustement du profil de pression	122
5.4	Conclusions	122

L'objectif premier des observations d'amas de galaxies avec NIKA2 est de mesurer les propriétés thermodynamiques du milieu intra-amas. En particulier, l'amplitude du signal tSZ mesuré en direction d'un amas permet la mesure de sa distribution de pression, qui peut ensuite être combinée à des observations X pour estimer la distribution de masse de l'amas, ainsi que d'autres propriétés thermodynamiques, comme décrit en section 3.4.3 (page 67).

Le logiciel PANCO (*Pipeline for the Analysis of NIKA2 Cluster Observations*) développé par F. Ruppin [162] a pour objectif d'automatiser ce processus dans le contexte du grand programme SZ de NIKA2. Au cours de ma thèse, après avoir utilisé ce logiciel, j'ai décidé d'en développer une version

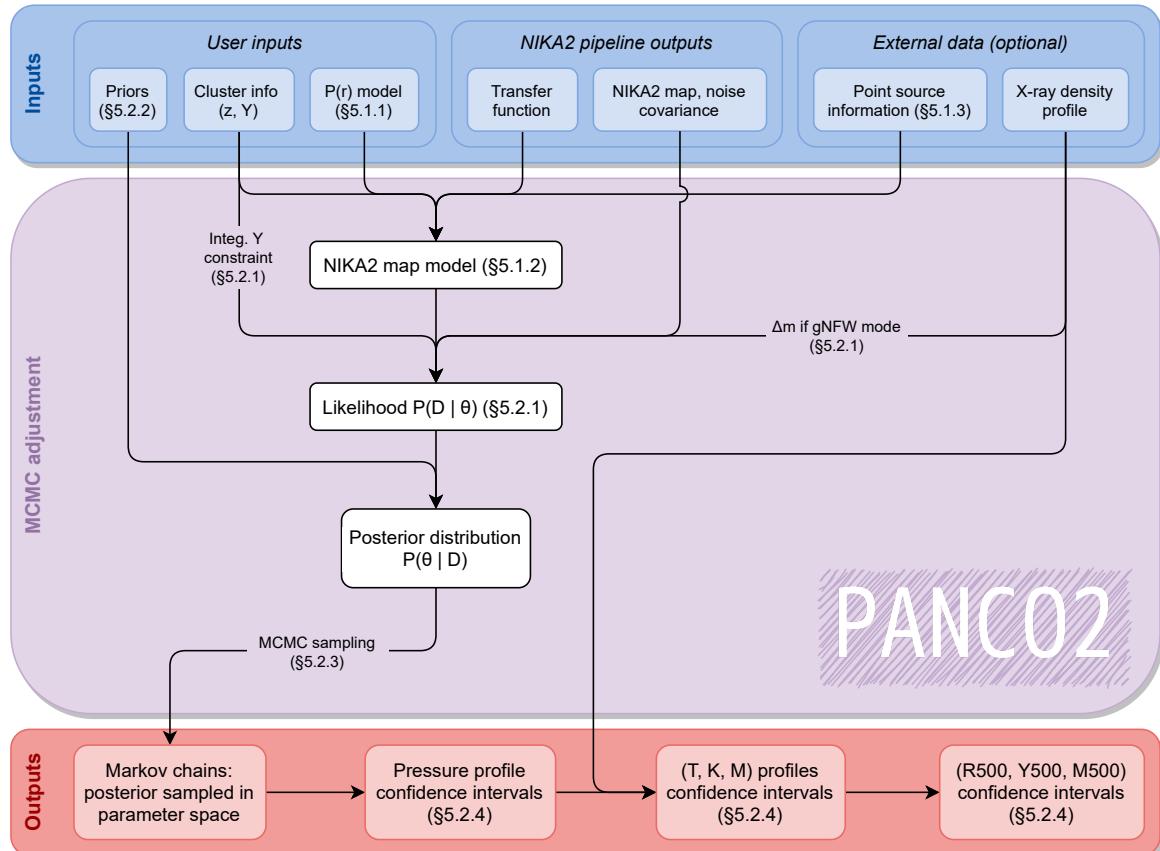


Figure 5.1 : Schéma représentant le principe de fonctionnement de PANCO2. Les données d’entrée sont les produits standards pour les observations NIKA2 d’amas de galaxies, produits par le pipeline d’analyse des données brutes présenté au chapitre 4, ainsi que des données externes telles que les profils radiaux mesurés en X ou les informations sur les sources ponctuelles.

plus optimisée, basée sur l’utilisation de librairies Python dédiées au calcul numérique. L’objectif de ce développement était double. D’une part, il était nécessaire de fournir à la collaboration NIKA2 un code plus rapide, le temps d’exécution de PANCO étant trop grand pour permettre l’analyse d’échantillons d’amas. D’autre part, il s’agissait d’une occasion de créer un outil standard pour le grand programme SZ de NIKA2, là où PANCO avait été développé pour l’analyse jointe de plusieurs cartes d’un même amas.

Ce nouveau logiciel, nommé PANCO2, a été rendu disponible pour la collaboration NIKA2, après avoir été testé sur des simulations. Le temps d’exécution a été réduit d’un facteur dix à cent par rapport à PANCO. Ce gain en performance a motivé le choix de PANCO2 comme pipeline SZ officiel de la collaboration NIKA2. Une généralisation du code est en cours, qui permettra d’utiliser PANCO2 pour l’exploitation de cartes issues d’autres instruments, et qui sera livrée à la communauté comme un code public. Ce chapitre décrit PANCO2, du principe de la mesure des propriétés thermodynamiques de l’ICM grâce à l’effet SZ à l’algorithme employé pour effectuer cette mesure.

Un schéma synoptique du fonctionnement de PANCO2 est présenté en figure 5.1. Il répertorie les différentes données d’entrée (bleu) nécessaire à la mesure des propriétés thermodynamiques (rouge). Les données principales sont les sorties du pipeline de la collaboration NIKA2, décrites au chapitre précédent. Des informations sur la contamination par des sources ponctuelles, issues du logiciel PSTools détaillé en section 4.5, peuvent également être fournies – elles seront discutées en 5.1.3 (page 105). On note aussi la possibilité de donner en entrée un profil de densité mesuré en X, qui permettra la com-

binaison des données X et SZ pour une caractérisation détaillée des propriétés thermodynamiques du milieu intra-amas, discutée en 5.2.4.

5.1 Modélisation du signal SZ

L'objectif principal de PANC02 est la mesure du profil de pression du milieu intra-amas à partir d'une carte de l'effet SZ. Pour cela, une approche *forward modelling* est employée. Une modélisation de l'état physique du système, basée sur un nombre de paramètres, est établie, puis affectée des différents effets ayant donné naissance aux données observées. Cela permet dans notre cas d'obtenir, à partir d'hypothèses physiques, un modèle de l'effet SZ mesuré par NIKA2, qui sera comparé aux données afin d'estimer les propriétés physiques de l'amas. Cette section décrit les étapes successives nécessaires à l'obtention d'un tel modèle.

5.1.1 Profil de pression et effet Sunyaev-Zeldovich

Comme décrit dans le chapitre 2, l'effet Sunyaev-Zeldovich est une distorsion spectrale du fond diffus cosmologique due à la diffusion Compton inverse des photons sur les électrons du milieu intra-amas. Dans le cadre de l'effet Sunyaev-Zeldovich thermique, dominant dans les observations NIKA2 d'amas de galaxies, l'énergie transférée par les électrons provient de leur grande agitation thermique. L'amplitude de la distorsion observée aux coordonnées θ est alors appelée paramètre de Compton y , et est proportionnelle à la pression électronique P_e du milieu intra-amas intégrée le long de la ligne de visée (LdV) correspondante. D'après l'équation (2.11, page 33), on a :

$$y(\theta) = \frac{\sigma_T}{m_e c^2} \int_{\text{LdV}(\theta)} P_e(l) dl. \quad (5.1)$$

Par conséquent, la cartographie de l'effet SZ est une mesure directe de la distribution intégrée de pression du milieu intra-amas dans le plan céleste. Bien que cette distribution ne soit pas mesurable le long de la ligne de visée, des hypothèses peuvent être faites sur la symétrie du milieu intra-amas. Dans le cadre de l'hypothèse de symétrie sphérique, la distribution de pression ne dépend que de la distance au centre de l'amas, et est alors décrite par un profil radial de pression $P_e(r)$. Une carte de l'effet tSZ permet alors de contraindre ce profil de pression. Par transformation d'Abel, l'équation (5.1) devient :

$$y(R) = 2 \frac{\sigma_T}{m_e c^2} \int_R^{+\infty} \frac{r}{\sqrt{r^2 - R^2}} P_e(r) dr. \quad (5.2)$$

L'objectif principal de PANC02 est la mesure du profil de pression dans l'hypothèse de symétrie sphérique. Dans PANC02, je propose deux modélisations du milieu intra-amas, illustrées sur la figure 5.2 :

- Le modèle Navarro-Frenk-White généralisé (noté « gNFW »), dans lequel le profil de pression est décrit par [163, 164] :

$$P_e(r) = P_0 \left(\frac{r}{r_p} \right)^{-c} \left[1 + \left(\frac{r}{r_p} \right)^a \right]^{\frac{c-b}{a}}, \quad (5.3)$$

où P_0 représente la normalisation du profil, b et c les pentes externe et interne du profil de pression, et r_p et a le rayon de transition entre les deux régimes et le caractère abrupt de cette transition ;

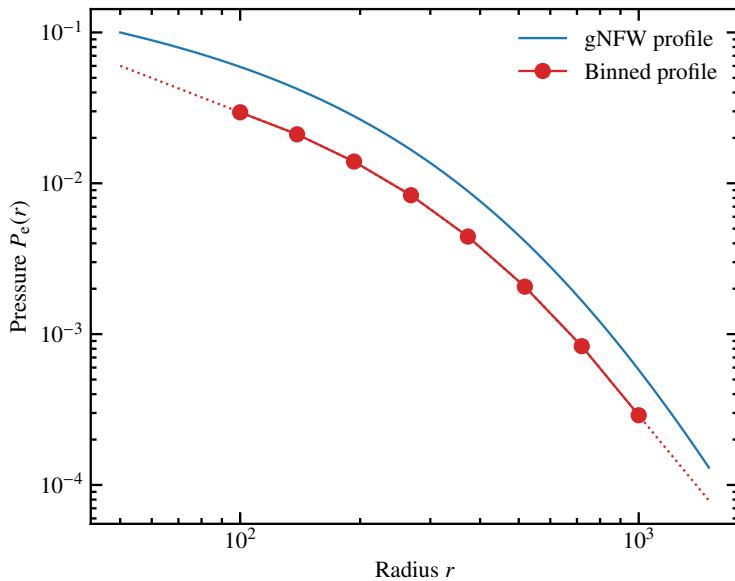


Figure 5.2 : Comparaison entre un profil de pression gNFW (bleu) et non-paramétrique (rouge). La valeur de la pression du profil non-paramétrique est divisée par deux à des fins de visibilité. Les lignes pointillées représentent l'extrapolation du profil non-paramétrique au-delà du premier et dernier point, en considérant une loi de puissance identique à celle entre le premier et deuxième point ($\text{à } r < R_{\min}$) et entre l'avant-dernier et le dernier point ($\text{à } r > R_{\max}$).

- Un modèle dit «non-paramétrique¹», dans lequel le profil de pression est décrit par des coquilles sphériques concentriques au sein desquelles la pression diminue avec le rayon en loi de puissance selon [165, 166] :

$$P_e(r) = P_i \left(\frac{r}{R_i} \right)^{-\alpha_i}. \quad (5.4)$$

Chacune des modélisations présente ses avantages et inconvénients. Le profil gNFW est une fonction continue du rayon, aisément extrapolable et avec un nombre fixe de paramètres, le rendant plus directement comparable à la littérature dans laquelle il est le plus répandu. Le profil non-paramétrique offre quant à lui une approche moins dépendante du modèle, et donc la possibilité de détecter des irrégularités dans le profil de pression, telles que des régions de surpression locales dues par exemple à des chocs. Le plus grand avantage du profil non-paramétrique provient cependant des corrélations plus faibles entre ses paramètres, qui offre de meilleures performances dans l'ajustement, comme décrit dans la section 5.3.4 (page 122).

5.1.2 Modèle de carte de brillance de surface SZ

Le profil de pression considéré est ensuite intégré le long de la ligne de visée afin d'obtenir une carte de paramètre de Compton y dans la région de l'amas. Dans le cas d'un modèle gNFW, l'intégrale de l'équation (5.1) n'a pas de solution analytique, et une intégration numérique est utilisée, en considérant un grand nombre de pas espacés logarithmiquement le long de la ligne de visée afin de limiter les artefacts d'intégration. Dans le cas d'un modèle non-paramétrique, l'intégrale de l'équation (5.2) est la somme des intégrales analytiques du profil de pression dans chacune des coquilles sphériques [166].

La carte de paramètre de Compton obtenue doit ensuite être convertie en unités de brillance de surface afin d'être comparable aux données observées. Le facteur de conversion dépend des bandes

¹Il serait plus juste de parler de profil «binné», car le modèle est de fait paramétrique : ses paramètres sont les valeurs du profil de pression aux rayons spécifiés. Nous parlerons toutefois de modèle non-paramétrique par abus de langage, et par opposition au modèle gNFW dans lequel le nombre de paramètres est fixé par l'expression du modèle.

passantes effectives de NIKA2 (et donc des conditions atmosphériques au moment des observations), mais également du lobe de l'instrument. Il dépend également de la forme du spectre de l'effet SZ, et donc de la température du milieu intra-amas du fait des corrections relativistes à l'effet SZ, comme décrit en section 2.2.3 (page 32). En pratique, PANCO2 traite ce facteur de conversion comme un paramètre de nuisance de l'analyse, ce qui permet de tenir compte de l'incertitude systématique liée à la variabilité de ce coefficient due aux conditions d'observations.

Les cartes de l'effet SZ réalisées par NIKA2 sont affectées par deux formes de filtrage. Tout d'abord, l'image du ciel est convoluée par le lobe instrumental caractéristique du couplage entre NIKA2 et le télescope (décris au chapitre 3). Le modèle de carte de brillance de surface SZ est donc convolué par un filtre gaussien de largeur à mi-hauteur spécifiée par l'utilisateur. Les cartes sont également affectées par le filtrage découlant de la procédure de soustraction du bruit des données en temps de NIKA2. Ce filtrage est estimé par le calcul de la fonction de transfert présenté au chapitre 4, et la carte de modèle est également convoluée par cette fonction de transfert. La carte résultante est donc une carte étalonnée en unités de brillance de surface, affectée par le lissage des petites échelles angulaires caractéristique du couplage entre NIKA2 et le télescope, et par le filtrage des grandes échelles provenant de l'analyse des données brutes. Elle peut donc être comparée à la carte de brillance de surface de l'effet SZ observée par NIKA2.

5.1.3 Contamination par des sources ponctuelles

Le signal astrophysique mesuré par NIKA2 lors des observations d'amas de galaxies est la somme des contributions des différents objets présents le long de la ligne de visée. Les contributeurs majeurs sont le milieu intra-amas par effet SZ, constituant l'observable d'intérêt pour les études présentées dans cette thèse ; ainsi que l'émission de sources ponctuelles. Cette émission, par nature positive, peut compenser partiellement ou totalement le décrément SZ détecté dans les cartes NIKA2 à 150 GHz, et peut donc induire un biais négatif dans le profil de pression du milieu intra-amas reconstruit à partir des observations NIKA2. La contamination résultante doit donc être prise en compte dans les mesures de propriétés du milieu intra-amas afin de ne pas souffrir de ce biais lors de l'utilisation des résultats du grand programme SZ en cosmologie.

Cette contamination peut être prise en compte en soustrayant des modèles de sources des données NIKA2. Pour cela, il est nécessaire de connaître la valeur du flux de chaque source, et une mauvaise estimation de ces flux peut induire une erreur dans les propriétés thermodynamiques mesurées. PANCO2 offre la possibilité de traiter cette contamination comme un paramètre de nuisance dans l'ajustement du profil de pression du milieu intra-amas. Cette approche, présentée plus en détail dans le chapitre 6, permet de propager l'incertitude sur les flux des différentes sources ponctuelles à la mesure des propriétés thermodynamiques du milieu intra-amas. Pour cela, des sources modélisées comme le lobe de NIKA2² peuvent être ajoutées au modèle de carte de brillance de surface présenté précédemment avant que celle-ci ne soit comparée aux données. Les flux de ces sources sont alors considérés comme des paramètres supplémentaires du modèle. La distribution de probabilité du flux de chaque source, obtenue avec PSTools (voir section 4.5), peut alors être utilisée comme *a priori* sur le paramètre, comme nous le verrons en section 5.2.2.

²Voir section 3.3.4.

5.1.4 Signal SZ intégré

La couverture angulaire des observations SZ réalisées par NIKA2 est limitée par le filtrage des grandes échelles angulaires issu de la soustraction du bruit atmosphérique. Ce filtrage se caractérise par une perte de signal en périphérie de l'amas, ce qui empêche l'établissement de contraintes fortes sur le profil de pression du milieu intra-amas dans ces régions. Une information sur ce signal est contenue dans le signal SZ intégré de l'amas, mesuré dans les relevés à plus grand champ de vue desquels est issu le catalogue du grand programme SZ, à savoir les relevés *Planck* [5] et ACT [136]. Ce signal est proportionnel à l'intégrale sphérique du profil de pression considéré jusqu'à un rayon donné R :

$$Y_R = 4\pi \frac{\sigma_t}{m_e c^2} \int_0^R r^2 P_e(r) dr. \quad (5.5)$$

Cette valeur peut donc être calculée à partir du modèle de profil de pression considéré, et comparée aux valeurs reportées dans les relevés SZ, imposant une contrainte effective sur le profil de pression dans les régions non contraintes par les observations NIKA2. Les valeurs de Y_{500} sont disponibles pour tous les amas du grand programme SZ de NIKA2, mesurées par *Planck* ou ACT, et sont systématiquement utilisées afin d'améliorer la contrainte du profil de pression dans la périphérie de l'amas.

5.2 Ajustement par méthode Monte Carlo à Chaînes de Markov

Afin de déterminer le profil de pression du milieu intra-amas, le modèle de carte de brillance de surface décrit dans la section précédente est ajusté sur les données NIKA2 en utilisant un algorithme Monte Carlo à Chaînes de Markov (MCMC). L'objectif est donc de trouver la distribution de probabilités des paramètres étant données les mesures effectuées par NIKA2. Cette distribution, dite postérieure, est donnée par le théorème de Bayes :

$$P(\vartheta | D) = \frac{P(D | \vartheta) P(\vartheta)}{P(D)}, \quad (5.6)$$

où ϑ est le vecteur formé par les différents paramètres de l'analyse, résumés dans le tableau 5.1, et D représente les données mesurées, qui dans le cas de PANCO2, sont la carte de brillance de surface mesurée par NIKA2 et le signal SZ intégré mesuré par ACT ou *Planck*.

La distribution $P(D | \vartheta)$ est la fonction de vraisemblance comparant le modèle aux données, et $P(\vartheta)$ est notre connaissance *a priori* sur les valeurs des paramètres de l'analyse. $P(D)$ est traitée comme une constante de normalisation permettant à l'équation (5.6) d'être intégrable à 1.

5.2.1 Fonction de vraisemblance

La fonction de vraisemblance, ou *likelihood* en anglais, est la probabilité qu'un jeu de données D soit issu d'un modèle ϑ . Dans PANCO2, j'ai choisi d'utiliser une fonction de vraisemblance Gaussienne multivariée dont chacun des pixels de la carte NIKA2 est une dimension :

$$\log P(D | \vartheta) = -\frac{1}{2} (D_N - \mathcal{M}_N(\vartheta))^T \Sigma_N^{-1} (D_N - \mathcal{M}_N(\vartheta)) - \frac{1}{2} \left(\frac{Y_R^{\text{meas.}} - Y_R(\vartheta)}{\Delta Y_R^{\text{meas.}}} \right)^2 - \Delta_{\text{mass}}, \quad (5.7)$$

où D_N représente la carte NIKA2 à 150 GHz, $\mathcal{M}_N(\vartheta)$ le modèle de carte calculé à partir des paramètres ϑ , et Σ_N la matrice de covariance du bruit ; $Y_R^{\text{meas.}}$ et $\Delta Y_R^{\text{meas.}}$ sont la mesure de signal SZ intégré dans

un rayon R et son incertitude, et $Y_R(\vartheta)$ le signal SZ intégré correspondant au modèle considéré (dans la suite, nous utiliserons $R = R_{500}$).

Cette fonction de vraisemblance permet donc de comparer un modèle aux observations réalisées avec NIKA2, en tenant également compte de la contrainte sur le paramètre de Compton intégré mesuré par *Planck* ou *ACT*. Dans le cadre d'un modèle gNFW, et si des données X sont fournies par l'utilisateur, le terme Δ_{mass} est ajouté afin d'interdire les points pour lesquels le profil de masse décroît avec le rayon :

$$\Delta_{\text{mass}} = \begin{cases} 0 & \text{si } dM^{\text{HSE}}/dr \geq 0 \forall r, \\ -\infty & \text{sinon,} \end{cases} \quad (5.8)$$

où M^{HSE} est calculé à partir de l'équation de l'équilibre hydrostatique (2.3).

Ce terme permet de restreindre l'espace des paramètres aux profils de pression ayant un sens physique, la masse contenue dans un rayon croissant ne pouvant pas diminuer. Il n'est toutefois pas implémenté lors de l'ajustement d'un modèle non-paramétrique. En effet, de telles analyses peuvent avoir pour objectif la mesure de surpressions locales, par exemple dues à des chocs. Dans ces cas où l'équilibre hydrostatique n'est pas respecté, l'imposition d'une telle contrainte empêcherait la détection de telles particularités du profil de pression. Cette contrainte n'est par conséquent imposée qu'au cours de l'ajustement de modèles gNFW, pour lesquels des fluctuations de pression ne sont de toute façon pas détectables, et qui bénéficient de la réduction de l'espace des paramètres au vu des fortes corrélations entre ces derniers.

5.2.2 Distribution *a priori* des paramètres

L'approche bayésienne utilisée dans PANCO2 requiert la définition d'une distribution de probabilité *a priori* des paramètres ϑ qui sera combinée à la fonction de vraisemblance pour préciser la connaissance des paramètres. Cette distribution est modélisée comme le produit de distributions définies individuellement pour chaque paramètre, soit en considérant chaque paramètre comme indépendant. PANCO2 propose plusieurs possibilités de distribution *a priori* pour chaque paramètre, résumées dans le tableau 5.1. Notons en particulier que les paramètres correspondant au flux des sources ponctuelles peuvent prendre comme *a priori* la distribution de probabilité issue de PSTools, obtenue par ajustement du spectre des sources (voir section 4.5).

5.2.3 Échantillonnage de la distribution de probabilités

L'ajustement du profil de pression est réalisé par échantillonnage Monte Carlo à Chaînes de Markov (MCMC) de la distribution postérieure définie par l'équation (5.6). L'avantage du MCMC par rapport à une maximisation de fonction de vraisemblance est l'obtention d'un ensemble de points de l'espace des paramètres distribués selon la postérieure de notre analyse, qui pourra ensuite être transposé dans l'espace des propriétés physiques du milieu intra-amas.

Principe du MCMC

Le principe de fonctionnement du MCMC repose sur une exploration de l'espace des paramètres par marche aléatoire, conditionnée par la valeur de la distribution de probabilités aux points visités. Cette marche aléatoire peut être décrite par une série d'étapes successives :

1. La marche aléatoire commence en un point de l'espace des paramètres de coordonnées définies par un vecteur ϑ . La valeur de la distribution de probabilités considérée est $P(\vartheta)$.

Paramètre	Description	Distribution <i>a priori</i>
<i>Paramètres d'intérêt pour un modèle gNFW</i>		
P_0, r_p, a, b, c	Paramètres du profil de pression	<ul style="list-style-type: none"> – Gaussiennes définies par l'utilisateur – Uniformes avec limites choisies par l'utilisateur
<i>Paramètres d'intérêt pour un modèle non-paramétrique</i>		
$P_i, i = 0 \dots n_{\text{bins}}$	Valeurs de pression	Uniformes entre 0 et $1 \text{ keV} \cdot \text{cm}^{-3}$
<i>Paramètres de nuisance</i>		
C_{conv}	Conversion de y en Jy/beam	Gaussienne définie par l'utilisateur
Z	Niveau zéro de la carte	Gaussienne $\mathcal{N}(0, 0.1 \text{ mJy}/\text{beam})$
$F_i, i = 0 \dots n_{\text{ps}}$	Flux des sources ponctuelles	<ul style="list-style-type: none"> – Distribution issue de PSTools (section 4.5) – Uniformes avec limites définies par l'utilisateur

Table 5.1 : Liste des paramètres utilisés par PANCO2 pour modéliser les cartes SZ construites à partir d'observations NIKA2. La dernière colonne indique les distributions *a priori* utilisables pour chaque paramètre.

2. Une position proche dans l'espace des paramètres, ϑ' , est tirée au hasard d'après une fonction de proposition prédéfinie.
3. Si $P(\vartheta') \geq P(\vartheta)$, alors ϑ' est automatiquement acceptée comme nouvelle position de l'échantillonneur ; sinon, elle est aléatoirement acceptée ou non avec une probabilité d'autant plus faible que $P(\vartheta') < P(\vartheta)$, nommée probabilité d'acceptance.
4. La procédure est répétée autant de fois que nécessaire pour obtenir un nombre d'échantillons jugé suffisant.

Ce processus peut être réalisé en parallèle par plusieurs marcheurs, créant chacun simultanément un ensemble de points nommé chaîne de Markov. L'ensemble des positions ϑ acceptées par l'échantillonneur est alors distribué selon la fonction de distribution de probabilités considérée.

Un exemple commun de MCMC est donné par l'algorithme de Metropolis-Hastings [167, 168], où la fonction de proposition est une distribution Gaussienne de largeur constante et la probabilité d'acceptance est donnée par $p = P(\vartheta') < P(\vartheta)$. Celui-ci est limité par la nécessité de définir au préalable une fonction de proposition immuable au cours de l'échantillonnage. PANCO2 utilise la librairie emcee [157], proposant plusieurs implémentations en Python d'échantillonneurs adaptant leur fonction de proposition au cours de l'échantillonnage, introduits par Goodman et Weare [169]. En particulier, dans le cas de PANCO2, l'échantillonneur retenu est celui offrant les meilleures performances en termes d'échantillons indépendants générés par unité de temps, l'échantillonneur DEMove [170].

Convergence des chaînes et autocorrélation

L'échantillonnage MCMC réalisé par PANCO2 est automatiquement stoppé lorsque la convergence des chaînes a été atteinte, c'est-à-dire à partir de l'instant où l'ajout de nouveaux points aux chaînes de Markov n'apporterait pas d'information significative par rapport à ceux déjà générés. La convergence est basée sur trois critères : l'adaptation des chaînes, leur mélange, et leur autocorrélation, décrite ci-après.

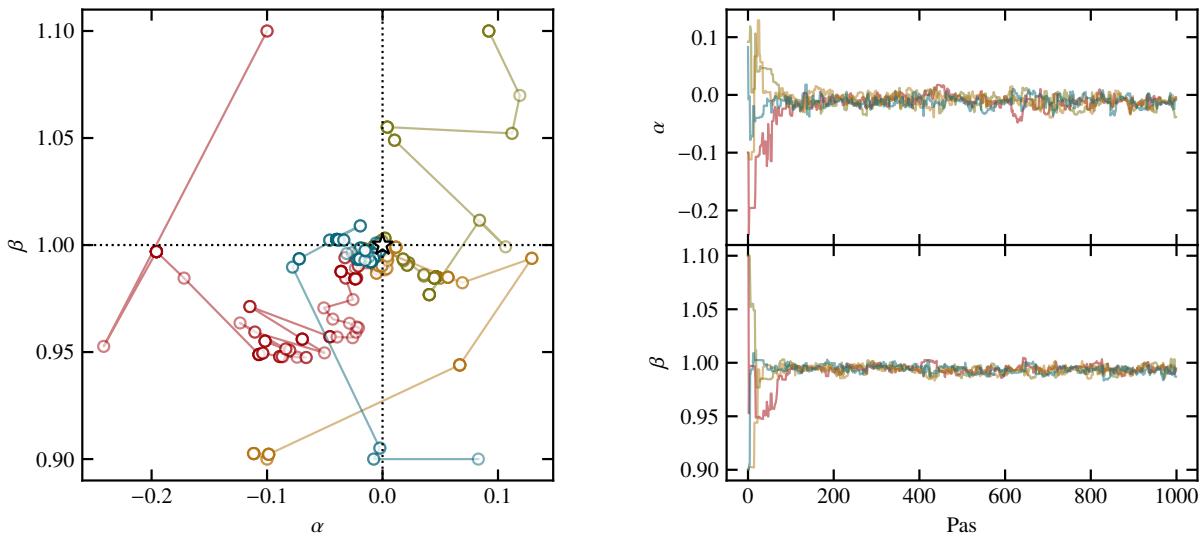


Figure 5.3 : Illustration du mouvement des chaînes d'un échantillonneur MCMC dans le cas d'une régression à deux paramètres (α, β), dans l'espace des paramètres (*gauche*) et projeté pour chacun des paramètres individuellement (*droite*). Chaque couleur représente une chaîne de Markov. L'étoile dans le panneau gauche représente le maximum de la distribution postérieure considérée.

Le point de départ du MCMC dans l'espace des paramètres est arbitraire et ne se trouve pas nécessairement dans la région d'intérêt. Par conséquent, il est nécessaire d'amputer chaque chaîne d'une longueur d'adaptation, correspondant au nombre de pas nécessaires à chaque marcheur pour rejoindre cette région d'intérêt. Cette longueur, qualifiée de *burn-in*, est systématiquement ignorée afin de ne considérer que les échantillons correspondant à des chaînes stationnaires et ne dépendant pas du point de départ arbitraire. La figure 5.3 illustre cette adaptation : le panneau gauche montre le déplacement des marcheurs vers la région d'intérêt, symbolisée par l'étoile noire, et le panneau droit permet d'estimer une longueur de *burn-in* d'environ une centaine de pas.

Le mélange des chaînes quantifie la similarité entre les distributions de points générés par chacune d'entre elles. Des chaînes mélangées seront donc stationnaires et statistiquement similaires, comme illustré dans le panneau droit de la figure 5.3 dans le régime stationnaire. Un critère permettant d'évaluer ce mélange est le critère de Gelman et Rubin [171] :

$$\hat{R} \equiv \sqrt{\frac{V}{W}} < 1.02 \quad (5.9)$$

où W est la variance moyenne des échantillons d'une chaîne et V la variance entre les différentes chaînes.

Les échantillons produits par la marche aléatoire décrite ci-dessus ne sont pas indépendants. En effet, la position de chaque point dépend par construction de la position précédente. On parle alors d'autocorrélation des chaînes. La longueur moyenne séparant deux points pouvant être considérés indépendants peut être calculée d'après la série de positions acceptées formant la chaîne de Markov :

$$\tau = 1 + 2 \sum_{i=1}^n \rho(\vartheta_0, \vartheta_i) \quad (5.10)$$

où n est la longueur de la chaîne considérée, et $\rho(\vartheta_0, \vartheta_i)$ est le coefficient de corrélation entre ϑ_0 et ϑ_i .

À intervalle régulier au cours de l'échantillonnage, PANC02 évalue ces trois critères, et accepte la convergence si :

- Toutes les chaînes ont parcouru une longueur supérieure à cinq fois la longueur de *burn-in* définie par l'utilisateur;
- Plus des deux tiers des chaînes passent le test de Gelman et Rubin (5.9);
- Ces chaînes ont parcouru plus de cinquante fois leur longueur d'autocorrélation.

Les valeurs numériques utilisées dans ces tests sont des paramètres de l'analyse pouvant être modifiés par l'utilisateur. Le choix particulier de deux tiers des chaînes et cinquante longueurs d'autocorrélation, utilisé par défaut, provient des tests sur simulations développés en section 5.3.

5.2.4 Évaluation des propriétés thermodynamiques

Les chaînes de Markov produites par l'échantillonneur constituent un ensemble de points suivant la distribution de probabilité postérieure dans l'espace des paramètres du modèle. On peut les écrire comme un vecteur $\boldsymbol{\vartheta}$, de dimension n le nombre de points dans les chaînes, dont chaque composante $i \in [1 \dots n]$ est un vecteur dans l'espace des paramètres. Cela permet de définir un vecteur de profils de pression $\mathbf{P}(r)$, dont chaque composante i est le profil associé aux paramètres du point $\boldsymbol{\vartheta}_i$ des chaînes de Markov :

$$\mathbf{P}(r) = P_e(r, \boldsymbol{\vartheta}_i), \quad (5.11)$$

où $P_e(r)$ est le modèle de profil de pression choisi, gNFW ou non-paramétrique.

La distribution des composantes de $\mathbf{P}(r)$ permet donc d'étudier la distribution postérieure dans l'espace non plus des paramètres, mais des profils de pression. On peut ainsi calculer l'incertitude sur le profil de pression comme la dispersion des composantes de $\mathbf{P}(r)$, ou encore comme les quantiles de cette distribution (les deux options sont implémentées dans PANC02, et peuvent être choisis par l'utilisateur).

Interpolations du modèle non-paramétrique

Dans le cas d'un modèle non-paramétrique, les paramètres du modèle sont les valeurs du profil de pression à différents rayons choisis par l'utilisateur. La mesure de la pression à un rayon r quelconque nécessite donc une interpolation (ou l'extrapolation) à partir de ces points. PANC02 propose trois méthodes différentes, à choisir par l'utilisateur, pour interpoler le profil de pression pour la présentation des résultats finaux :

- L'interpolation en loi de puissance des points. C'est la méthode utilisée au cours de l'ajustement (voir équation 5.4), et donc la plus naturelle. Cependant, comme nous l'avons décrit en 5.1.1 (page 103), le modèle non-paramétrique est une fonction dont la dérivée n'est pas continue. Le profil de masse hydrostatique étant justement proportionnel à cette dérivée, il peut présenter des discontinuités, rendant difficile l'estimation d'un profil de masse à partir de l'ajustement d'un tel modèle.
- Une interpolation en *splines* du profil de pression. Si cette option est choisie, chacune des composantes du vecteur $\mathbf{P}(r)$ est interpolée par une fonction *spline* dans le plan log-log, résultant en

un profil de pression lisse et de dérivée continue. Cette méthode présente l'avantage de réaliser un lissage du profil, permettant l'estimation d'un profil de masse continu, tout en gardant l'aspect non-paramétrique du modèle utilisé pour l'ajustement. Cependant, son extrapolation est risquée, car les splines ne disposent d'aucune contrainte au delà de la région couverte par les rayons R considérés dans le modèle.

- Une interpolation par un modèle gNFW. Dans ce cas, un modèle gNFW est ajusté sur chacune des composantes de $\mathbf{P}(r)$, et le modèle ajustant le mieux chaque profil est utilisé comme fonction d'interpolation. Les profils obtenus sont lisses et de dérivées continues, et leur extrapolation est rendue possible par le peu de variété dans les formes que peut prendre le modèle gNFW. En revanche, l'identification d'écart à ce modèle n'est plus possible.

Quelle que soit la méthode choisie, elle permet de calculer le vecteur $\mathbf{P}(r)$ pour toute valeur de r choisie. La distribution des composantes du vecteur peut alors être étudiée pour évaluer l'incertitude sur le profil de pression. La différence entre les profils de pression et de masse hydrostatique obtenus par l'utilisation de ces trois méthodes sera détaillée dans la section suivante (page 117), et illustrée en figure 5.9.

Profils thermodynamiques par combinaison avec les mesures X

Une fois que le vecteur $\mathbf{P}(r)$ a été construit, en calculant un profil de pression pour chaque jeu de paramètres issu des chaînes de Markov, celui-ci est combiné aux mesures X pour caractériser plus en détail les propriétés thermodynamiques du milieu intra-amas. Pour cela, le vecteur $\mathbf{P}(r)$ est combiné avec le profil de densité X pour obtenir un vecteur de profils de température \mathbf{T} d'après l'équation (3.11) :

$$k_B \mathbf{T}(r) = \mathbf{P}(r) / n_e(r) \quad (5.12)$$

et d'entropie \mathbf{K} d'après l'équation (3.12) :

$$\mathbf{K}(r) = \mathbf{P}(r) n_e^{-5/3}(r), \quad (5.13)$$

et un vecteur de profils de masse à partir de l'équation de l'équilibre hydrostatique (2.3). Comme pour le profil de pression, la dispersion des composantes de ces vecteurs donne l'incertitude sur les profils. Comme nous l'avons vu en 3.4.3 (page 67), ces quantités permettent l'étude de la structure de l'amas et de son état dynamique, en exploitant la synergie entre observations SZ à haute résolution et X. Une fois mesurées pour tout l'échantillon, elles pourront mener à des travaux sur le lien entre la physique de ces objets et les incertitude systématiques liées à leur exploitation cosmologique.

Grandeurs caractéristiques intégrées

Dans le cas du grand programme SZ de NIKA2, l'objectif est la mesure du profil de pression des amas, mais également de la relation d'échelle liant le paramètre de Compton Y_{500} intégré dans un rayon R_{500} à la masse M_{500} contenue dans ce même rayon. PANCO2 permet de mesurer ces grandeurs en calculant le profil de masse hydrostatique à partir de l'équation de l'équilibre hydrostatique (2.3), puis en déduisant le profil de contraste de densité $\delta_c(r)$:

$$\delta_c(r) = \frac{M_{\text{HSE}}(< r)}{\rho_c(z) \times \frac{4}{3} \pi r^3}. \quad (5.14)$$

où $\rho_c(z)$ est la densité critique de l’Univers au redshift z de l’amas.

La procédure utilisée pour calculer les grandeurs intégrées est la suivante :

1. Le vecteur de profils de masse hydrostatique $\mathbf{M}_{\text{HSE}}(r)$ est utilisé pour calculer un vecteur de profils de contraste de densité par l’équation (5.14) ;
2. Une valeur de $R_{500,i}$ est calculée pour chacune des composantes i du vecteur de profils de contraste, en trouvant le rayon pour lequel il est égal à 500 ;
3. Une valeur de $Y_{500,i}$ est calculée pour chacune des composantes i du vecteur $\mathbf{P}(r)$ par :

$$Y_{500,i} = 4\pi \frac{\sigma_t}{m_e c^2} \int^{R_{500,i}} r^2 P_i(r) dr, \quad (5.15)$$

et une valeur de $M_{500,i}$ à partir des profils de masse hydrostatique :

$$M_{500,i} = M_i(R_{500,i}). \quad (5.16)$$

Le résultat de cette procédure est donc un vecteur pour chacune des grandeurs caractéristiques R_{500} , Y_{500} et M_{500} , dont chaque composante i correspond à la grandeur estimée pour le point Φ_i des chaînes de Markov. La distribution des composantes représente alors la distribution postérieure échantillonnée par le MCMC, dans l’espace des grandeurs intégrées, et peut être utilisée pour calculer les incertitudes entre les paramètres. On note que les grandeurs intégrées sont par construction corrélées, comme déjà mentionné au chapitre 2. Cette méthode d’estimation des grandeurs intégrées permet la mesure de cette corrélation en calculant la covariance des grandeurs intégrées mesurées pour chaque échantillon des chaînes de Markov. La connaissance de cette covariance est cruciale à l’estimation des relations d’échelle, comme nous le verrons au chapitre 7. Les produits de PANC02 pourront donc être utilisés directement pour la mesure de la relation d’échelle $Y_{500} - M_{500}$ par la collaboration NIKA2.

5.3 Validation de PANC02 sur des amas simulés

Afin de s’assurer de la capacité de PANC02 à produire des mesures correctes de profil de pression, il est nécessaire de le tester sur des cartes d’amas dont le profil de pression est connu, et ainsi de comparer le profil évalué avec la vérité. Pour cela, il est possible soit d’utiliser des observations réelles d’amas connus au profil de pression déjà caractérisé par d’autres instruments, soit de s’appuyer sur des simulations. La première approche présente l’avantage de faire appel à des données réelles, mais est limitée par la connaissance imparfaite de l’amas observé. L’analyse de simulations permet de connaître parfaitement les propriétés physiques d’entrée, mais est limitée par la nécessité de créer des données réalistes et représentatives d’observations réelles. La validation de PANC02 emploie la seconde approche, utilisant des amas synthétiques analytiques, puis un échantillon d’amas issu de la simulation hydrodynamique MUSIC [172].

5.3.1 Simulations d’amas simples

La première analyse de validation de PANC02 est réalisée sur une carte d’amas synthétique similaire à l’un des amas du LPSZ. Nous avons choisi l’amas ACT-CL J0215.4+0030, un des amas de faible masse et haut redshift de l’échantillon, dont les observations réelles seront présentées et analysées au chapitre 6.

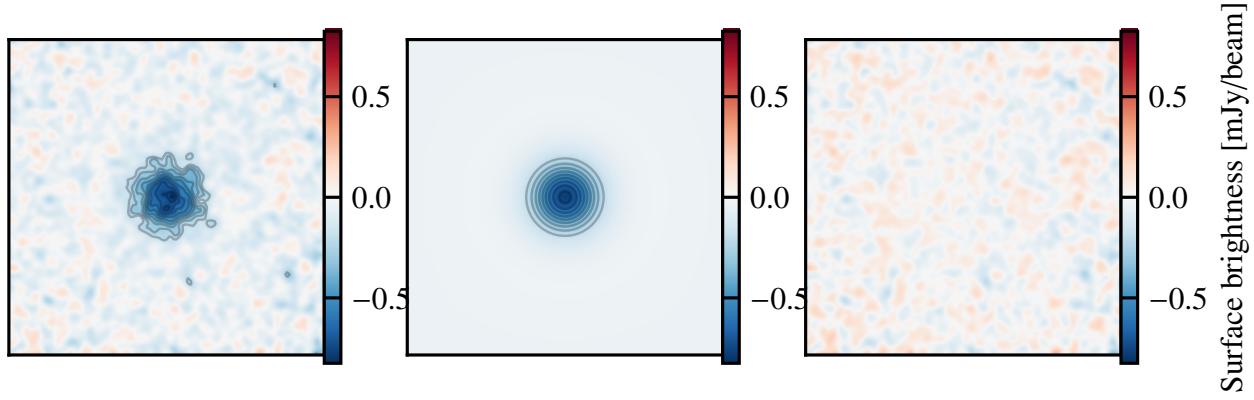


Figure 5.4 : Carte simulée de ACT-CL J0215.4+0030 (*gauche*), carte de modèle ajustée par PANCO2 (*centre*), et différence entre les deux cartes (*droite*). La région représentée est la même pour les trois cartes, et représente un carré de 5' de côté centré sur le centre simulé de l'amas. Les contours gris sont les niveaux de signal-sur-bruit, commençant à 3σ et espacés de 1σ .

Le milieu intra-amas de cette source synthétique est modélisé par un profil de pression gNFW, correspondant à celui mesuré par l'analyse effectuée sur les observations NIKA2 de l'amas [138]. Des cartes de l'amas sont ensuite produites en intégrant ce profil de pression le long de la ligne de visée, et en filtrant la carte obtenue par le lobe de NIKA2 et par la fonction de transfert réelle de l'analyse. Enfin, une réalisation de bruit blanc correspondant au niveau de bruit résiduel dans la carte de ACT-CL J0215.4+0030 est ajoutée à la carte.

Les cartes obtenues sont ajustées avec PANCO2 afin de tester sa validité sur des cas simples d'amas sphériques idéaux. Cette analyse utilise un profil de pression non-paramétrique, dont les valeurs de rayons R sont définies pour couvrir le signal attendu avec NIKA2, de son lobe à son champ de vue. Cette distribution de rayons sera détaillée en 5.3.2 (page 118). Dans cette section, nous décrivons les résultats pour cet amas synthétique, et présentons par la même occasion les figures produites par PANCO2. Notons qu'une analyse similaire a été menée pour l'ajustement d'un modèle gNFW. Les résultats sont très similaires pour les deux modèles (offrant une reconstruction non-biaisée du profil de pression), aussi ne présentons-nous ici que les résultats de l'ajustement d'un modèle non-paramétrique.

Résultats de l'ajustement

Les résultats de l'ajustement permettent d'observer un accord entre le modèle ajusté par PANCO2 et la vérité, utilisée pour générer les cartes. La figure 5.4 montre la carte d'entrée et le meilleur ajustement de cette carte obtenu par PANCO2. La différence entre ces deux cartes, représentant les résidus de l'ajustement, est compatible avec du bruit et ne permet pas d'identifier d'écart entre la vérité et l'ajustement.

Chaînes de Markov et distribution postérieure des paramètres

La figure 5.5 montre les chaînes de Markov, à savoir l'évolution des positions acceptées des marcheurs dans l'espace des paramètres avec le temps. Chaque couleur correspond à un marcheur. On voit que les chaînes ont atteint un état stationnaire, où elles n'évoluent plus à grande échelle et oscillent autour d'une valeur de paramètre correspondant à la valeur maximisant la distribution postérieure. On voit également l'évolution à partir du point de départ, au cours des ~ 200 premiers pas. La longueur

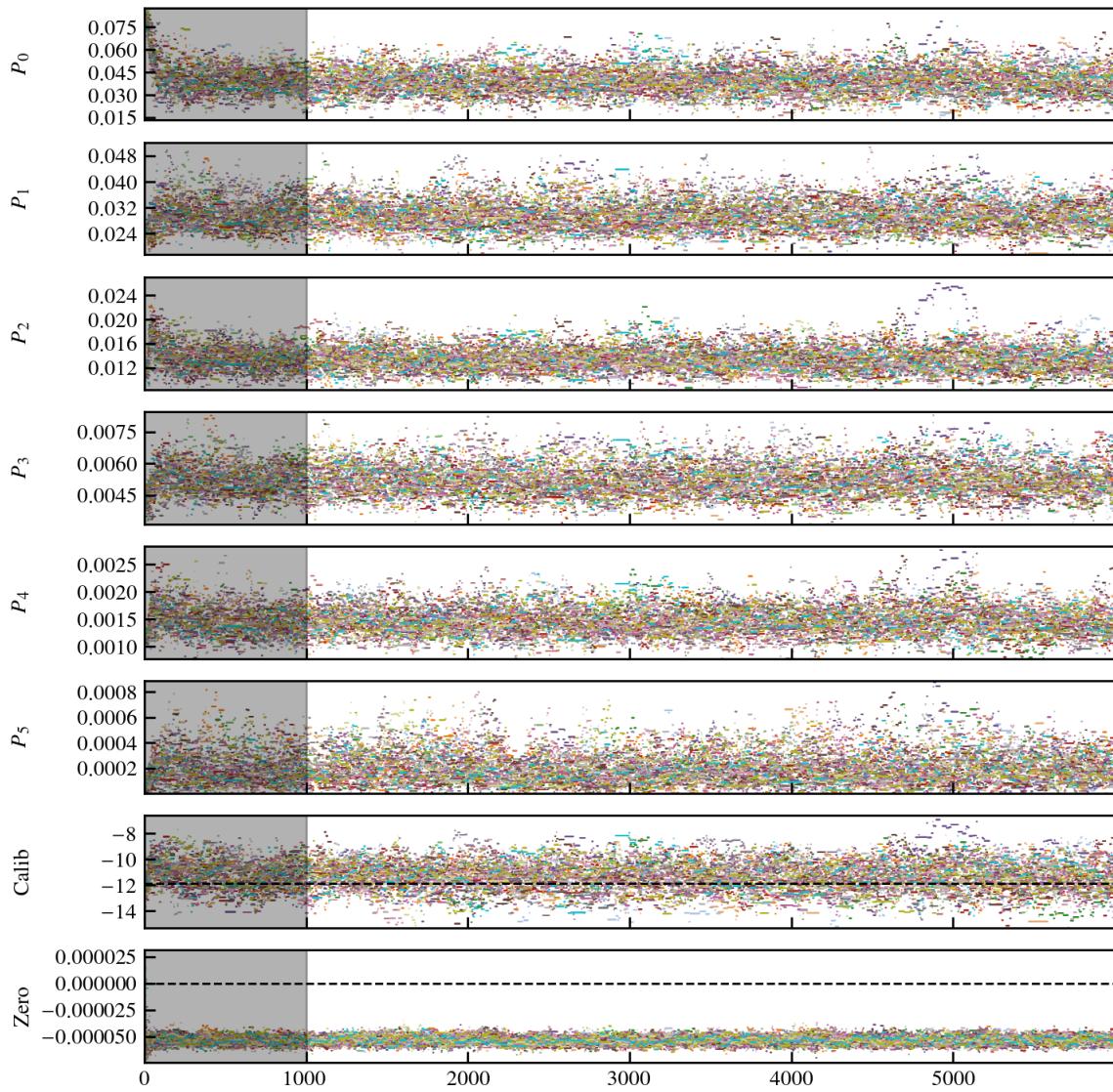


Figure 5.5 : Évolution des chaînes de Markov dans l'espace des paramètres avec le nombre de pas. Chaque couleur représente une chaîne distincte. La région grise correspond au *burn-in* retiré pour l'analyse. Les lignes noires pointillées correspondent au vraies valeurs du coefficient de conversion et du niveau zéro de la carte.

de *burn-in* retirée, représentée en gris, est de 1000 pas.

La figure 5.6 montre la distribution des positions des chaînes de Markov, c'est-à-dire la distribution postérieure, marginalisée sur chacun des paramètres (diagonale) et sur chaque combinaison de deux paramètres (éléments extra-diagonaux). Cette figure représente donc la distribution postérieure échantillonnée dans l'espace des paramètres, qui sont les paramètres du profil de pression P_i , le coefficient de conversion et le niveau zéro de la carte. Cette figure illustre donc à la fois les distributions de probabilité des paramètres et les corrélations entre ces paramètres. On voit par exemple une forte corrélation entre le dernier point en pression P_5 et le niveau zéro de la carte : en effet, ajouter de la pression à très grand rayon, du fait de l'intégration le long de la ligne de visée, revient à ajouter un plateau quasi-constant à la carte. La distribution de χ^2 par degrés de liberté est également représentée, et est centrée en une valeur proche de 1, indiquant un bon ajustement des données par le modèle.

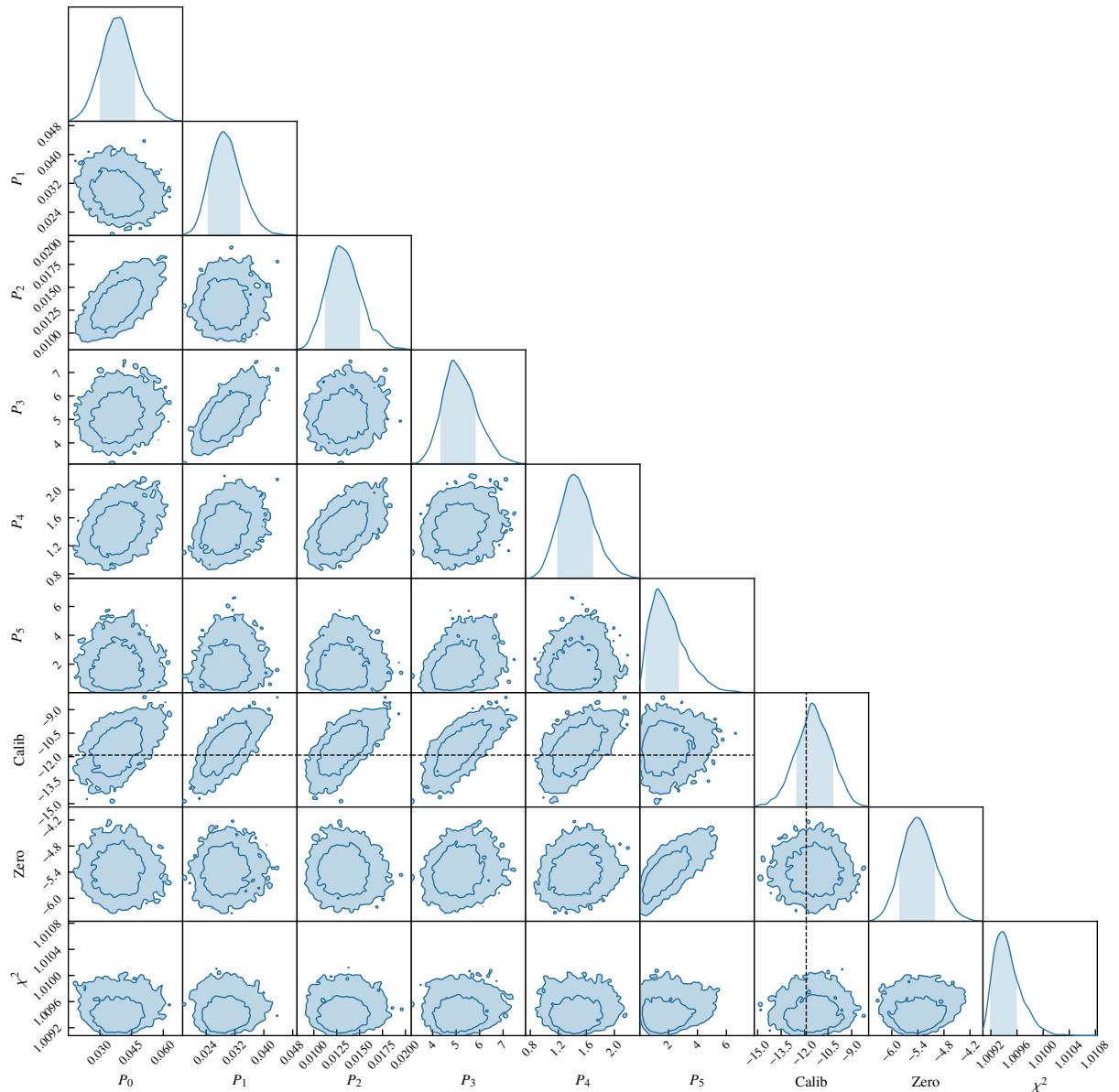


Figure 5.6 : Distribution postérieure dans l'espace des paramètres. Les éléments diagonaux représentent les distributions marginalisées sur chacun des paramètres, et les extra-diagonaux les distribution à deux dimensions pour chaque couple de paramètres. Les contours représentés sont les intervalles de confiance à 1σ et 2σ . La ligne noire pointillée correspond à la vraie valeur du coefficient de conversion.

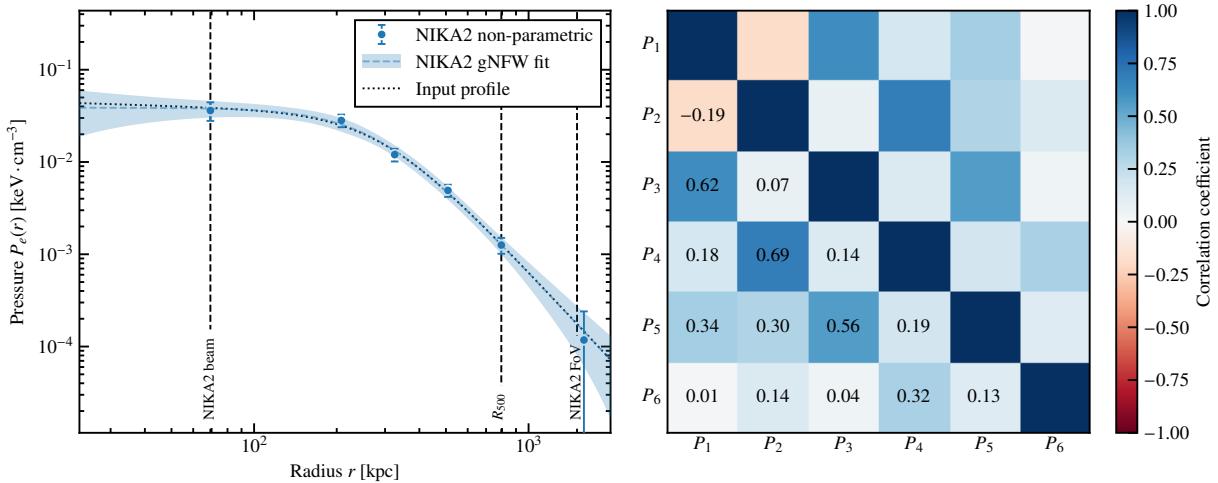


Figure 5.7 : **Gauche** : Profil de pression obtenu avec PANC02. Les points bleus représentent le meilleur ajustement en non-paramétrique. L'enveloppe bleue représente l'incertitude sur ce profil obtenue par interpolation gNFW des résultats non-paramétriques. La courbe pointillée noire est le profil de pression utilisé pour générer la carte de l'amas. **Droite** : Matrice de corrélation entre les points du profil de pression, obtenue à partir des chaînes de Markov.

Profils thermodynamiques

Le profil de pression issu de l'analyse est présenté en figure 5.7. On voit que le profil de pression mesuré par PANC02 (bleu) est en accord avec la vérité (noir) du lobe de NIKA2 jusqu'au delà de R_{500} , ce qui montre le succès de PANC02 pour cet ajustement. L'interpolation utilisée pour estimer l'incertitude sur le profil de pression est l'interpolation gNFW, décrite en section 5.2.4. Les résultats des autres méthodes seront discutés par la suite (page 117). Comme on peut le voir, le résultat est bien l'interpolation des points non-paramétriques par une fonction lisse : dans ce cas, un tel résultat était attendu, puisque le profil d'entrée est une fonction gNFW. La matrice de corrélation entre les points du profil de pression est également représentée. Celle-ci permet d'évaluer l'interdépendance des points du profil de pression, afin de déterminer si la distribution des rayons considérés pour le modèle est adaptée aux données. Deux points trop proches seront en général fortement anti-correlés, traduisant le fait que pour deux points très proches, augmenter l'un revient à diminuer l'autre. La matrice de corrélation sera utilisée pour déterminer la distribution radiale optimale pour des amas synthétiques en section 5.3.2.

Comme nous le verrons au chapitre suivant, ACT-CL J0215.4+0030 a été observé par *XMM-Newton*, et ces observations permettent d'extraire un profil de densité (représenté en figure 6.4). Puisque le profil de pression utilisé pour générer la carte NIKA2 est celui obtenu par ajustement des observations NIKA2 de l'amas réel, nous pouvons combiner ce profil de densité avec le profil de pression mesuré par PANC02 pour représenter les propriétés thermodynamiques de cet amas synthétique. La procédure employée est décrite en 5.2.4 (page 111). Les profils obtenus par cette combinaison sont représentés en figure 5.8. On y trouve les profils de température et d'entropie du milieu intra-amas, ainsi que le profil de masse hydrostatique. Les points bleus représentent les profils obtenus en non-paramétrique. Comme pour le profil de pression, les enveloppes sont obtenues par interpolation gNFW, ce qui permet d'avoir un profil de masse continu.

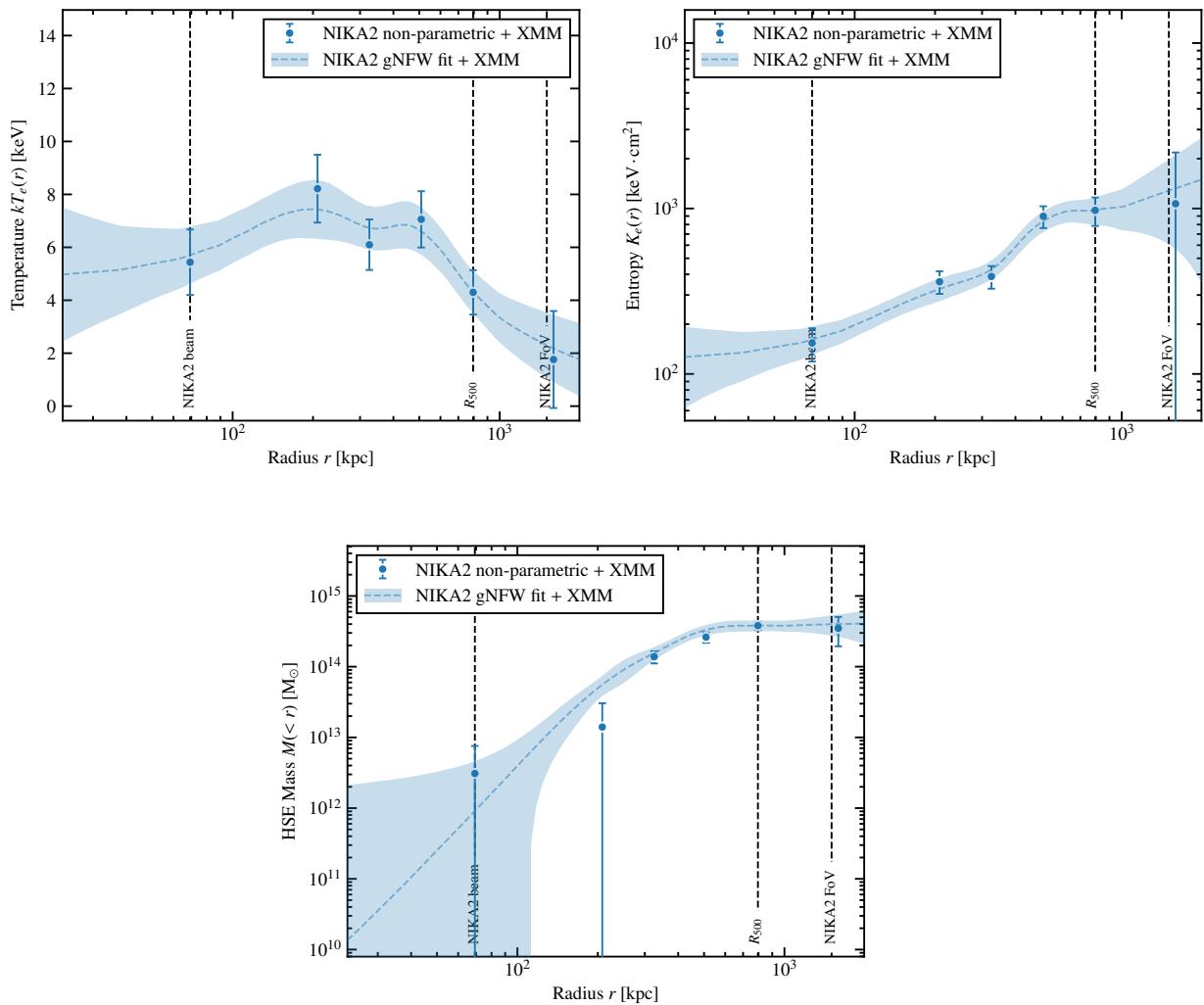


Figure 5.8 : Profils de température (haut gauche), d'entropie (haut droit), et de masse hydrostatique (bas) obtenus par combinaison du profil de pression ajusté par PANCO2 et du profil de densité mesuré en X pour ACT-CL J0215.4+0030. Les points bleus représentent la combinaison du meilleur ajustement du profil de pression en non-paramétrique avec le profil de densité mesuré en X. Les enveloppes bleues représentent l'incertitude sur ces profils obtenue par interpolation gNFW des résultats non-paramétriques.

Effet de la méthode d'interpolation

Les résultats trois interpolations possibles pour le profil de pression, présentées en section 5.2.4, sont présentées en figure 5.9. Les profils de masse correspondants sont également représentés. Les résultats des trois méthodes sont similaires. Dans les différences notables, on note une discontinuité dans le profil de masse à $r \simeq 300$ kpc pour l'interpolation non-paramétrique, due à la nature discontinue de ce modèle. On observe également une incompatibilité entre les profils interpolés en *splines* et les autres méthodes dans le centre de l'amas ($r \lesssim 100$ kpc). Au-delà de 400 kpc, les trois méthodes sont en excellent accord. Le rayon caractéristique R_{500} de cet amas étant de l'ordre de 800 kpc, cette similarité traduit un accord entre les trois méthodes pour l'estimation de la masse caractéristique M_{500} . Le choix de la méthode d'interpolation dépend donc des objectifs scientifiques de l'analyse : la détection de suppressions locales nécessitera une interpolation en loi de puissance, alors que l'interpolation par un profil gNFW sera plus adaptée à l'étude d'un profil de masse continu dans le coeur de l'amas.

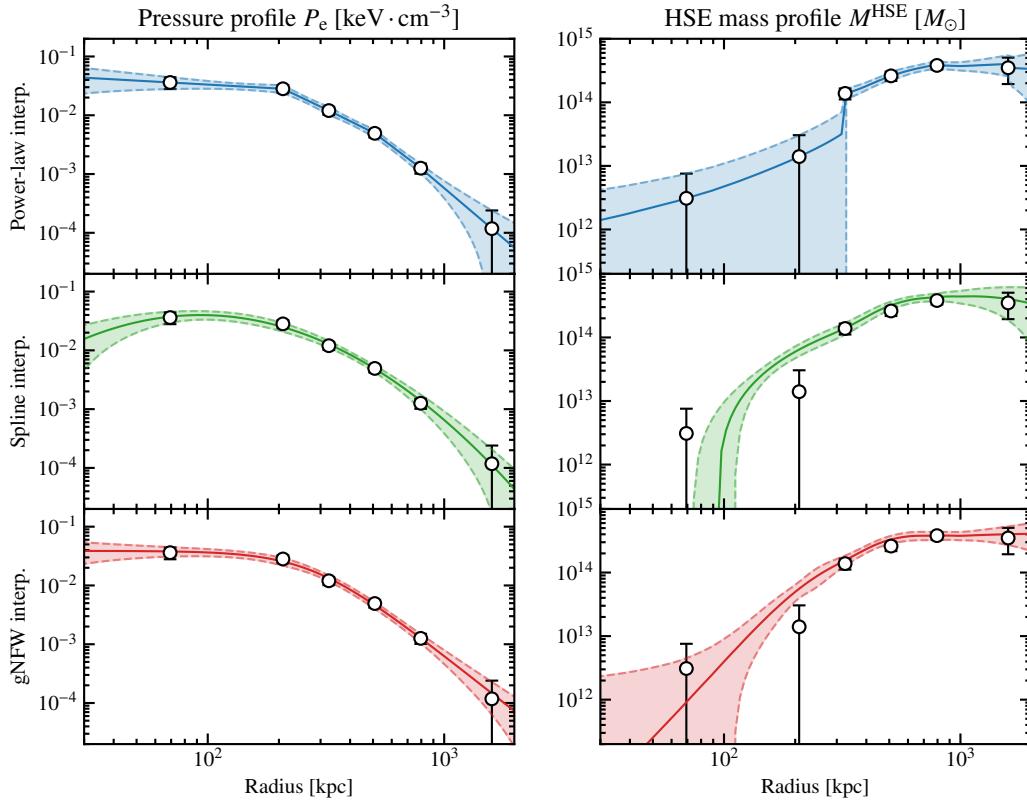


Figure 5.9 : **Gauche :** Interpolations possibles du profil de pression non-paramétrique (points blancs) pour un amas simulé : interpolation en loi de puissance (bleu), en *splines* (vert), et en gNFW (rouge). La ligne continue montre le profil correspondant à l'interpolation du profil de pression ajustant le mieux les données **Droite** : Profils de masse hydrostatique obtenus par combinaison des profils de pression SZ et de densité X pour les trois méthodes d'interpolation. La ligne continue montre le profil correspondant au profil de masse obtenu à partir du profil de pression interpolé à partir du meilleur ajustement. Les régions colorées correspondent aux intervalles de confiance à 1σ sur les profils.

Grandeur caractéristiques intégrées

Enfin, la combinaison des données X et SZ permet, à travers le profil de masse hydrostatique, de calculer les grandeurs intégrées caractéristiques de l'amas. La procédure utilisée est décrite en 5.2.4 (page 111). La distribution de probabilité obtenue par l'application de cette procédure à l'issue de l'ajustement de l'amas synthétique avec PANC02 est présentée en figure 5.10. On y voit la forte corrélation entre les paramètres, de l'ordre de 80% pour $R_{500} - Y_{500}$ et $M_{500} - Y_{500}$, et 95% pour $R_{500} - M_{500}$ (qui sont calculés à partir du même profil de masse). La distribution dans le plan $M_{500} - Y_{500}$ pour chaque amas du LPSZ est une donnée d'entrée pour l'ajustement de la relation d'échelle $Y_{500} - M_{500}$, discutée au chapitre 7.

5.3.2 Optimisation de la distribution radiale

Deux amas synthétiques considérés comme extrêmes pour le grand programme SZ ont été étudiés : un amas compact C1, de basse masse et haut redshift ; et un amas étendu C2, de grande masse et bas

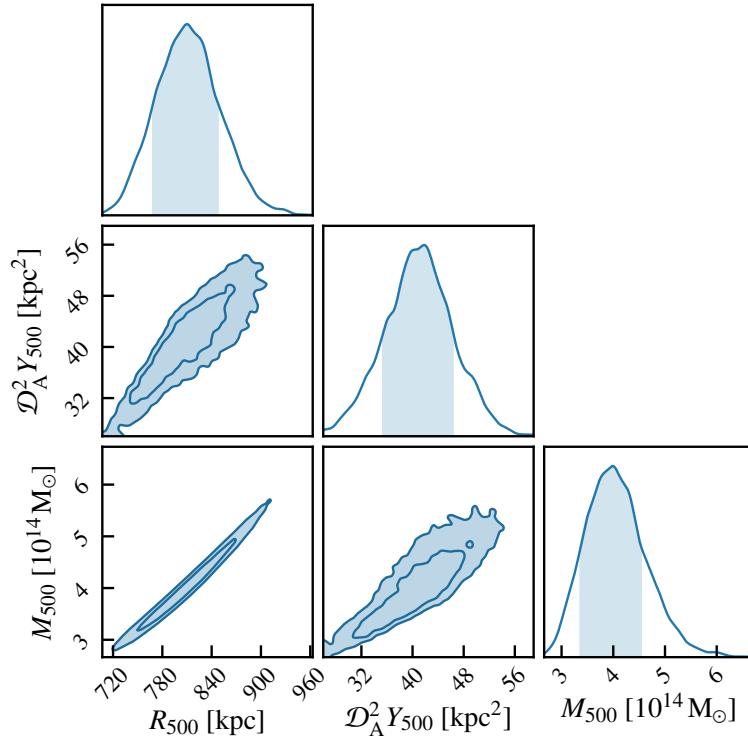


Figure 5.10 : Distributions des grandeurs caractéristiques intégrées R_{500} , Y_{500} , et M_{500} obtenues par l'ajustement avec PANC02 de la carte simulée d'un amas synthétique. Les distributions présentées sont les distributions marginalisées à une dimension (éléments diagonaux) et à deux dimensions (éléments extra-diagonaux). Les contours présentés sont les intervalles de confiance à 1σ et 2σ .

redshift :

$$\begin{aligned} \text{C1 : } & M_{500} = 3.5 \times 10^{14} M_\odot, z = 0.9; \\ \text{C2 : } & M_{500} = 9.5 \times 10^{14} M_\odot, z = 0.5. \end{aligned} \quad (5.17)$$

Les propriétés thermodynamiques de ces amas ont été simulées en considérant une symétrie sphérique, et un profil de pression universel tel que mesuré par Arnaud *et al.* [94]. Comme pour l'étude précédente, des cartes NIKA2 de ces amas sont simulées par intégration le long de la ligne de visée, conversion, filtrage par le lobe et par une fonction de transfert réaliste, puis ajout de bruit blanc réaliste.

Les cartes des amas C1 et C2 ont ensuite été utilisées afin d'optimiser le *binning* du profil de pression, c'est-à-dire le choix des rayons R_i considérés dans le modèle non-paramétrique de l'équation (5.4). Ainsi, chacune de ces cartes a été ajustée pour plusieurs *binning* différents, définis comme suit :

- Le premier rayon R_0 correspond à la demi-largeur à mi-hauteur (HWHM) du lobe de NIKA2, rayon en dessous duquel toute fluctuation de signal est lissée par la résolution de l'instrument ;
- n points sont ensuite ajoutés, espacés logarithmiquement entre $3 \times R_0$ et le rayon R_{500} de l'amas ;
- Un dernier point à $2 \times R_{500}$ pour contraindre la pente externe du profil de pression.

L'ajustement est réalisé avec PANC02 sur les deux cartes d'amas simulées entre $n = 2$ et $n = 10$ dans le but de trouver la valeur de n la plus adaptée aux amas du grand programme SZ. Le choix du meilleur n est basé sur deux critères. Le profil de pression reconstruit par PANC02 doit être aussi proche que

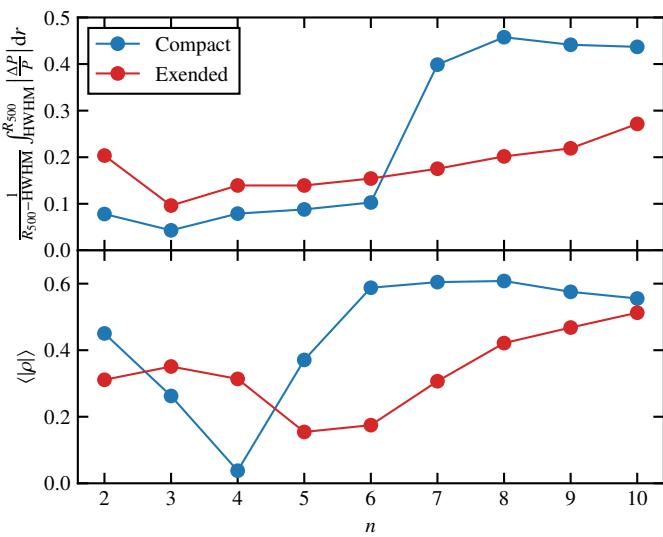


Figure 5.11 : Influence du nombre de rayons considérés dans le modèle non-paramétrique sur les résultats de PANC02. n est le nombre de points situés entre trois fois la demi-largeur à mi-hauteur du lobe de NIKA2 et le rayon caractéristique R_{500} de l'amas (inclus). Les amas considérés sont des amas synthétiques ayant un profil de pression universel, considérant un amas compact (C1, bleu) et un amas étendu (C2, rouge) pour le grand programme SZ de NIKA2 (équation 5.17). Le panneau du *haut* représente l'évolution de la différence relative entre le résultat de PANC02 et la vérité intégrée sur la plage de rayons considérés dans le modèle. Le panneau du *bas* représente l'évolution de la corrélation moyenne entre deux points successifs du profil de pression.

possible du profil réel utilisé en entrée de la simulation ; et la corrélation entre deux points successifs du profil de pression (P_i, P_{i+1}) doit être minimale, symbolisant des mesures indépendantes du profil de pression à différents rayons.

Les résultats sont représentés sur la figure 5.11. Ils montrent qu'un amas étendu peut être modélisé avec un plus grand nombre de rayons qu'un amas compact, puisque l'extension de l'amas est plus grande par rapport au lobe de NIKA2 et dispose d'une meilleure couverture angulaire. Pour l'amas compact C1, on voit qu'il est raisonnable de considérer jusqu'à $n = 6$ points entre trois fois la demi-largeur à mi-hauteur du lobe de NIKA2 et le rayon caractéristique R_{500} de l'amas avant que les résultats ne diffèrent grandement de la vérité. D'autre part, on observe que la corrélation moyenne entre deux mesures de pressions successives augmente rapidement à partir de $n = 4$. Pour l'amas C2, les conclusions sont similaires : la justesse des résultats ne se dégrade que graduellement avec n , mais la corrélation entre deux points successifs augmente à partir de $n = 5$. Ces conclusions nous montrent que, quel que soit le critère choisi, les valeurs de n optimales pour les amas les plus étendus et les plus compacts du grand programme SZ ne sont différentes que d'un point. Par la suite, nous choisirons $n = 5$, apparaissant comme la plus grande valeur utilisable avant que la corrélation des mesures de pression ne dépasse les 50% pour les amas les plus compacts du programme.

5.3.3 La simulation MUSIC et l'échantillon jumeau du grand programme SZ

La simulation MUSIC (*Marenostrum-MultiDark SImulations of galaxy Clusters*, [172]) est une simulation hydrodynamique d'amas de galaxies. Elle est basée sur la technique de la resimulation par *zoom-in* [173], dont le principe de fonctionnement est rappelé ici. Une simulation cosmologique à N -corps à relativement basse résolution est réalisée pour simuler l'évolution de l'Univers, à partir de conditions initiales à haut redshift jusqu'à $z = 0$. Les résultats de cette simulation sont utilisés pour trouver les zones correspondant aux pics de densité de la toile cosmique formée, régions qui abritent les amas de galaxies dans l'Univers. Ces régions sont ensuite resimulées, en repartant des mêmes conditions initiales, mais en ne considérant qu'un plus petit volume, correspondant à la région de l'amas seulement, ce qui diminue le temps nécessaire à la production des simulations. Cette diminution de volume permet donc d'ajouter des ingrédients à la simulation, comme de la nouvelle physique, ainsi

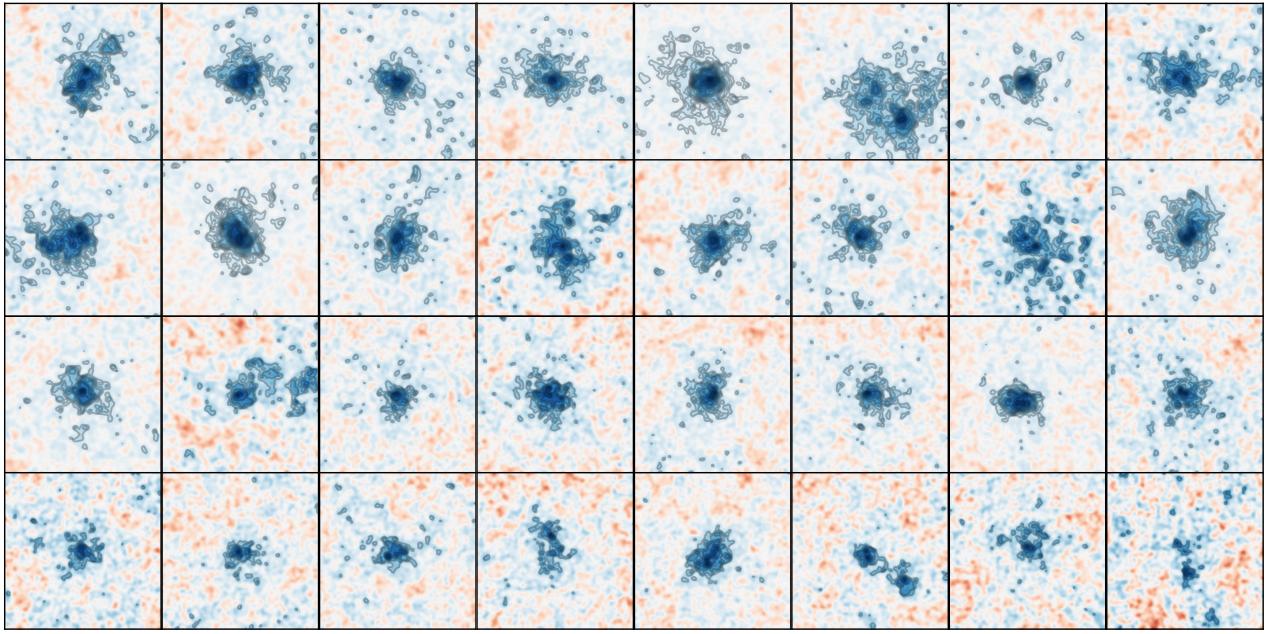


Figure 5.12 : Cartes simulées des 32 amas de l'échantillon utilisé pour la validation du code. Les échelles de couleurs sont arbitraires. Chaque carte est représentée dans un carré de $6.5''$ de côté, centré aux coordonnées du pic de densité du halo considéré. Les contours représentent les niveaux de signal sur bruit commençant à 3σ espacés de 1σ . Les cartes sont lissées par un noyau gaussien de largeur à mi-hauteur $10''$ pour des raisons visuelles.

qu'un plus grand nombre de particules, augmentant ainsi la résolution finale de la simulation.

Dans le cas de MUSIC, les simulations d'entrée sont *Marenostrum* [174] et *MultiDark* [175]. Un total d'environ 500 halos sont resimulés en incluant des phénomènes hydrodynamiques, de transfert radiatif et de formation stellaire. Les amas resimulés offrent donc une physique bien plus riche que les halos d'origine, en plus d'une résolution un ordre de grandeur plus élevée. Chaque amas est resimulé à partir de conditions initiales à un redshift $z = 9$ jusqu'à $z = 0$, et des *snapshots* sont enregistrés à plusieurs redshifts intermédiaires afin de pouvoir étudier l'évolution des amas au cours du processus de formation des structures.

L'échantillon considéré pour la validation de PANC02 est un sous-ensemble de l'échantillon MUSIC sélectionné pour être similaire à l'échantillon du grand programme SZ de NIKA2. Cet échantillon, créé par Ruppin *et al.* [176], comporte 32 amas, couvrant une gamme de masse similaire à celle couverte par le grand programme SZ, à des redshifts de 0.54 et 0.82. Les amas ont été sélectionnés afin d'avoir autant d'amas morphologiquement relaxés que d'amas perturbés. Cet échantillon est particulièrement adapté à la validation de PANC02 : il est composé d'amas réalistes similaires à ceux du grand programme SZ de NIKA2. Des observations NIKA2 simulées sont produites pour ces amas selon la procédure décrite dans [176]. La pression – connue – du milieu intra-amas est intégrée le long d'une ligne de visée arbitraire afin d'obtenir une carte de paramètre de Compton, qui est ensuite convertie en unités de brillance de surface. Celle-ci est ensuite convoluée par le lobe instrumental de NIKA2 et par une fonction de transfert fiduciale réaliste, et une réalisation de bruit corrélé est ajoutée suivant un spectre de puissance de bruit réaliste. Les cartes résultantes sont donc des simulations réalistes d'observations avec NIKA2 des amas de l'échantillon. Elles sont représentées sur la figure 5.12. On remarque une grande variété dans les morphologies d'amas, ainsi que dans le rapport signal sur bruit des observations.

5.3.4 Ajustement du profil de pression

Le profil de pression de chacun des amas de l'échantillon est estimé à l'aide de PANC02. Un modèle non-paramétrique est considéré : les profils de pression des amas MUSIC étant réalistes, toutes leurs caractéristiques ne peuvent pas être capturées par un modèle gNFW, rendant la comparaison d'un profil de pression ajusté par un tel modèle au vrai profil inappropriée. Nous choisissons le même *binning* que celui décrit dans la section 5.3.2 avec $n = 5$ points entre la résolution de NIKA2 et R_{500} . La fonction de transfert et la matrice de covariance du bruit utilisées pour générer les cartes simulées sont considérées dans l'analyse.

La convergence du MCMC est atteinte en environ une à deux heures par amas, représentant un gain d'un facteur ~ 40 par rapport à PANCO, qui réalisait l'ajustement des mêmes cartes en deux à trois jours par amas d'après Ruppin *et al.* [176]³. Il est à noter qu'en considérant le bruit comme blanc – c'est à dire en négligeant les éléments extra-diagonaux de la matrice de covariance du bruit – PANC02 est capable de réaliser l'ajustement en cinq à quinze minutes par amas seulement, soit environ dix fois plus vite. Ce constat motive grandement la poursuite du travail sur la soustraction du bruit corrélé dans les cartes NIKA2, afin d'obtenir des spectres de puissance de bruit résiduel les plus plats possibles et de pouvoir négliger les corrélations pixel-à-pixel du bruit dans les cartes. Ce gain de temps permettrait de mener des études systématiques plus approfondies, comme l'évolution des résultats avec le *bining* du profil de pression non-paramétrique sur l'ensemble des données du grand programme SZ de NIKA2.

Les résultats montrent un excellent accord entre les profils mesurés par PANC02 et ceux extraits de la simulation MUSIC. Trois exemples sont présentés dans la figure 5.13, illustrant la compatibilité entre le profil de pression réel et les résultats de PANC02 pour trois des amas de l'échantillon MUSIC. On y voit un excellent accord dans les barres d'erreur entre les profils de pression ajustés par PANC02 et les données d'entrée sur toute la couverture radiale considérée dans l'ajustement, du lobe de NIKA2 à $2 \times R_{500}$, soit hors du champ de vue pour les amas considérés. Cet accord permet de conclure que les résultats de PANC02 sont satisfaisants, même lorsque le logiciel est confronté à des données réalistes affectées des bruits et filtrages caractéristiques des observations d'amas avec NIKA2.

5.4 Conclusions

L'observation des amas de galaxies du grand programme SZ de NIKA2 permet de mesurer la distribution de pression dans le milieu intra-amas individuellement pour chaque source. L'approche choisie par la collaboration NIKA2 est le *forward modelling*, permettant une prise en compte des différents phénomènes affectant les observations de l'effet Sunyaev-Zeldovich avec NIKA2. Ce chapitre a présenté l'algorithme développé pour effectuer ces mesures en utilisant une procédure MCMC pour ajuster un modèle de profil de pression sur les données NIKA2.

Au début de ma thèse, cet algorithme était implémenté dans le logiciel PANCO, alors pipeline officiel du grand programme SZ de NIKA2. Ce chapitre a présenté le développement de PANC02, dont l'objectif était la livraison d'un logiciel offrant plus de possibilités d'analyse tout en étant plus rapide. Le désir d'optimisation du temps d'exécution provenait du fait que PANCO ne permettait l'analyse d'un échantillon de plusieurs dizaines d'amas qu'au prix de plusieurs semaines de calcul, limitant l'exploitation des données du grand programme SZ de NIKA2. L'analyse de cartes synthétiques d'un échantillon d'amas issus de la simulation hydrodynamique MUSIC a permis à la fois d'affirmer la capacité de

³Il est important de préciser que PANCO a été utilisé dans [176] pour ajuster simultanément des cartes simulées NIKA2 et *Planck* des amas MUSIC, alors que l'analyse PANC02 décrite ici n'a considéré que les cartes NIKA2.

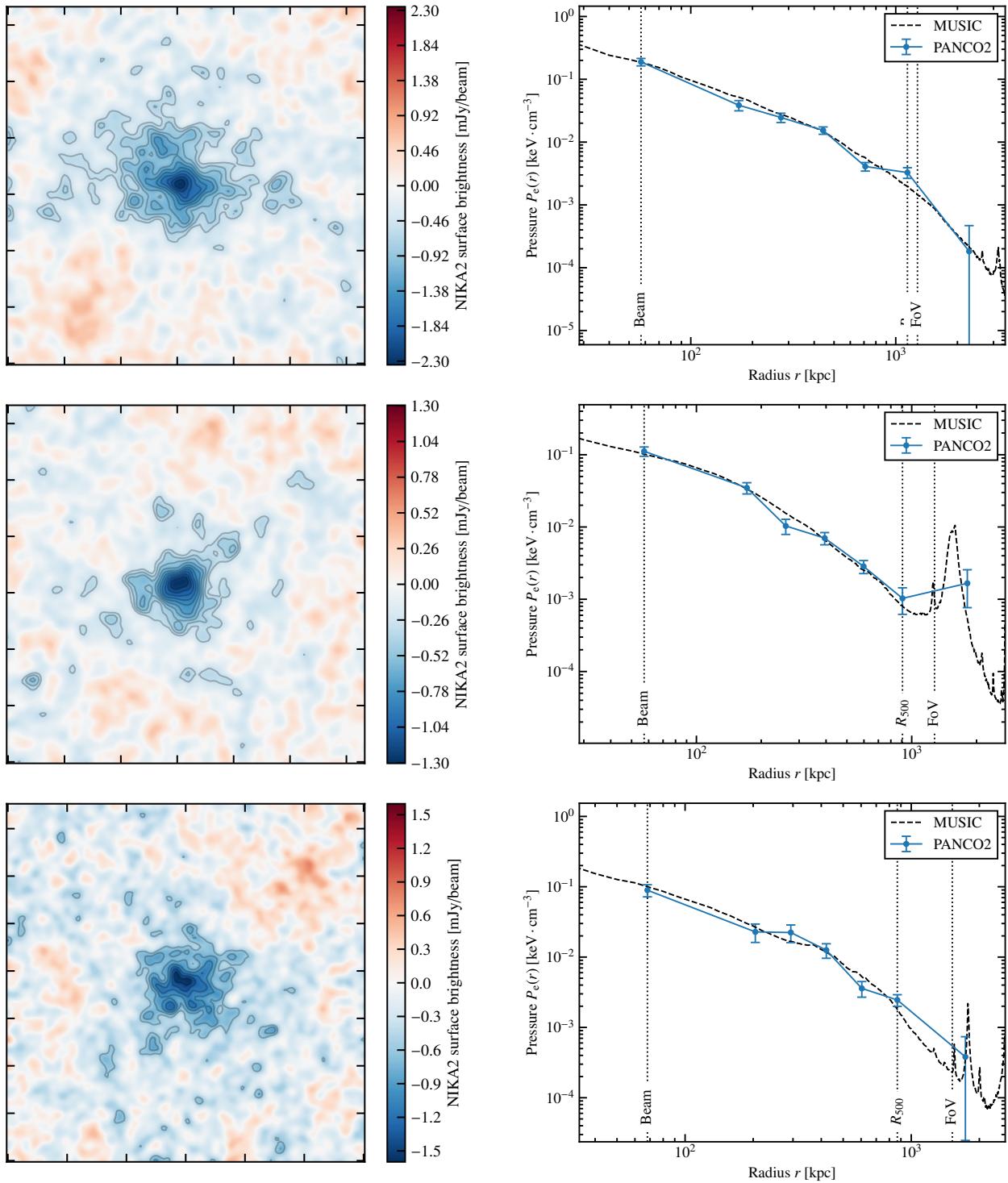


Figure 5.13 : Illustration des cartes NIKA2 simulées (gauche) et profils de pression mesurés (droite) pour trois amas MUSIC : deux à $z = 0.54$ (haut, milieu) et un à $z = 0.82$ (bas). La légende des cartes est la même que pour la figure 5.12. Les lignes pointillées noires représentent les profils de pression réels des amas tels qu’extraits de la simulation MUSIC. Les profils bleus sont ceux obtenus par ajustement avec PANCO2. Tous les profils sont centrés sur le pic de potentiel gravitationnel des amas tels que mesurés dans la simulation MUSIC.

PANCO2 à estimer correctement le profil de pression d'amas de galaxies à partir de données NIKA2 réalistes, mais également de comparer ses performances à celles de PANCO. Le temps d'exécution est diminué par un facteur d'environ quarante, rendant possible non seulement l'analyse d'échantillons d'amas en un temps raisonnable, mais également des études systématiques afin de tirer des informations plus fines sur les propriétés du milieu intra-amas à partir des observations NIKA2. Ceci a mené au choix de PANCO2 comme pipeline officiel du grand programme SZ de NIKA2, et à la préparation d'une mise à disposition de PANCO2 pour la communauté scientifique.

CHAPITRE 6

Mesure des propriétés thermodynamiques d'un amas distant de faible masse

6.1	Observations de ACT-CL J0215.4+0030	126
6.1.1	Détection de l'amas par l'Atacama Cosmology Telescope	126
6.1.2	Observations de l'amas avec NIKA2	126
6.1.3	Observations de l'amas avec <i>XMM-Newton</i>	128
6.2	Contamination par des sources ponctuelles	128
6.2.1	Identification des sources ponctuelles	128
6.2.2	Estimation de la contamination	129
6.2.3	Résultats	130
6.3	Mesures des propriétés thermodynamiques du milieu intra-amas	131
6.3.1	Modélisation du milieu intra-amas	133
6.3.2	Analyse MCMC	133
6.3.3	Auto-consistance des estimations de flux des sources ponctuelles	135
6.4	Résultats	135
6.4.1	Thermodynamique du milieu intra-amas	136
6.4.2	Grandeurs intégrées	138
6.5	Conclusions	139

Les observations des amas de galaxies du grand programme SZ de NIKA2 ont commencé en 2017, avec l'observation de l'amas PSZ2-G144.83+25.11. Cet amas a bénéficié d'un long temps d'observation dans le cadre de la phase de vérification scientifique de NIKA2. Par conséquent, les cartes de cet amas offrent une détection hautement significative de l'effet SZ, jusqu'à 13σ . Cette grande qualité de données a permis une analyse certes profonde et détaillée des propriétés thermodynamiques du milieu intra-amas [137], mais qui n'est pas représentative des observations standard du LPSZ.

Ce chapitre présente la deuxième analyse individuelle d'un amas de galaxies dans le LPSZ. Cette analyse présente plusieurs intérêts. D'une part, elle constitue la première analyse de données de qualité

standard pour le LPSZ. D'autre part, elle se focalise sur ACT-CL J0215.4+0030, un amas distant et de faible masse, ayant pour conséquence un signal SZ faible et peu étendu. Les cartes NIKA2 de cet amas sont de plus fortement contaminées par des sources ponctuelles, dont le flux compense une proportion non négligeable du signal SZ de l'amas (comme nous le verrons en page 128). Cet amas représente donc un banc d'essai pour les analyses des sources les plus complexes du grand programme SZ. Cette analyse et ses résultats sont publiés dans Kéruzoré *et al.*, A&A 644, A93, 2020 [138].

6.1 Observations de ACT-CL J0215.4+0030

6.1.1 Détection de l'amas par l'Atacama Cosmology Telescope

L'amas ACT-CL J0215.4+0030 a été découvert par le premier relevé SZ de l'Atacama Cosmology Telescope (ACT), à un redshift de 0.865 [136, 177]. Il y a été détecté avec une signifiance de 5.5σ . Faute d'avoir une résolution angulaire suffisante pour résoudre des amas distants, le catalogue ACT reporte pour cet amas uniquement une mesure de signal SZ intégré, Y_{500} . La masse de cet amas, estimée grâce à la relation d'échelle $Y_{500} - M_{500}$ publiée par Arnaud *et al.* [94], est de $M_{500} = 3.8 \times 10^{14} M_\odot$.

6.1.2 Observations de l'amas avec NIKA2

Les observations NIKA2 de ACT-CL J0215.4+0030 ont eu lieu lors de la quatorzième campagne d'observations de NIKA2. Elles sont réparties sur cinq jours, entre le 17 et le 22 janvier 2018. Toutes les observations de l'amas ont été effectuées entre 16h et 22h15 (UT). La focalisation du télescope a été vérifiée au début de chaque série d'observations. Le pointage du télescope et le lobe de NIKA2 ont quant à eux été vérifiés toutes les heures au cours des observations. L'amas a été observé 8.6 heures au total, dans des conditions stables, avec une opacité moyenne à 225 GHz de $\tau_{225} = 0.175 \pm 0.05$. Les distributions de l'opacité et de la transmission atmosphérique au cours des observations sont représentés sur la figure 6.1, montrant une grande proportion d'observations réalisées dans de mauvaises conditions (à des élévations inférieures à 30° , ou à des transmissions atmosphériques inférieures à 0.75). Puisque les objectifs de rapport signal sur bruit du LPSZ demandaient 9 heures d'observations avec une opacité¹ de $\tau_{225} = 0.1$, un bruit plus élevé que l'objectif est attendu dans les cartes NIKA2.

Les observations ont été réalisées sous forme de plusieurs séries de quatre scans OTF, décrits en section 3.3.1 de $8' \times 4'$, chacun composé de 25 *subscans*, avec des angles de 0, 45, 90 et 135 degrés par rapport à l'axe des ascensions droites. Le centre de ces scans a été choisi aux coordonnées de l'amas reportées par le catalogue ACT [177], c'est-à-dire $(\text{RA}, \text{Dec})_{J2000} = (02h\ 15m\ 28.8s\ +00^\circ 30' 33.0'')$

Le pipeline de la collaboration NIKA2 décrit dans le chapitre 4 a été utilisé pour le traitement des données brutes de NIKA2. La décorrélation du signal astrophysique et des différentes sources de bruit a été réalisée en utilisant la méthode *Common mode one block*, décrite dans le chapitre 4, ainsi que dans Perotto *et al.* [11] sous le nom de *most correlated pixels*. Le masque utilisé pour le calcul du mode commun a été défini comme un disque de diamètre $2'$ au centre de la carte. Les itérations sur le masque basées sur des niveaux de signal sur bruit n'ont montré aucune amélioration des cartes, soustrayant au fur et à mesure des itérations plus de signal sans diminuer le niveau de bruit. Par conséquent, la carte obtenue avec un masque circulaire a été retenue pour le reste de l'analyse.

La fonction de transfert, quantifiant le filtrage dû à la soustraction du bruit, ainsi que le spectre de puissance du bruit résiduel, sont produits automatiquement par le logiciel de réduction des données

¹correspondant à une transmission atmosphérique supérieure à 0.8, voir section 3.1.1 (page 50).

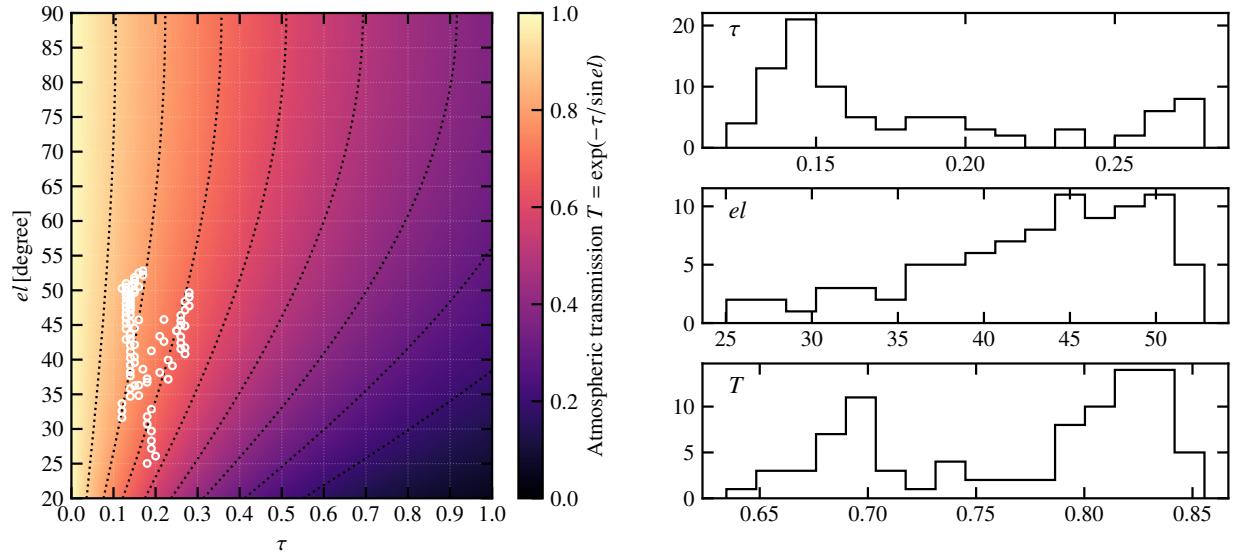


Figure 6.1 : Gauche : transmission atmosphérique T en fonction de l’opacité τ et de l’élévation el . Chaque point blanc représente un scan de ACT-CL J0215.4+0030. Droite : Distribution des opacités (haut), élévations (centre), et transmissions atmosphériques (bas) pour les scans de ACT-CL J0215.4+0030.

brutes utilisant le pipeline de la collaboration NIKA2 (cf. section 4.4). Ils sont tous deux présentés dans la figure 6.2. Il y apparaît que le signal est bien préservé aux petites échelles angulaires ($k > 0.5 \text{ arcmin}^{-1}$), et que le bruit résiduel est fortement corrélé.

Les cartes NIKA2 utilisées pour l’analyse sont présentées dans la figure 6.3. Dans la partie gauche, correspondant à la carte NIKA2 à 150 GHz, l’amas peut être identifié comme un décrément au centre de la carte, caractéristique de l’effet tSZ à des fréquences inférieures à 217 GHz. Le maximum de ce décrément est détecté avec un rapport signal sur bruit de 8.5σ . On remarque dans cette carte un haut niveau de bruit résiduel, se manifestant par un gradient dans le sens sud ouest – nord est.

La carte NIKA2 à 150 GHz montre également une indication de contamination par des sources ponctuelles. Celles-ci émettent un flux positif qui peut compenser le décrément de l’effet SZ, et créer des trous² dans la forme apparente du milieu intra-amas. Une telle contamination est particulièrement inquiétante dans le cas de cet amas. En effet, étant donné son faible signal (moins d’un mJy/beam au maximum) et sa petite extension spatiale (environ dix fois l’angle solide du lobe de NIKA2), la moindre contamination peut avoir de grandes conséquences sur les propriétés physiques reconstruites.

Dans la partie droite de la figure 6.3, correspondant à la carte NIKA2 à 260 GHz, aucun signal SZ n’est détecté. Le signal SZ positif attendu à ces fréquences ($\sim 0.2 \text{ mJy/beam}$) est inférieur au niveau de bruit au centre de la carte par un facteur quatre. En revanche, cette carte nous permet de confirmer la présence de sources ponctuelles proches du centre de l’amas, correspondant aux positions auxquelles des trous sont observés dans le signal SZ de l’amas. Cela indique qu’une prise en compte soigneuse de la contamination par ces sources sera nécessaire à la mesure des propriétés thermodynamiques du milieu intra-amas.

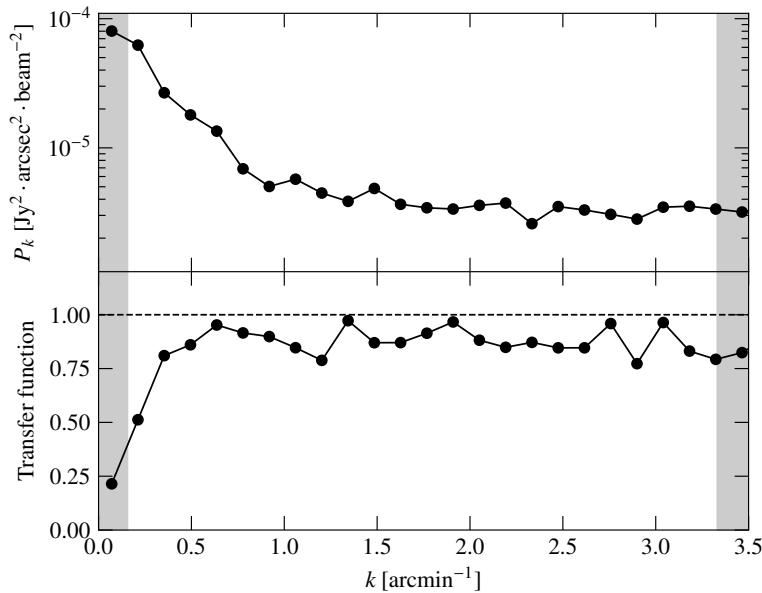


Figure 6.2 : Spectre de puissance du bruit résiduel (*haut*) et fonction de transfert (*bas*) correspondant à l’analyse des observations de ACT-CL J0215.4+0030 avec NIKA2 La procédure d’obtention est décrite dans 4.4. Le spectre de puissance montre un bruit résiduel corrélé à grande échelle, et la fonction de transfert montre le filtrage du signal aux grandes échelles angulaires.

6.1.3 Observations de l’amas avec *XMM-Newton*

ACT-CL J0215.4+0030 a été observé par *XMM-Newton* pendant 37 ks. Les données ont été récupérées depuis l’archive XMM, et les procédures standard (voir, par exemple, [178]) ont été suivies pour produire des fichiers étalonnés répertoriant les évènements enregistrés. Les données ont été corrigées des effets de vignettage et du signal de fond. Les sources ponctuelles ont été détectées et exclues. Après ce nettoyage, le temps d’observation était de 35 ks sur les caméras MOS, et de 29 ks sur la caméra pn.

La carte lissée par ondelettes montrée sur la figure 6.4 montre que le milieu intra-amas semble perturbé. Les profils de densité et température électroniques corrigées de la PSF ont été produits à partir des données de *XMM-Newton* en utilisant la déprojection décrite dans [111]. Au vu de la nature possiblement perturbée de l’amas, ces profils extraits aux pics X et SZ ont été comparés, mais aucune différence significative n’a été détectée. Les profils 3D extraits au pic SZ sont montrés dans la partie droite de la figure 6.4. Le profil de température projeté, extrait directement à partir des spectres 2D, est aussi montré pour référence, mais seul le profil 3D est utilisé dans le reste de l’analyse.

6.2 Contamination par des sources ponctuelles

Comme nous l’avons vu précédemment, une estimation précise de la contamination par des sources ponctuelles est nécessaire dans l’objectif de mesurer les propriétés du milieu intra-amas. Cette étude est cruciale dans le cas de cet amas, étant donné son faible signal et sa faible extension spatiale. Dans cette section, nous présentons la méthodologie utilisée pour estimer l’amplitude de cette contamination.

6.2.1 Identification des sources ponctuelles

Dans le cas de ACT-CL J0215.4+0030, quatre sources sont identifiées avec un rapport signal sur bruit supérieur à 3 dans la carte NIKA2 à 260 GHz. Leurs flux sont plus élevés à 260 GHz qu’à 150

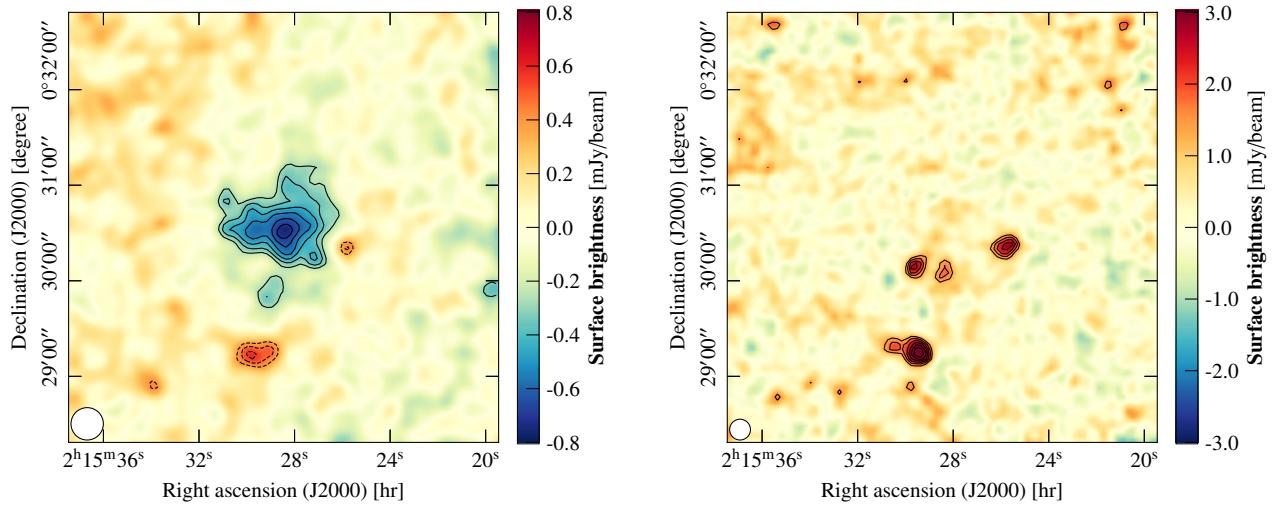


Figure 6.3 : Cartes NIKA2 du champ de l’amas à 150 GHz (gauche) et 260 GHz (droite). Chaque carte est montrée dans un champ carré de $4.5' \times 4.5'$ centré sur les coordonnées observées. Dans chacune des cartes, les contours noirs montrent les niveaux de signal sur bruit commençant à 3σ et espacés de 1σ . Les cartes de gauche et droite sont lissées avec un noyau Gaussien de $10''$ et $6''$ respectivement pour des raisons visuelles. Pour chaque carte, la FWHM de sa résolution effective est illustrée par un disque blanc en bas de la carte.

GHz, favorisant l’hypothèse de sources sub-milimétriques (SMG) plutôt que radio. Cette hypothèse est confirmée par l’identification de ces sources dans le catalogue HerS82 [179], construit grâce à l’instrument SPIRE à bord du satellite *Herschel* qui a observé le ciel dans le domaine sub-millimétrique. Ce catalogue nous permet également d’identifier une cinquième source dans le nord-est de l’amas.

La figure 6.5 montre une carte composite multifréquence de la région de l’amas. Les cercles verts sont centrés sur les positions des cinq SMG identifiées dans le catalogue HerS82, et le diamètre de chaque cercle correspond à la FWHM du lobe instrumental de SPIRE à $250 \mu\text{m}$, bande à laquelle la résolution angulaire de SPIRE est la meilleure. Les régions rouges représentent les niveaux de rapport signal sur bruit dans la carte NIKA2 à 260 GHz lorsque celui-ci est supérieur à 3σ . NIKA2 et SPIRE sont en accord pour les positions des trois sources les plus proches de l’amas : de gauche à droite, SMG1, SMG2, et SMG4. Dans le nord-est de l’amas, on trouve SMG3, qui est identifiée dans le catalogue HerS82 mais a un signal sur bruit inférieur à 3σ dans NIKA2 à 260 GHz. Dans le sud de l’amas, deux sources semblent être résolues par NIKA2, mais seulement la plus faible des deux (S5) est identifiée dans le catalogue HerS82. Il est par conséquent difficile d’estimer l’amplitude de leur contamination du signal SZ, puisqu’il n’existe pas de moyen de déterminer si le flux de S5 dans les bandes SPIRE provient d’une seule source ou de la somme des deux détectées dans NIKA2. Ces sources étant jugées suffisamment éloignées de l’amas et de la zone dans laquelle le signal SZ est significatif, elles sont soustraites de la carte NIKA2 à 150 GHz dans la suite de l’analyse.

6.2.2 Estimation de la contamination

Il est important de noter dans cette analyse l’absence d’information sur l’émission des sources à des fréquences supérieures à celles des bandes passantes de SPIRE ; par exemple, les études précédentes d’observations d’amas avec NIKA et NIKA2 tirent parti de l’existence de données publiques de l’instrument PACS [155], qui ne sont pas disponibles ici. Nous utilisons le logiciel PSTools, présenté en section 4.5 (page 93), pour estimer l’amplitude de la contamination. Le flux de chacune des sources est mesuré dans la carte NIKA2 à 260 GHz en ajustant un modèle de lobe instrumental de largeur

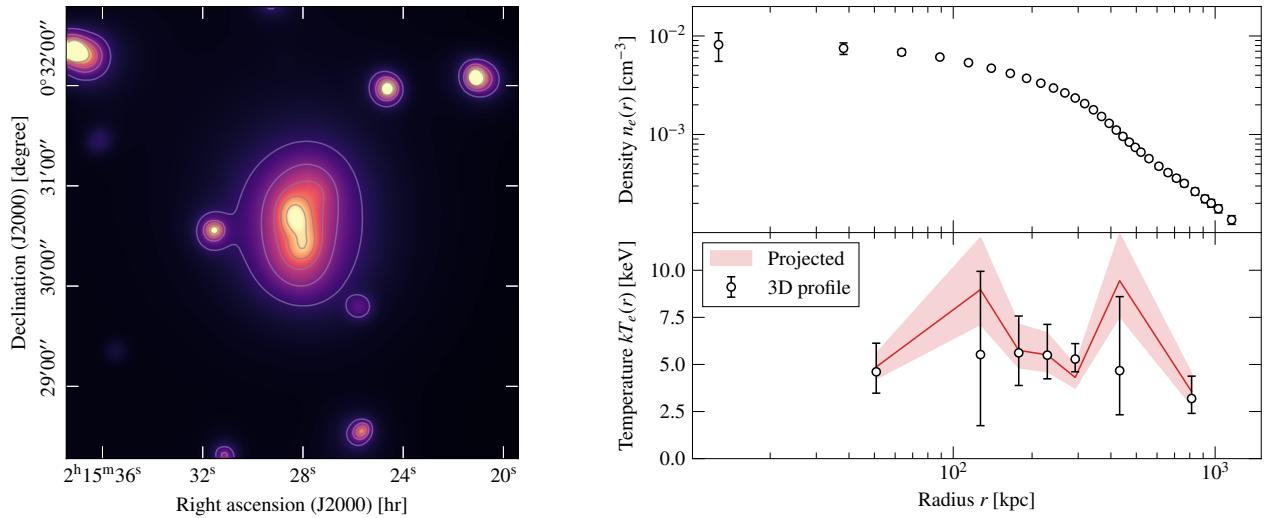


Figure 6.4 : **Gauche :** Carte d'évènements *XMM-Newton* lissée par ondelettes. La région montrée est la même que dans la figure 6.3. Les iso-contours sont montrés en gris. **Droite :** Profils de densité (*haut*) et température (*bas*) électronique extraits à partir des observations XMM de l'amas. Les profils sont extraits aux coordonnées du pic SZ, comme décrit dans le texte (6.1.3). Les barres d'erreur montrent les incertitudes à 1σ . Pour le profil de température, la ligne rouge représente le profil projeté, tandis que les points blancs représentent le profil 3D reconstruit; pour une explication détaillée de la différence entre les deux quantités, voir [111].

fixe FWHM=12.5'', d'amplitude et de position libre, dans une portion carrée de largeur 1' autour de la position de la source dans le catalogue HerS82. Les sources SMG1 et SMG2 étant proches, elles sont ajustées simultanément afin d'éviter une surestimation de leurs flux respectifs. La SED de chaque source est ensuite ajustée par MCMC sur les points SPIRE et NIKA2 à 260 GHz, en considérant le modèle de corps noir modifié donné par l'équation (4.19) et les *a priori* sur les paramètres présentés en 4.5. Les échantillons dans l'espace des paramètres sont ensuite utilisés pour calculer une SED qui est extrapolée dans la bande passante de NIKA2 à 150 GHz, ce qui donne une estimation de la densité de probabilité de l'amplitude de la contamination.

La SED ajustée et la distribution de probabilité du flux à 150 GHz de chaque source sont présentés en figure 6.6. On voit que l'ajustement de SED est correct pour toutes les sources, avec des modèles passant bien par les points. On remarque également que la distribution de probabilité du flux de chaque source à 150 GHz n'est pas gaussienne, et apparaît asymétrique. La capacité de PSTools à fournir ces distributions comme produit de sortie permettra de les utiliser comme *a priori* sur les flux des sources dans notre analyse, plutôt que de supposer une distribution gaussienne avec une valeur moyenne et une dispersion.

6.2.3 Résultats

La table 6.1 présente les flux et positions des six sources considérées dans cette analyse. La dernière colonne présente le flux attendu dans la carte NIKA2 à 150 GHz de chaque source, obtenu par ajustement puis extrapolation de SED. Il est important de noter que ces flux peuvent atteindre jusqu'à 0.74 mJy, alors que le pic d'émission SZ de l'amas est inférieur à 0.8 mJy/beam. Cela accentue l'importance de la prise en compte de la contamination par les sources ponctuelles dans cette analyse.

L'une des sources (SMG2) présente également une émission radio. Celle-ci est détectée dans le dernier catalogue FIRST [180] avec une densité de flux à 1.4 GHz de 2.22 ± 0.10 mJy. L'absence

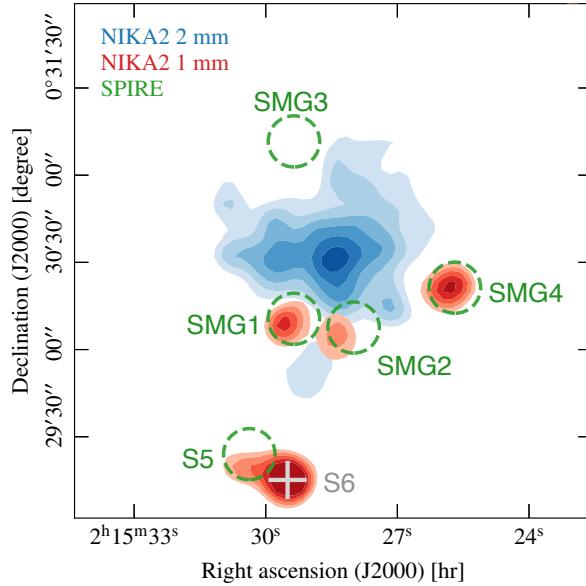


Figure 6.5 : Carte composite multifréquence de la région de l'amas. Les niveaux de bleu correspondent aux niveaux de signal sur bruit du décrément SZ dans la carte NIKA2 à 150 GHz, commençant à 3σ et espacés de 1σ . De même, les niveaux de rouge correspondent au signal sur bruit des sources ponctuelles dans la carte NIKA2 à 260 GHz. Les cercles verts sont centrés sur les positions des sources dans le catalogue HerS82 [179] et leur diamètre est la FWHM du lobe instrumental de SPIRE à 250 μ m. La croix grise marque la position de S6, la source identifiée par NIKA2 qui n'a pas de contrepartie dans le catalogue HerS82.

de couverture de la région de l'amas à d'autres fréquences radio rend impossible l'ajustement de la SED de l'émission radio de cette source. Elle est donc modélisée par une loi de puissance $F(\nu) = F(\nu_0) \times (\nu/\nu_0)^\alpha$ avec un indice spectral $\alpha = -0.7 \pm 0.2$, qui est une bonne approximation de la SED de galaxies émettrices en radio [181]. L'extrapolation de cette loi de puissance dans la bande passante à 150 GHz de NIKA2 donne une estimation de la densité de flux radio de cette source, $F_{\text{radio}}(150 \text{ GHz}) = 83^{+127}_{-50} \mu\text{Jy}$. Cette contribution est ajoutée à celle calculée par extrapolation de SED dans le domaine sub-millimétrique pour obtenir une estimation de la densité de flux totale de la source, bien que la contribution radio soit négligeable devant l'émission thermique de la poussière.

Les distributions de probabilité des flux à 150 GHz obtenues seront utilisées dans cette analyse afin de considérer la contamination par les sources ponctuelles comme des paramètres de nuisance lors de l'ajustement du profil de pression de ACT-CL J0215.4+0030. Cette approche est préférée à une simple soustraction des sources dans le but de propager l'incertitude sur les flux des sources aux mesures des propriétés du milieu intra-amas, en traitant la contamination comme un effet systématique affectant l'analyse. Cette procédure est décrite dans la section suivante.

6.3 Mesures des propriétés thermodynamiques du milieu intra-amas

La mesure des propriétés thermodynamiques du milieu intra-amas de ACT-CL J0215.4+0030 a été réalisée avant le développement de PANCO2, et a donc utilisé le logiciel PANCO [162]. L'algorithme en est décrit dans le chapitre 5 (de même que les différences entre PANCO et PANCO2), aussi ne décrirons nous que brièvement la procédure employée.

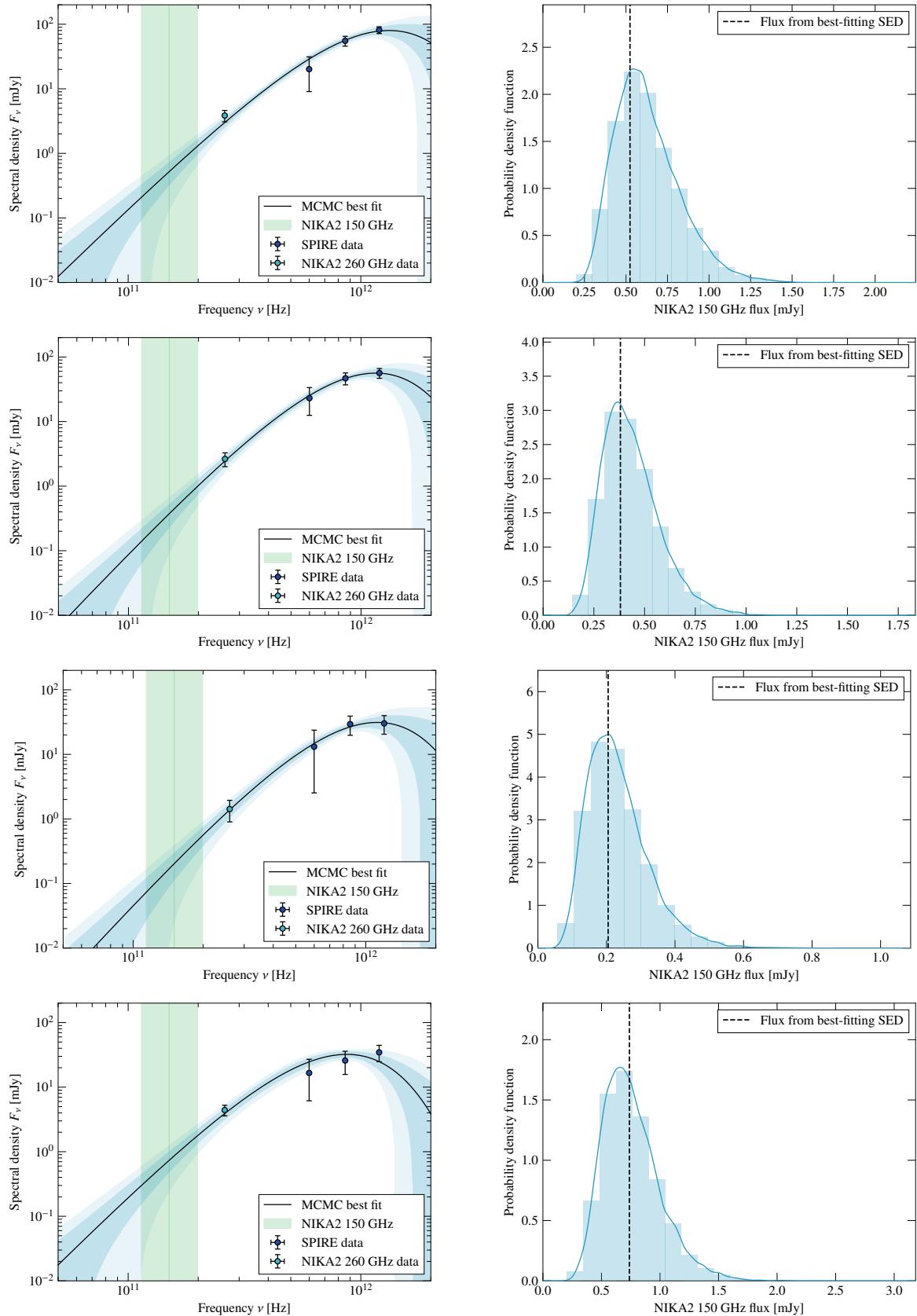


Figure 6.6 : De haut en bas, résultats de PStools pour les sources SMG1 à SMG4 dans le champ de l'amas ACT-CL J0215.4+0030. **Gauche** : ajustement de la SED sur les flux SPIRE (bleu foncé) et NIKA2 à 260 GHz (bleu clair). La région verte correspond à la bande passante de NIKA2 à 150 GHz. Les régions bleues correspondent aux intervalles de confiance à 1σ et 2σ sur l'ajustement de la SED. **Droite** : distribution de probabilité du flux à 150 GHz. L'histogramme représente les flux obtenus pour les chaînes de Markov, et la courbe bleue est la distribution de probabilité résultante.

Source	Coordonnées J2000	1200 GHz [mJy]	860 GHz [mJy]	600 GHz [mJy]	260 GHz [mJy]	150 GHz [mJy]
SMG1	02 ^h 15 ^m 29.5 ^s +00°30'07.2"	81.55 ± 9.77	55.39 ± 9.63	20.14 ± 11.08	3.85 ± 0.77	0.62 ± 0.19
SMG2	02 ^h 15 ^m 28.3 ^s +00°30'03.5"	56.57 ± 9.83	46.87 ± 9.83	23.07 ± 10.57	2.64 ± 0.64	0.38 ± 0.14 ^(a)
SMG3	02 ^h 15 ^m 29.1 ^s +00°31'10.2"	30.19 ± 9.68	29.45 ± 9.69	13.13 ± 10.6	1.42 ± 0.52	0.21 ± 0.09
SMG4	02 ^h 15 ^m 25.7 ^s +00°30'21.3"	34.6 ± 9.86	25.84 ± 10.11	16.58 ± 10.43	4.42 ± 0.83	0.74 ± 0.25
S5	02 ^h 15 ^m 30.0 ^s +00°29'12.6"	—	—	—	—	0.66 ± 0.19 ^(b)
S6	02 ^h 15 ^m 29.0 ^s +00°29'14.2"	—	—	—	—	0.61 ± 0.19 ^(b)

Table 6.1 : Catalogue des positions et flux des six sources sub-millimétriques détectées dans un rayon de 2.5' autour du centre de ACT-CL J0215.4+0030. Les flux à 1200, 860 et 600 GHz proviennent du catalogue HerS82 [179]. Les coordonnées et flux à 260 GHz sont obtenues par ajustement de modèles de lobe dans la carte NIKA2 à 260 GHz. Les flux à 150 GHz et leurs barres d'erreur sont les moyennes et écarts type des distributions de flux dans la bande passante de NIKA2. ^(a) Flux incluant l'émission radio. ^(b) Flux mesurés directement dans la carte NIKA2 à 150 GHz.

6.3.1 Modélisation du milieu intra-amas

Comme décrit dans les chapitres 2 et 5, l'amplitude de l'effet SZ thermique, nommée paramètre de Compton y , est proportionnelle à la pression du milieu intra-amas intégrée le long de la ligne de visée (équation 2.11 page 33). Nous modélisons le milieu intra-amas comme une structure de gaz à symétrie sphérique, dont la distribution radiale de pression est décrite par un modèle de profil Navarro-Frenk-White généralisé (gNFW, [164]) :

$$P_e(r) = P_0 \left(\frac{r}{r_p} \right)^{-c} \left[1 + \left(\frac{r}{r_p} \right)^a \right]^{(c-b)/a}, \quad (6.1)$$

où P_0 est la normalisation du profil de pression, b et c sont les pentes du profil dans les régions externes et internes de l'amas respectivement, et r_p et a quantifient respectivement le rayon et la vitesse de la transition entre les deux régimes.

6.3.2 Analyse MCMC

PANCO utilise une procédure Monte Carlo à Chaînes de Markov (MCMC) pour ajuster le profil de pression du milieu intra-amas sur la carte NIKA2 à 150 GHz. À chaque pas du MCMC, un profil de pression gNFW est intégré le long de la ligne de visée pour obtenir un profil de paramètre de Compton y . Ce profil est ensuite projeté sur une grille pixelisée de la même façon que la carte NIKA2 pour obtenir une carte de paramètre de Compton. Cette carte est convoluée par le lobe de NIKA2, modélisé par une Gaussienne à deux dimensions de largeur à mi-hauteur FWHM = 18.5'', en accord avec l'étalonnage des cartes NIKA2 [11]. La carte est centrée aux coordonnées du pic de brillance de surface SZ dans la carte NIKA2 à 150 GHz. La carte est ensuite convertie en unités de brillance de surface à l'aide d'un coefficient de conversion C_{conv} . Ce coefficient dépend de la forme de la distorsion spectrale de l'effet SZ, mais également des bandes passantes de NIKA2, de son lobe instrumental, et de l'opacité atmosphérique, qui varie au cours des observations. Par conséquent, le coefficient est traité comme un paramètre de nuisance et ajusté avec les paramètres du profil de pression, avec un *a priori* Gaussien d'écart type 10% dans le but de tenir compte de l'incertitude sur l'étalonnage absolue des cartes NIKA2.

La contamination par les sources ponctuelles est ajustée simultanément avec le profil de pression

du milieu intra-amas, plutôt que simplement soustraite en utilisant les flux des sources obtenus par extrapolation de SED. Chaque source est modélisée par un modèle de lobe de NIKA2 d'amplitude libre ajouté au modèle de carte de l'effet SZ. Le flux de chaque source est traité comme un paramètre de nuisance, avec un *a priori* donné par la distribution de probabilité calculée dans la section 6.2. Cette procédure présente l'avantage de tenir compte de l'incertitude sur les densités de flux extrapolées dans l'ajustement du profil de pression. Les sources S5 et S6 sont quant à elles jugées assez éloignées de l'amas pour être directement soustraites des données. Le modèle de carte de brillance de surface obtenu, modélisant simultanément l'effet SZ et la contamination par les sources ponctuelles, est convolué par la fonction de transfert présentée dans la figure 6.2 afin de tenir compte du filtrage créé par le pipeline NIKA2.

La couverture angulaire de NIKA2 et le filtrage des grandes échelles par la soustraction du bruit corrélé limite la possibilité de contraindre le profil de pression dans la périphérie de l'amas. De telles contraintes peuvent toutefois être établies en utilisant la valeur du signal SZ intégré dans un rayon de R_{500} :

$$\mathcal{D}_A^2 Y_{500} = 4\pi \frac{\sigma_r}{m_e c^2} \int_0^{R_{500}} r^2 P_e(r) dr, \quad (6.2)$$

où \mathcal{D}_A est la distance diamètre angulaire jusqu'à l'amas. Le paramètre de Compton intégré de l'amas a été mesuré dans le relevé ACT, avec une valeur publiée de [136]

$$\mathcal{D}_A^2 Y_{500}^{\text{ACT}} = (4.07 \pm 1.13) \times 10^{-5} \text{ Mpc}^2 \quad (6.3)$$

Cette contrainte est incluse dans la fonction de vraisemblance de l'ajustement du profil de pression : à chaque pas du MCMC, la valeur de Y_{500} est calculée et comparée à cette mesure, imposant une contrainte effective sur le signal SZ à grande échelle.

En résumé, à chaque pas du MCMC, un modèle de carte est calculé à partir d'un jeu de paramètres ϑ comprenant :

- P_0, r_p, a, b , et c : les paramètres décrivant le profil de pression gNFW du milieu intra-amas ;
- $F_{150}^i, i = 1 \dots 4$: les densités de flux des sources ponctuelles proches de l'amas ;
- C_{conv}, Z : le coefficient de conversion de y vers Jy/beam et le niveau zéro de la carte.

La probabilité qu'un jeu de paramètres décrive nos données est obtenue en combinant la distribution *a priori* sur les paramètres et la fonction de vraisemblance \mathcal{L} comparant le modèle de carte \mathcal{M} aux données \mathcal{D} :

$$\log \mathcal{L}(\vartheta) = -\frac{1}{2} (\mathcal{M}(\vartheta) - \mathcal{D})^T \Sigma^{-1} (\mathcal{M}(\vartheta) - \mathcal{D}) - \frac{1}{2} \left(\frac{Y_{500}^{\text{ACT}} - Y_{500}(\vartheta)}{\Delta Y_{500}^{\text{ACT}}} \right)^2, \quad (6.4)$$

où Σ est la matrice de covariance du bruit, évaluée à partir de 10^5 réalisations aléatoire de bruit corrélé générées à partir du spectre de puissance présenté sur la figure 6.2. Cette distribution de probabilité est échantillonnée à l'aide de la librairie Python emcee [157]. La convergence des chaînes est contrôlée à intervalle régulier à l'aide du test de Gelman-Rubin [171]. Une fois la convergence atteinte, une longueur de *burn-in* est enlevée au début de chaque chaîne.

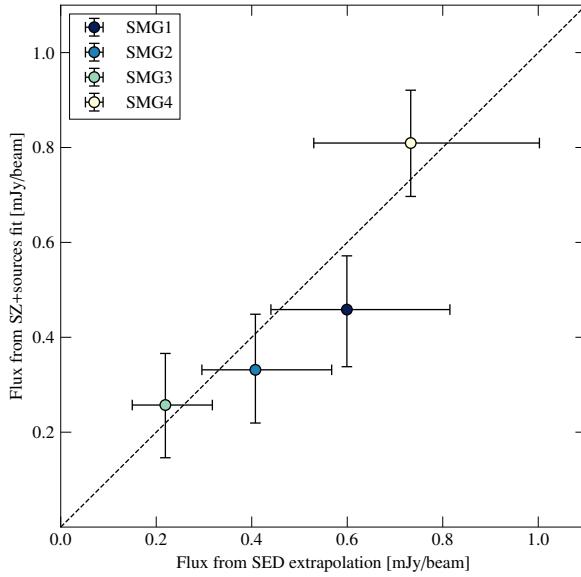


Figure 6.7 : Comparaison des densités de flux des sources ponctuelles estimées par extrapolation de SED (abscisse) et par l’analyse jointe SZ–sources ponctuelles avec un *a priori* uniforme sur les flux (ordonnée, voir section 6.3.3). Les barres d’erreur représentent les 16èmes et 84èmes percentiles des distributions. Les deux estimateurs sont compatibles dans ces barres d’erreur pour chacune des sources.

6.3.3 Auto-consistance des estimations de flux des sources ponctuelles

Les paramètres évalués lors de la reconstruction du milieu intra-amas incluent les flux des sources ponctuelles proches de l’amas. La comparaison des flux retrouvés par l’analyse MCMC avec ceux obtenus par extrapolation de SED révèle un accord des mesures. Cependant, puisque les résultats utilisant l’extrapolation de SED sont utilisés comme *a priori* dans l’analyse jointe du milieu intra-amas et des sources, les résultats ne sont pas indépendants. À fin de comparaison, l’analyse MCMC décrite dans la section précédente a donc été répétée sans utiliser d’*a priori* sur les flux des sources. Les résultats de cette comparaison sont représentés sur la figure 6.7. Pour chaque source, les mesures de flux sont compatibles à 1σ . Ce résultat renforce la confiance dans l’estimation des flux des sources, et donc dans l’estimation du signal SZ dans les régions contaminées. De plus, bien que cette compatibilité montre que l’extrapolation de SED n’apporte que peu d’information utile au MCMC, l’utilisation d’*a priori* sur les flux des sources réduit grandement le volume de l’espace des paramètres accessible par le MCMC, réduisant ainsi la durée nécessaire à l’atteinte de la convergence. Enfin, il est à noter que l’ajustement sans *a priori* sur le flux des sources donne une importance plus grande à l’hypothèse de sphéricité du milieu intra-amas, et son utilisation systématique pourrait avoir des conséquences sur la capacité à identifier des amas perturbés à partir des données NIKA2.

6.4 Résultats

L’échantillonnage par MCMC décrit dans la section précédente est utilisé pour déterminer le jeu de paramètres ϑ qui ajuste le mieux la carte NIKA2 à 150 GHz. Les modèles de carte de l’effet SZ, de contamination par les sources ponctuelles, et de la somme des deux contributions sont représentés dans la figure 6.8. Puisqu’aucune structure de rapport signal sur bruit significative n’est détectée dans les résidus, il est possible d’affirmer qu’il n’y a pas de signe de sous-structures ou d’asphéricité dans la carte NIKA2 de cet amas.

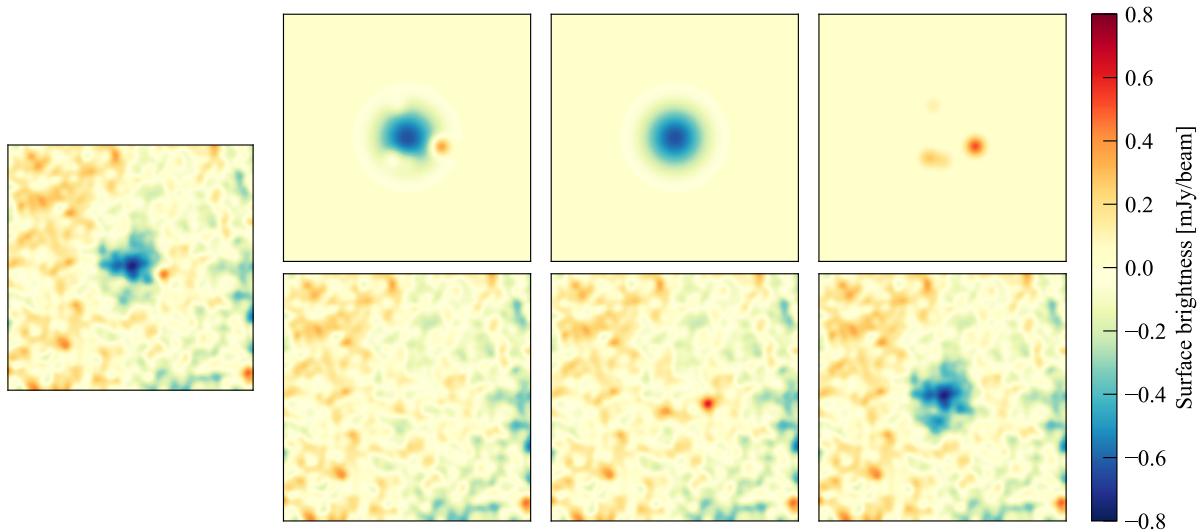


Figure 6.8 : Résultats de l’ajustement de la carte NIKA2 à 150 GHz de ACT-CL J0215.4+0030 (panneau de gauche). Le modèle ajustant le mieux les données et les résidus sont respectivement montrés dans les panneaux haut et bas de la colonne du centre gauche. Les colonnes de centre droit et droite représentent le meilleur modèle et les résidus en considérant seulement le signal SZ et les sources ponctuelles, respectivement. Les cartes sont représentées dans la même région et lissées par le même filtre que dans la figure 6.3. Les deux sources dans le sud de la carte (S5 et S6) ont été soustraite.

6.4.1 Thermodynamique du milieu intra-amas

Le profil de pression obtenu à l’issue de l’analyse MCMC est présenté en haut à gauche de la figure 6.9. La ligne bleue correspond au maximum de vraisemblance de l’ajustement. Les enveloppes d’erreur sont obtenues en calculant un profil de pression pour chaque jeu de paramètres échantillonné par le MCMC, puis la dispersion de ces profils en fonction de la distance au centre. Ce profil est comparé à celui obtenu par combinaison des profils de densité et de température obtenus en X (représenté en blanc dans la figure 6.9). Cette comparaison est intéressante car les deux profils représentent deux mesures indépendantes de la distribution de pression dans le milieu intra-amas d’un amas distant et peu massif. Les deux mesures sont en accord ; le profil de pression obtenu en SZ semble marginalement plus élevé, mais cet effet n’est pas significatif au vu des incertitudes sur les deux profils.

Plusieurs autres propriétés thermodynamiques peuvent être obtenues en combinant le profil de pression obtenu en SZ avec le profil de densité obtenu en X : le profil de température $T_e(r)$, d’entropie $K_e(r)$, et de masse hydrostatique $M_{\text{HSE}}(< r)$ sont obtenus grâce à :

$$k_B T_e(r) = P_e(r) / n_e(r), \quad K_e(r) = P_e(r) n_e^{-5/3}(r), \quad (6.5)$$

et à l’équation de l’équilibre hydrostatique :

$$M_{\text{HSE}}(< r) = -\frac{1}{\mu m_p G} \frac{r^2}{n_e(r)} \frac{dP_e}{dr}, \quad (6.6)$$

où m_p est la masse du proton, μ le poids moléculaire moyen du gaz, et G la constante de gravitation. Ces profils sont présentés sur la figure 6.9, en bleu tels qu’obtenus par combinaison de la densité X et de la pression SZ, et en blanc par combinaison de la densité et de la température X. Comme pour la pression, les profils sont en accord, illustrant la complémentarité entre observations SZ à haute résolution et X.

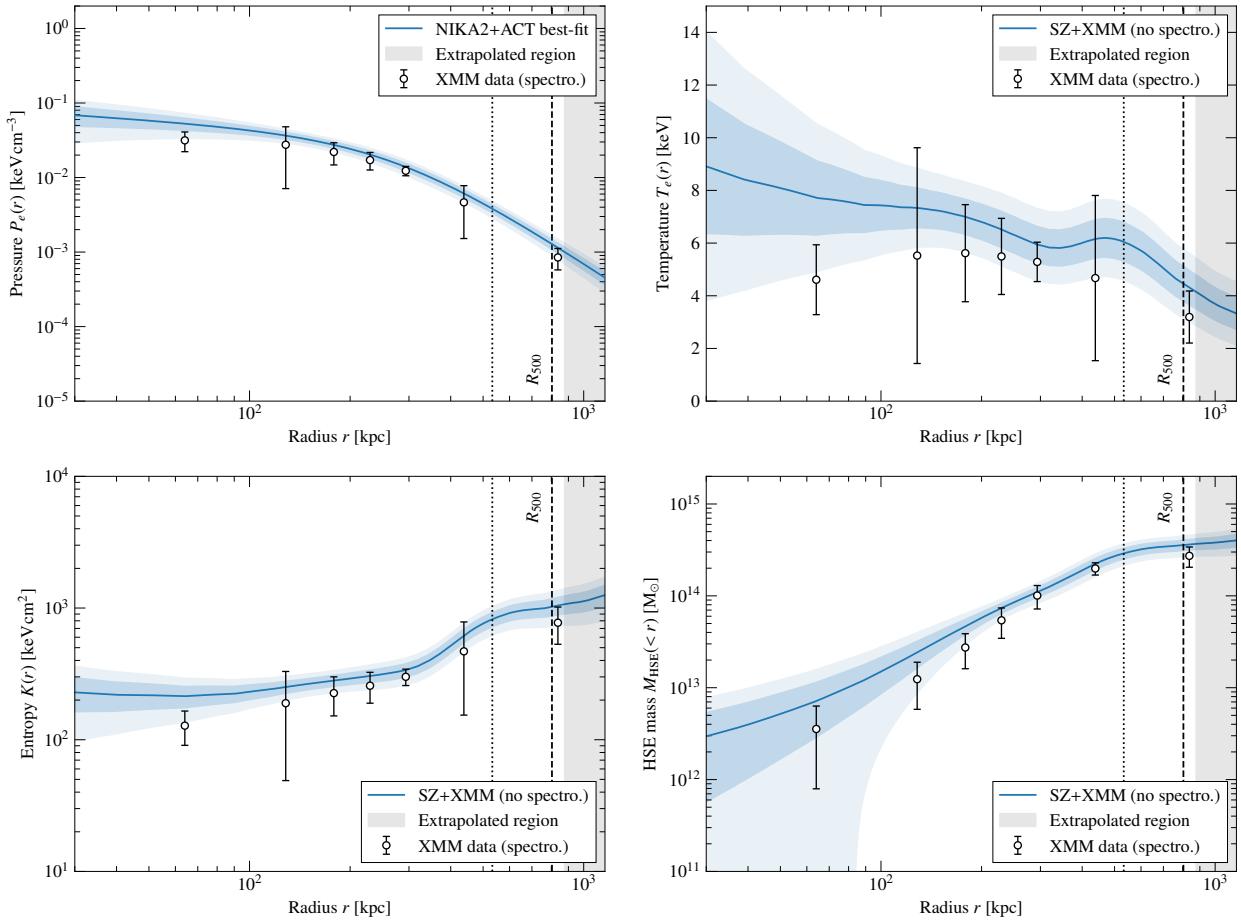


Figure 6.9 : Profils radiaux de la pression (haut gauche), température (haut droit), entropie (cas gauche), et masse hydrostatique (bas droit) du milieu intra-amas de ACT-CL J0215.4+0030. Pour le profil de pression, la courbe bleue représente le profil de pression correspondant au meilleur ajustement des données NIKA2. Pour les autres profils, elle montre la combinaison entre ce profil de pression et le profil de densité obtenu avec *XMM-Newton*. Les enveloppes bleues montrent les intervalles de confiance à 1σ et 2σ sur ces profils. Les points blancs représentent les profils obtenus en combinant les profils de densité et de température mesurés en X, avec des erreurs à 1σ . La ligne noire en pointillés montre la limite à l'intérieur de laquelle le rapport signal sur bruit sur le profil de brillance de surface de la carte NIKA2 à 150 GHz est supérieur à 3σ . La zone grisée représente la région dans laquelle les profils ne sont pas contraints par les données SZ, c'est-à-dire au delà de R_{500}^{ACT} .

Bien que cet amas ait bénéficié de temps d'observation similaires avec NIKA2 et *XMM-Newton*, cette compatibilité illustre la grande qualité des mesures pouvant être obtenue en combinant les deux longueurs d'onde, minimisant la dépendance en observations de spectroscopie X pour la caractérisation des propriétés thermodynamiques du milieu intra-amas. Enfin, cette combinaison permet d'atteindre une plus grande résolution sur le profil de température du milieu intra-amas que ce qui peut être atteint en spectroscopie X. La figure 6.9 montre que l'extrapolation de la pente interne du profil de pression obtenu en spectroscopie X donne une température de cœur bien plus faible que celle obtenue par combinaison X-SZ. Les incertitudes sur le profil de température sont également plus faibles en moyenne pour le profil NIKA2+XMM. Par conséquent, les profils de masse et d'entropie présentent également des incertitudes plus faibles en combinant les observations X et SZ qu'en utilisant la spectroscopie X, résultant en une mesure plus précise des propriétés thermodynamiques du milieu intra-amas.

Les propriétés thermodynamiques de ACT-CL J0215.4+0030 fournissent des renseignements sur l'état dynamique de son milieu intra-amas. Le profil de pression est très plat dans la partie centrale

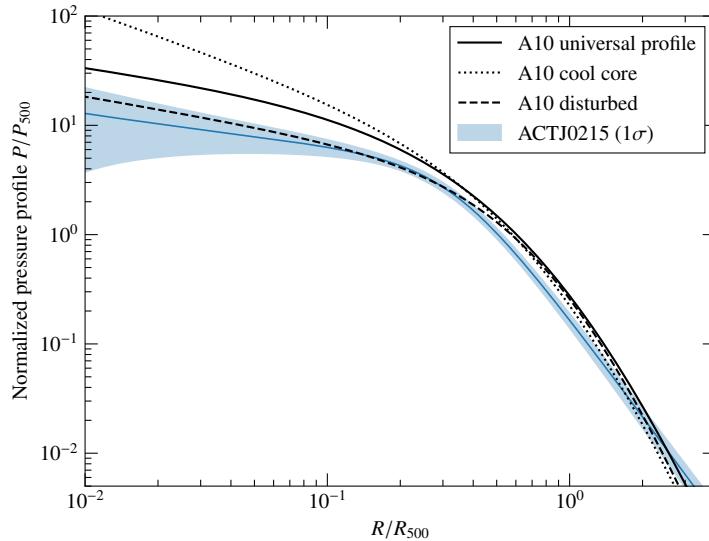


Figure 6.10 : Comparaison du profil de pression obtenu à partir des données NIKA2 (bleu) avec les profils moyens de Arnaud *et al.* [94] (noir). ACT-CL J0215.4+0030 présente un profil similaire aux amas identifiés comme ayant un cœur perturbé (ligne noire pointillée).

de l’amas, indiquant un cœur perturbé. La figure 6.9 compare le profil de pression obtenu dans cette analyse avec les profils de pression moyens obtenus en X et à bas redshift pour les amas de l’échantillon REXCESS [94]. Le profil de pression de ACT-CL J0215.4+0030 est mal décrit par le profil de pression dit universel, particulièrement dans le cœur de l’amas. En revanche, il est bien décrit par le profil moyen des amas de l’échantillon REXCESS possédant un cœur perturbé. La valeur de l’entropie ($K_{e,\text{core}} \gtrsim 200 \text{ keV} \cdot \text{cm}^2$) et de la densité électronique ($n_{e,\text{core}} \lesssim 10^{-2} \text{ cm}^{-3}$) indiquent également un amas au cœur perturbé [145]. La figure 6.9 compare le profil de pression obtenu dans cette analyse avec les profils de pression moyens obtenus en X et à bas redshift pour les amas de l’échantillon REXCESS [94]. Le profil de pression de ACT-CL J0215.4+0030 est mal décrit par le profil de pression universel, particulièrement dans le cœur de l’amas. En revanche, il est bien décrit par le profil moyen des amas perturbés de l’échantillon REXCESS. Des conclusions similaires ont également été tirées par Ricci *et al.* [182], en utilisant NIKA2 et *XMM-Newton* pour caractériser le milieu intra-amas d’un autre amas distant de faible masse. Ces conclusions sont intéressantes dans le cadre de discussions sur l’évolution de la fraction d’amas à cœur perturbé avec le redshift (par exemple, [183, 184]).

6.4.2 Grandeurs intégrées

Pour des relevés SZ tels que ACT ou *Planck*, la résolution angulaire ne permet pas de résoudre les amas distants, et la distribution radiale de leurs propriétés thermodynamiques ne peut donc pas être mesurée. Par conséquent, les relations d’échelles utilisées pour calculer les masses des amas relient leur signal SZ intégré dans un rayon donné à la masse de l’amas contenue dans ce rayon. Par exemple, le relevé ACT [136] utilise le rayon R_{500} , correspondant au rayon à l’intérieur duquel la densité moyenne est 500 fois supérieure à $\rho_c(z)$, la densité critique de l’Univers au redshift de l’amas. Ce rayon peut être calculé comme le rayon pour lequel le contraste de densité δ_c vaut 500, c’est-à-dire en résolvant

$$\delta_c(R_{500}) = \frac{M_{\text{HSE}}(< R_{500})}{\rho_c(z) \times \frac{4}{3}\pi R_{500}^3} = 500. \quad (6.7)$$

	ACT	XMM-Newton	NIKA2+XMM
R_{500} [kpc]	877.8 ± 46.2	780.9 ± 19.8	810.1 ± 41.9
$\mathcal{D}_A^2 Y_{500}$ [10^{-5} Mpc^2]	4.07 ± 1.13	–	3.76 ± 0.39
M_{500} [10^{14} M_\odot]	3.5 ± 0.8	2.48 ± 0.70	3.79 ± 0.58

Table 6.2 : Grandeur intégrées caractéristiques de ACT-CL J0215.4+0030 obtenue avec l’analyse des données ACT, XMM-Newton, et NIKA2+XMM-Newton. Les valeurs correspondant à l’analyse ACT proviennent de la table 8 de [136].

En utilisant le profil de masse hydrostatique obtenu par combinaison des profils X et SZ et présenté en figure 6.9, on trouve

$$R_{500} = 810.1 \pm 41.9 \text{ kpc}, \quad (6.8)$$

en accord avec la mesure du relevé ACT [136]. Le paramètre de Compton intégré dans ce rayon peu alors être calculé :

$$\mathcal{D}_A^2 Y_{500} = (3.76 \pm 0.39) \times 10^{-5} \text{ Mpc}^2, \quad (6.9)$$

de même que la masse hydrostatique contenue dans le même rayon :

$$M_{500}^{\text{HSE}} = M_{\text{HSE}}(< R_{500}) = (3.79 \pm 0.58) \times 10^{14} \text{ M}_\odot. \quad (6.10)$$

Ces résultats sont présentés dans la table 6.2. À des fins de comparaison, deux autres mesures de R_{500} et M_{500} y sont également rapportés, respectivement obtenus à partir de l’algorithme de détection d’amas dans le relevé ACT sous l’hypothèse d’un profil de pression universel, et à partir de l’analyse des seules données XMM-Newton de ACT-CL J0215.4+0030. L’incertitude sur la mesure de Y_{500} est réduite d’un facteur 3 par rapport à la mesure du relevé ACT, indiquant un gain significatif en précision grâce aux observations SZ à haute résolution. La valeur de M_{500} obtenue à partir du profil de masse hydrostatique est compatible avec celle publiée dans le catalogue ACT, calculée avec la relation d’échelle $Y_{500} - M_{500}$ de [94]. L’incertitude sur la masse est également plus faible dans cette étude que celle estimée dans le catalogue ACT, alors même que cette dernière est sous-estimée puisqu’elle ne tient pas compte de l’incertitude sur la relation d’échelle utilisée pour l’estimation (voir table 8 de [136]). La valeur de masse obtenue par combinaison X-SZ est également compatible avec celle obtenue à partir de l’analyse X seule, bien que cette dernière soit plus faible en conséquence d’un rayon R_{500} plus faible, et plus précise.

6.5 Conclusions

Le travail présenté dans ce chapitre constitue la première analyse individuelle d’un amas de galaxies du grand programme SZ de NIKA2 en utilisant des données de qualité standard. Cette analyse a soulevé plusieurs difficultés. Tout d’abord, l’amas analysé, ACT-CL J0215.4+0030, est un amas de basse masse et haut redshift dans l’échantillon du LPSZ, et de ce fait une source faible et compacte. De plus, l’amas a été observé dans des conditions estimées suffisantes pour atteindre un rapport signal sur bruit de 3σ à R_{500} sur le profil de brillance de surface SZ, objectif qui n’est pas atteint puisque le signal sur bruit ne vaut 3σ que jusqu’à $0.7 \times R_{500}$. Enfin, le signal SZ détecté dans NIKA2 est fortement affecté par la

présence de sources ponctuelles. Cette contamination est particulièrement importante dans le cas de cet amas à cause du flux élevé des sources par rapport à celui de l'amas, mais également de sa faible extension spatiale.

En dépit de ces difficultés, de fortes contraintes ont pu être posées sur les propriétés thermodynamiques du milieu intra-amas de ACT-CL J0215.4+0030. La contamination par les sources ponctuelles a été prise en compte dans une approche multifréquence basée sur une analyse jointe du signal SZ et des sources, permettant une propagation de l'incertitude sur l'amplitude de cette contamination aux résultats finaux. Cette analyse a permis de mesurer la distribution radiale de pression du milieu intra-amas de l'amas, et de combiner ce résultat avec des mesures X pour caractériser plus largement la thermodynamique du milieu intra-amas de ACT-CL J0215.4+0030.

Les conclusions tirées de ces résultats sont les suivantes :

- Le profil de pression de ACT-CL J0215.4+0030 est incompatible avec le profil de pression dit universel des amas de galaxies publié par Arnaud *et al.* [94], mais en accord avec le profil des amas perturbés rapporté dans la même publication. Cela indique un cœur d'amas perturbé, en accord avec les autres propriétés thermodynamiques obtenues par combinaison avec des mesures X (telles que l'entropie au cœur de l'amas).
- Les propriétés thermodynamiques obtenues en combinant les mesures NIKA2 et *XMM-Newton* sans spectroscopie sont compétitives avec celles obtenues en utilisant des mesures X seules. Puisque les temps d'observation nécessaires à l'extraction d'information spectroscopique en X sont largement supérieurs à ceux nécessaires à l'obtention d'un profil de densité, et puisque ces temps augmentent rapidement avec le redshift, la combinaison d'observations SZ à haute résolution et X peu profondes apparaît comme un outil idéal pour la caractérisation du milieu intra-amas d'amas distants.
- Les observations NIKA2 de cet amas ont permis d'améliorer la précision des mesures des profils thermodynamiques, de la masse et du signal SZ intégré de ACT-CL J0215.4+0030, à la fois par rapport au relevé ACT et à l'analyse X seule.

Cette analyse étant focalisée sur l'un des amas les plus complexes du LPSZ, il est raisonnable d'attendre de telles conclusions pour la plupart des analyses individuelles de l'échantillon. Ces résultats sont donc très prometteurs pour l'étude du profil de pression moyen et de la relation d'échelle $Y_{500} - M_{500}$ grâce au LPSZ.

CHAPITRE 7

La relation d'échelle SZ – Masse

7.1	Modélisation et état de l'art	142
7.1.1	Paramètre de Compton intégré et masse	142
7.1.2	État de l'art des mesures	144
7.2	Modélisation probabiliste de la relation d'échelle	146
7.2.1	Relation linéaire masse-observable	146
7.2.2	Effets de sélection et représentativité	149
7.3	Ajustement de la relation d'échelle par MCMC	152
7.3.1	Distribution de probabilité <i>a priori</i> des paramètres	152
7.3.2	Analyse MCMC par échantillonnage de Gibbs	153
7.4	Validation sur données simulées	155
7.4.1	Incertitudes de mesure et dispersion intrinsèque	155
7.4.2	Estimateur de masse dispersé	157
7.4.3	Estimateur de masse biaisé	157
7.4.4	Données tronquées	158
7.5	Application au grand programme SZ de NIKA2	158
7.5.1	Simulation d'échantillons	158
7.5.2	Modélisation et impact de la sélection	160
7.5.3	Incertitudes sur les paramètres d'intérêt	163
7.5.4	Extensions du grand programme SZ	164
7.6	Conclusions et perspectives	166

L'un des objectifs principaux du grand programme SZ de NIKA2 est la mesure de la relation d'échelle liant l'observable SZ des relevés cosmologiques à la masse des amas de galaxies. Ainsi, tout au long de cette thèse, nous avons fait référence à ce chapitre en soulignant l'importance de plusieurs éléments dans le calcul de la relation d'échelle. Nous avons vu au chapitre 2 que l'exploitation cosmologique d'échantillons d'amas dépendait d'une connaissance précise de la relation d'échelle,

et l'utilité du grand programme SZ dans ce contexte au chapitre 3. De même, nous avons vu au chapitre 5 que les produits du logiciel PANCO2 étaient les données d'entrée utilisées pour l'ajustement de cette relation. Ainsi, l'évaluation de la relation d'échelle représente l'aboutissement de ce travail de thèse, faisant appel aux différents travaux menés en amont pour évaluer cette relation de manière aussi juste et précise que possible, c'est-à-dire en minimisant le biais et la dispersion des estimateurs des paramètres de la relation d'échelle obtenus par régression.

Ce chapitre traite de la méthodologie de l'ajustement de la relation d'échelle dans le cadre du grand programme SZ de NIKA2. Nous présentons tout d'abord le lien entre masse et paramètre de Compton intégré, ainsi que quelques unes des mesures récentes de cette relation. Nous détaillons ensuite l'utilisation d'un modèle bayésien hiérarchique de la relation d'échelle pour l'ajustement, qui permet de tenir compte des différents effets systématiques propres aux relations masse-observable. Nous utiliserons le logiciel LIRA [185] pour effectuer l'échantillonnage de la distribution de probabilité des paramètres par échantillonnage de Gibbs. Nous présenterons la validation de ce code sur des données simulées comportant plusieurs caractéristiques pertinentes pour les relations d'échelle masse-observable. Nous présenterons ensuite l'analyse d'échantillons simulés réalistes, similaires au LPSZ, afin de prédire les biais et incertitudes sur les mesures de paramètres pouvant être effectuées avec cet échantillon.

7.1 Modélisation et état de l'art

7.1.1 Paramètre de Compton intégré et masse

La relation d'échelle étudiée au sein du grand programme SZ de NIKA2 est celle liant le paramètre de Compton intégré Y_{500} dans un rayon caractéristique R_{500} à la masse des amas M_{500} contenue dans ce même rayon (voir section 2.1.2, page 26). Le paramètre de Compton intégré a été introduit par Kravtsov *et al.* [186] comme une observable fortement corrélée à la masse, du fait du lien étroit entre ce dernier et l'énergie thermique au sein de l'amas détaillé en section 2.2.3 (page 32). En effet, dans l'hypothèse autosimilaire (voir par exemple [186, 187]), il est possible de relier le paramètre de Compton intégré à la masse d'un amas de façon analytique. Considérons une sphère contenant un gaz parfait à température T . L'équation de l'équilibre hydrostatique (2.3) s'écrit :

$$\frac{k_B T}{m} \frac{d\rho}{dr} = -\frac{\rho GM(< r)}{r^2}, \quad (7.1)$$

où ρ et T sont la masse volumique et la température du gaz, et m est la masse des particules le constituant.

Cette équation peut se réécrire comme :

$$\frac{r^2}{\rho} \frac{d\rho}{dr} = -4\pi \frac{Gm}{k_B T} \int_0^r r^2 \rho(r) dr. \quad (7.2)$$

On peut dériver cette équation par rapport au rayon r :

$$\frac{d}{dr} \left[r^2 \frac{d \ln \rho}{dr} \right] = -4\pi \frac{Gm}{k_B T} r^2 \rho(r). \quad (7.3)$$

Cette équation différentielle admet des solutions en loi de puissance. En y injectant une solution de la forme $\rho(r) = \alpha r^\beta$, on trouve $\beta = 2$ et $\alpha = k_B T / 2\pi Gm$, soit :

$$\rho(r) = \frac{k_B T}{2\pi Gm} \frac{1}{r^2}. \quad (7.4)$$

On peut évaluer la masse $M(< r)$ contenue dans une sphère de rayon r en intégrant cette équation :

$$\begin{aligned} M(< r) &= 4\pi \int_0^r r^2 \rho(r) dr \\ &= \frac{2k_B T}{Gm} r. \end{aligned} \quad (7.5)$$

La masse contenue dans un rayon r est donc proportionnelle à ce rayon. On parle alors d'autosimilarité des amas de galaxies : les amas sont des répliques à l'échelle les uns des autres. Cette hypothèse joue un rôle important en cosmologie avec des amas de galaxies, par exemple pour le calcul du profil de pression moyen des amas (voir section 2.2.4, page 38).

Le rayon caractéristique R_Δ est défini en section 2.1.2 (page 26) comme le rayon de la sphère contenant une densité moyenne égale à $\Delta \rho_c(z)$, où $\rho_c(z)$ est la densité critique de l'Univers au redshift de l'amas (équation 1.9), soit :

$$R_\Delta = \left[\frac{2GM_\Delta}{\Delta \times H^2(z)} \right]^{1/3}. \quad (7.6)$$

D'après l'équation (7.5), la masse M_Δ contenue dans un rayon R_Δ s'écrit :

$$M_\Delta = \frac{2k_B T}{Gm} R_\Delta \quad (7.7)$$

En combinant les équations (7.6) et (7.7), on aboutit à :

$$k_B T = m \Delta^{1/3} \left[\frac{GM_\Delta H(z)}{4} \right]^{2/3}. \quad (7.8)$$

Le paramètre de Compton Y_Δ intégré dans le rayon R_Δ est quant à lui défini par l'intégrale sphérique du profil de pression dans la sphère, comme décrit dans l'équation (2.14) :

$$Y_\Delta = 4\pi \frac{\sigma_T}{m_e c^2} \int_0^{R_\Delta} P_e(r) r^2 dr. \quad (7.9)$$

L'équation d'état des gaz parfaits et l'équation (7.8) permettent d'écrire le profil de pression des électrons du gaz comme :

$$\begin{aligned} P_e(r) &= n_e(r) k_B T_e \\ &= \mu m_p n_e(r) \Delta^{1/3} \left[\frac{GM_\Delta H(z)}{4} \right]^{2/3}. \end{aligned} \quad (7.10)$$

où μ est le poids moléculaire moyen du gaz et m_p la masse du proton, de sorte à ce que $m = \mu m_p$.

Ce profil de pression peut alors être injecté dans l'équation (7.9) :

$$Y_\Delta = 4\pi \frac{\sigma_t}{m_e c^2} \mu m_p \Delta^{1/3} \left[\frac{GM_\Delta H(z)}{4} \right]^{2/3} \int_0^{R_\Delta} n_e(r) r^2 dr. \quad (7.11)$$

On peut introduire la fraction massique de gaz dans l'amas $f \equiv M_\Delta^{\text{gaz}}/M_\Delta$, avec la masse de gaz dans un rayon R_Δ définie comme :

$$M_\Delta^{\text{gaz}} = 4\pi \mu_g m_p \int_0^{R_\Delta} n_e(r) r^2 dr, \quad (7.12)$$

où μ_g est le poids moléculaire moyen du gaz.

La relation entre le paramètre de Compton intégré Y_Δ à la masse M_Δ s'écrit donc :

$$Y_\Delta = \frac{\sigma_t}{m_e c^2} \left(\frac{G}{4} \right)^{2/3} f \frac{\mu}{\mu_g} \Delta^{1/3} H^{2/3}(z) M_\Delta^{5/3}, \quad (7.13)$$

Dans le cadre de cette étude, nous nous intéressons à une valeur de contraste $\Delta = 500$, conventionnellement utilisée pour les observations SZ et X. On a alors :

$$Y_{500} = \frac{\sigma_t}{m_e c^2} \left(\frac{G}{4} \right)^{2/3} f \frac{\mu}{\mu_g} 500^{1/3} H^{2/3}(z) M_{500}^{5/3}, \quad (7.14)$$

La relation d'échelle liant le paramètre de Compton intégré Y_{500} à la masse M_{500} prend donc la forme d'une loi de puissance. Dans le scénario autosimilaire développé ici, modélisant le milieu intra-amas comme une sphère isotherme, Y_{500} évolue comme $M_{500}^{5/3}$, et le coefficient de proportionnalité est connu. On peut réécrire l'équation (7.14) pour tenir compte de possibles déviations des amas de galaxies au comportement autosimilaire et isotherme, en l'écrivant :

$$E^{-2/3}(z) \left[\frac{Y_{500}}{10^{-4} \text{Mpc}^2} \right] = 10^\alpha \left[\frac{M_{500}}{3 \times 10^{14} M_\odot} \right]^\beta, \quad (7.15)$$

où $E(z) \equiv H(z)/H_0$ est le paramètre de Hubble réduit.

La valeur de α dépend directement des propriétés physiques des amas de galaxies, comme par exemple la fraction de gaz f et la composition chimique du milieu intra-amas μ_g . On peut utiliser des valeurs typiques de ces grandeurs issues de simulations numériques : en prenant $f = 0.175$, $\mu = 0.59$ et $\mu_g = 1.14$ (voir par exemple [94]), on trouve $\alpha = -0.53$. En revanche, la valeur de $\beta = 5/3$ est fixe dans le scénario autosimilaire. L'ajustement de cette relation sur un jeu de données permet alors de mesurer les paramètres α et β .

7.1.2 État de l'art des mesures

La relation d'échelle $Y_{500} - M_{500}$ a été étudiée à plusieurs reprises au cours des dernières années. Ces études ont débuté avec l'introduction du paramètre de Compton intégré comme variable corrélée à la masse par Kravtsov *et al.* [186], qui ont mesuré le lien entre Y_{500} et M_{500} au travers de simulations hydrodynamiques.

Des mesures plus récentes de la relation d'échelle existent, utilisant aussi bien des simulations que des données réelles. L'une de ces mesures est celle effectuée par la collaboration *Planck* [148], présentée en figure 7.1. L'échantillon d'amas utilisé est constitué de 62 amas à des redshifts $z < 0.45$.

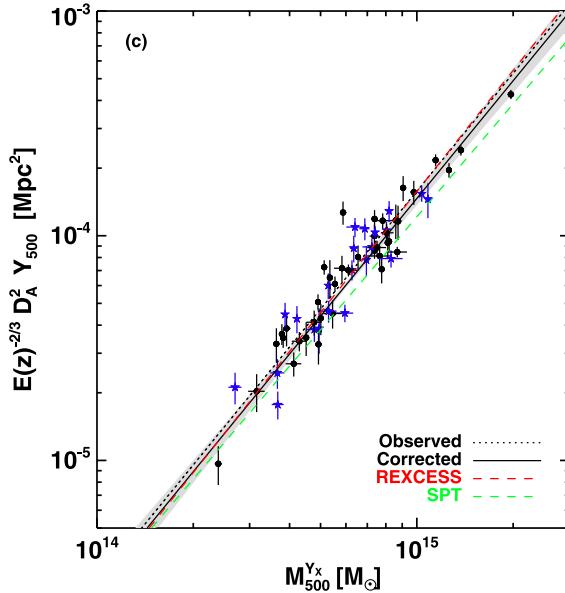


Figure 7.1 : Mesure de la relation d'échelle $Y_{500} - M_{500}$ par la collaboration *Planck*. Chaque point représente un amas de galaxies. La relation ajustant le mieux les données est présentée en noir, et comparée aux relations issues des analyses de l'échantillon REXCESS en X (rouge, [94]) et SPT en SZ (vert, [190]). Figure extraite de [148].

Le paramètre de Compton intégré est mesuré à partir des observations SZ à basse résolution des amas par *Planck*. Les masses des amas ont quant à elles été obtenues à l'aide d'observations X seules en spectroscopie, réalisées avec le satellite *XMM-Newton*. Les valeurs des paramètres offrant le meilleur ajustement de la relation d'échelle de l'équation (7.15) sur les données sont les suivantes :

$$\alpha = -0.19 \pm 0.02 ; \quad \beta = 1.79 \pm 0.08 ; \quad \sigma = 0.075 \pm 0.01, \quad (7.16)$$

où σ est la dispersion intrinsèque de la relation, qui sera discutée en 7.2.1 (page 147).

L'ajustement a été réalisé en utilisant l'algorithme BCES (*bivariate correlated errors and scatter*, [188]), qui offre un estimateur ponctuel des paramètres de la relation d'échelle, et ne permet pas de propager les incertitudes systématiques liées, par exemple, aux incertitudes sur les mesures de masse et d'observable. Par conséquent, il est à noter que les incertitudes sur les paramètres sont obtenues par *bootstrap* [189], et sont donc entièrement statistiques.

Plus récemment, dans le cadre de l'étude CoMaLit, Sereno *et al.* [191] ont utilisé les valeurs de paramètre de Compton intégré issues du catalogue d'amas de la collaboration *Planck* pour mesurer la relation d'échelle en utilisant plusieurs mesures de masses différentes. Sont notamment comparés les résultats obtenus pour des masses d'amas mesurées en X, mais également par lentillage gravitationnel. La méthodologie employée pour l'ajustement des relations est basée sur un modèle bayésien hiérarchique, similaire à celui que nous utiliserons au cours de ce chapitre. Comme nous le verrons en section 7.2, cette approche permet de tenir compte des différents effets systématiques affectant l'ajustement de la relation d'échelle. Les résultats obtenus par Sereno *et al.* [191] sont similaires à ceux de la collaboration *Planck* lors de l'analyse du même échantillon, mais avec des incertitudes augmentées d'environ un ordre de grandeur, correspondant à une prise en compte plus complète des incertitudes.

À la date de l'écriture de cette thèse, il n'existe pas d'étude ayant mesuré la relation d'échelle $Y_{500} - M_{500}$ en utilisant la combinaison d'observations SZ à haute résolution et X, ni d'étude ayant réalisé cette mesure à des redshifts supérieurs à 0.5. Le grand programme SZ de NIKA2 représente

donc une avancée majeure pour la connaissance du lien entre cette observable SZ et la masse des amas de galaxies.

7.2 Modélisation probabiliste de la relation d'échelle

Cette section détaille la modélisation probabiliste de la relation d'échelle liant Y_{500} et M_{500} , et les différentes subtilités devant être prises en compte lors de l'ajustement de ce modèle sur des données. Elle est inspirée du travail de Sereno [185], qui offre un cadre probabiliste à l'étude des relations linéaires dans le contexte de l'astronomie. L'ajustement d'une telle relation peut être effectué de plusieurs façons ; nous nous intéressons ici au cadre général d'un modèle *errors in variable* ([192-194], chapitre 4.3 de [195]). Ce modèle est caractérisé par une approche bayésienne hiérarchique de la régression linéaire, dans laquelle le modèle est découpé en plusieurs sous-modèles, chacun caractérisé comme une distribution de probabilité. Des approches fréquentistes permettant de répondre aux besoins créés par les caractéristiques des jeux de données d'astronomie existent également, comme l'algorithme BCES (*bivariate correlated errors and scatter*, [188]). De telles approches sont décrites dans la revue d'Andreon et Hurn [105].

7.2.1 Relation linéaire masse-observable

À l'instar de la relation $Y_{500} - M_{500}$, la plupart des relations d'échelle masse-observable utilisées en cosmologie avec des amas de galaxies sont des lois de puissance évoluant avec le redshift. De telles relations peuvent prendre une forme linéaire en s'intéressant aux logarithmes de la masse et de l'observable ; on peut alors écrire

$$Y = \alpha_{YZ} + \beta_{YZ}Z + \gamma T(z) + \delta T(z)Z, \quad (7.17)$$

où $T(z) = \log_{10} E(z)$.

Dans cette équation, Y est le logarithme de l'observable, et Z celui de la masse¹. Dans la littérature statistique traitant des régressions linéaires, Z est nommée variable latente ou indépendante, tandis que Y est nommée variable de réponse. α_{YZ} et β_{YZ} sont l'ordonnée à l'origine et la pente de la relation, et γ et δ paramétrisent l'évolution avec le redshift de la normalisation et de la pente de la relation.

En procédant par association entre les équations (7.15) et (7.17), on voit qu'elles sont équivalentes pour :

$$\begin{aligned} Y &\equiv \log_{10} \left[\frac{Y_{500}}{10^{-4} \text{Mpc}^2} \right], \\ Z &\equiv \log_{10} \left[\frac{M_{500}}{3 \times 10^{14} M_\odot} \right], \\ \gamma &= 2/3, \\ \delta &= 0. \end{aligned} \quad (7.18)$$

Ainsi, l'équation (7.17) est plus générale que (7.15), puisque la possibilité de considérer des valeurs de $\gamma \neq 2/3$ et $\delta \neq 0$ permet d'étudier l'écart à l'autosimilarité et l'évolution d'une relation d'échelle avec le redshift.

¹Par la suite, nous parlerons d'observable et de masse pour Y et Z par simplicité.

L'ajustement d'une relation d'échelle est donc celui d'une relation linéaire. Toutefois, plusieurs caractéristiques des données et du modèle rendent cet ajustement complexe. Il est notamment important de tenir compte de l'existence d'une dispersion intrinsèque autour de la relation d'échelle, de la différence entre estimateur de masse et masse réelle des amas, ainsi que des incertitudes de mesures sur les deux quantités, qui peuvent être corrélées. Ces caractéristiques constituent les sous-modèles du modèle bayésien hiérarchique, que nous décrivons ci-après.

Dispersion intrinsèque

En réalité, la relation masse-observable ne suit pas parfaitement une loi de puissance. Dans le cas de la relation $Y_{500} - M_{500}$, il existe un grand nombre de processus physiques complexes qui peuvent altérer le contenu en énergie thermique d'un amas sans être associés à un changement de masse. Ainsi, l'équation (7.17) n'est qu'une tendance, et pas une relation strictement suivie par tous les amas de galaxies. On parle alors de dispersion intrinsèque autour de la relation.

La dispersion intrinsèque d'une relation d'échelle revêt une importance capitale dans l'exploitation cosmologique d'un échantillon d'amas de galaxies. En effet, sa valeur peut être interprétée comme une mesure de la qualité d'une observable en tant que *proxy* de masse². Ainsi, une relation masse-observable très dispersée est caractéristique d'une observable peu corrélée à la masse. L'étalonnage en masse des amas est alors peu précis, et les contraintes sur les paramètres cosmologiques sont moins fortes. Ainsi, la valeur de la dispersion intrinsèque est un paramètre d'intérêt de l'ajustement d'une relation d'échelle, au même titre que sa pente et son ordonnée à l'origine.

La prise en compte de l'existence d'une dispersion intrinsèque autour de la relation d'échelle n'est pas triviale, et ne peut pas être obtenue simplement en généralisant un estimateur de χ^2 . Prenons l'exemple de l'ajustement d'un groupe de points $\{x_i, y_i\}, i \in [1 \dots N]$ par une droite d'équation $f(x) = \alpha + \beta x$. Les valeurs des paramètres α et β décrivant le modèle en meilleur accord avec les données peuvent être trouvées en minimisant :

$$\chi^2 = \frac{1}{N} \sum_i \frac{(y_i - (\alpha + \beta x_i))^2}{\sigma_{y,i}^2}, \quad (7.19)$$

où $\sigma_{y,i}$ représente l'incertitude de mesure sur la valeur y_i .

Afin de tenir compte d'une dispersion intrinsèque σ_{int} autour de la relation, on peut modifier l'équation (7.19) de sorte à inclure la dispersion intrinsèque au dénominateur, soit $\sigma_{y,i}^2 \rightarrow \sigma_{y,i}^2 + \sigma_{\text{int}}^2$. Si ce remplacement permet en effet de tenir compte de l'existence d'une dispersion intrinsèque, il ne permet pas de mesurer la valeur de ce paramètre : la valeur de σ_{int} minimisant l'équation (7.19) est $\sigma_{\text{int}} \rightarrow \infty$.

Il est possible de tenir compte du caractère non-déterministe de la relation masse-observable en la modélisant non pas comme une relation stricte, mais comme une distribution de probabilité. On peut notamment considérer une distribution gaussienne :

$$P(Y|Z) = \mathcal{N}\left(\alpha_{YZ} + \beta_{YZ}Z + \gamma T(z) + \delta T(z)Z, \sigma_{YZ}^2\right), \quad (7.20)$$

où $\mathcal{N}(\mu, \nu)$ est la distribution normale de moyenne μ et variance ν .

On note σ_{YZ} la dispersion intrinsèque de la relation d'échelle, encapsulant la variabilité des valeurs

²Le terme de *proxy* de masse fait référence à une observable corrélée à la masse, qui peut être utilisée pour prédire la masse d'un amas à l'aide d'une relation d'échelle.

d'observable Y pouvant être associés à une valeur de masse Z . Cette équation constitue l'un des sous-modèles permettant de structurer le modèle hiérarchique de relation d'échelle.

Estimateurs de masse

Comme nous l'avons discuté au chapitre 2, la masse hydrostatique mesurée par combinaison d'observations X et SZ est un estimateur biaisé de la vraie masse d'un amas. Aucun estimateur de masse n'est parfait. En effet, la masse hydrostatique est un estimateur certes biaisé, mais peu dispersé ; à l'inverse, les estimateurs de masse par lentillage gravitationnel sont moins biaisés, mais plus dispersés (voir par exemple [62, 63]). Une revue sur les estimateurs de masse d'amas de galaxies, discutant leur justesse et leur précision propre, peut être trouvée en [43].

Il est nécessaire de tenir compte des biais et dispersions caractéristiques des estimateurs de masse dans l'étalonnage d'une relation d'échelle. En effet, l'utilisation d'une relation d'échelle étalonnée à l'aide d'un estimateur de masse biaisé entraînera une estimation biaisée des masses des amas, et donc un biais dans les analyses cosmologiques. De même, un estimateur de masse dispersé induit un biais sur la pente de la relation d'échelle, un effet connu sous le nom de biais d'Eddington ([196, 197]). Afin de tenir compte des imperfections caractéristiques de l'estimateur de masse utilisé, on peut définir l'estimateur de masse X , relié à la masse réelle de l'amas Z au travers d'une distribution de probabilité :

$$P(X|Z) = \mathcal{N}\left(\alpha_{X|Z} + \beta_{X|Z}Z, \sigma_{X|Z}^2\right), \quad (7.21)$$

où $\alpha_{X|Z}$ est le biais de l'estimateur de masse, $\beta_{X|Z}$ quantifie l'évolution de ce biais avec la masse de l'amas, et $\sigma_{X|Z}$ est la dispersion de l'estimateur.

Dans le cas de la masse hydrostatique, $\alpha_{X|Z}$ est le biais hydrostatique b discuté en section 2.1.3 (page 28). L'équation (7.21) constitue un autre des sous-modèles de l'approche hiérarchique.

Incertitudes de mesure

La propagation des incertitudes de mesure dans l'ajustement d'une relation d'échelle est également d'une importance capitale (voir [105] pour une revue). Dans le cadre de l'étude de relations d'échelle masse-observable, la nature des données crée une complication supplémentaire dans l'ajustement. En effet, les valeurs d'observable et de masse mesurées comportent toutes deux des incertitudes. Or, il est difficile de réaliser un ajustement comportant des incertitudes à la fois en abscisse et en ordonnée. Il est toutefois possible d'inclure ces incertitudes dans un estimateur de χ^2 . De la même façon que pour la dispersion intrinsèque, l'équation (7.19) peut être généralisée en remplaçant son dénominateur comme $\sigma_{y,i}^2 \rightarrow \sigma_{y,i}^2 + \beta \sigma_{x,i}^2$. Tout comme pour l'inclusion d'une dispersion intrinsèque, cette approche conduit à des valeurs de paramètres biaisées, puisque la valeur de χ^2 peut être diminuée avec une augmentation de β , ce qui crée un biais vers les grandes valeurs de pente.

Le modèle *errors in variable* permet de tenir compte des incertitudes de mesures sur les deux axes de la relation d'échelle et de leur éventuelle corrélation en modélisant chaque point de données (les valeurs de masse et d'observable de chaque amas considéré dans l'ajustement) comme une distribution de probabilités multivariée. Pour un amas de galaxies i , on peut alors écrire :

$$P(x_i, y_i | X_i, Y_i) = \mathcal{N}_{2D}(\{X_i, Y_i\}, V_i), \quad (7.22)$$

où x_i et y_i sont les valeurs mesurées de l'estimateur de masse et de l'observable, X_i et Y_i sont les vraies valeurs de ces deux grandeurs, et V_i est la matrice de covariance des incertitudes sur x_i et y_i .

Ainsi, dans cette modélisation de la relation d'échelle, les valeurs réelles des estimateurs de masse et d'observable sont des paramètres de nuisance du modèle. L'équation (7.22) comporte une distribution de probabilité pour chaque point de données i , et constitue autant de sous-modèles pour le modèle hiérarchique.

Nous avons vu en section 5.3.1 que PANCO2 permettait de mesurer, pour un amas, la distribution de probabilité dans le plan $M_{500} - Y_{500}$ (voir figure 5.10). Ainsi, l'analyse avec PANCO2 de tous les amas de galaxies du grand programme SZ de NIKA2 permettra de mesurer pour chaque amas les grandeurs x_i , y_i , et V_i , qui seront les données d'entrée de l'ajustement de la relation d'échelle au travers de l'équation (7.22).

7.2.2 Effets de sélection et représentativité

Nous avons vu dans le chapitre 2 que la fonction de sélection d'un relevé d'amas devait être prise en compte dans l'exploitation cosmologique de ce dernier, afin de calculer le nombre d'amas observables en fonction des paramètres cosmologiques. De la même façon, un échantillon d'amas utilisé pour ajuster une relation d'échelle a subi une sélection. Cette sélection peut avoir un effet important sur l'ajustement de la relation d'échelle, et doit donc être prise en compte.

Biais de Malmquist

L'un des effets de sélection les plus fréquemment rencontrés en astronomie est le biais de Malmquist [198]. Celui-ci intervient lorsque l'échantillon considéré pour la régression est tronqué à une valeur d'observable minimale. Il est donc omniprésent en astronomie, puisque les relevés du ciel sont souvent limités en flux du fait de leur sensibilité. Dans le cadre d'une relation d'échelle masse-observable, il est présent lorsque l'échantillon considéré pour l'ajustement est composé uniquement d'amas dont l'observable est supérieure à un seuil donné. Ainsi, la significativité de détection, souvent utilisée comme observable corrélée à la masse en optique et SZ (voir section 2.3.2, page 42), est intrinsèquement affectée par le biais de Malmquist, puisque les catalogues construits à partir de cette grandeur sont souvent limités à une valeur seuil.

Le biais de Malmquist se manifeste comme un biais sur les valeurs des paramètres mesurés lors de l'ajustement de relations d'échelle. Il est lié à une surreprésentation des amas au dessus de la valeur seuil d'observable. Afin d'illustrer cet effet, un exemple est présenté en figure 7.2. Un ensemble de 100 valeurs de x sont générées aléatoirement à partir d'une distribution log-normale. Les valeurs de y correspondantes sont générées en considérant une relation linéaire comportant une dispersion intrinsèque, avec les paramètres suivants :

$$\alpha = 0; \quad \beta = 1; \quad \sigma = 0.25. \quad (7.23)$$

La relation d'échelle correspondante est représentée par une ligne noire pointillée sur le panneau gauche de la figure 7.2, et la valeur des paramètres dans le plan (α, β) de la même façon dans le panneau droit. L'ajustement³ des 100 points de cet échantillon permet d'obtenir une relation proche de celle utilisée pour générer les points, représentée par la ligne noire pleine. Les intervalles de confiance sur les paramètres α et β sont représentés en noir sur le panneau gauche de la figure 7.2, et sont compatibles avec la vérité.

On considère ensuite seulement les 60 points correspondant aux valeurs de y les plus élevés, soit,

³La méthode d'ajustement par MCMC sera détaillée en 7.3.

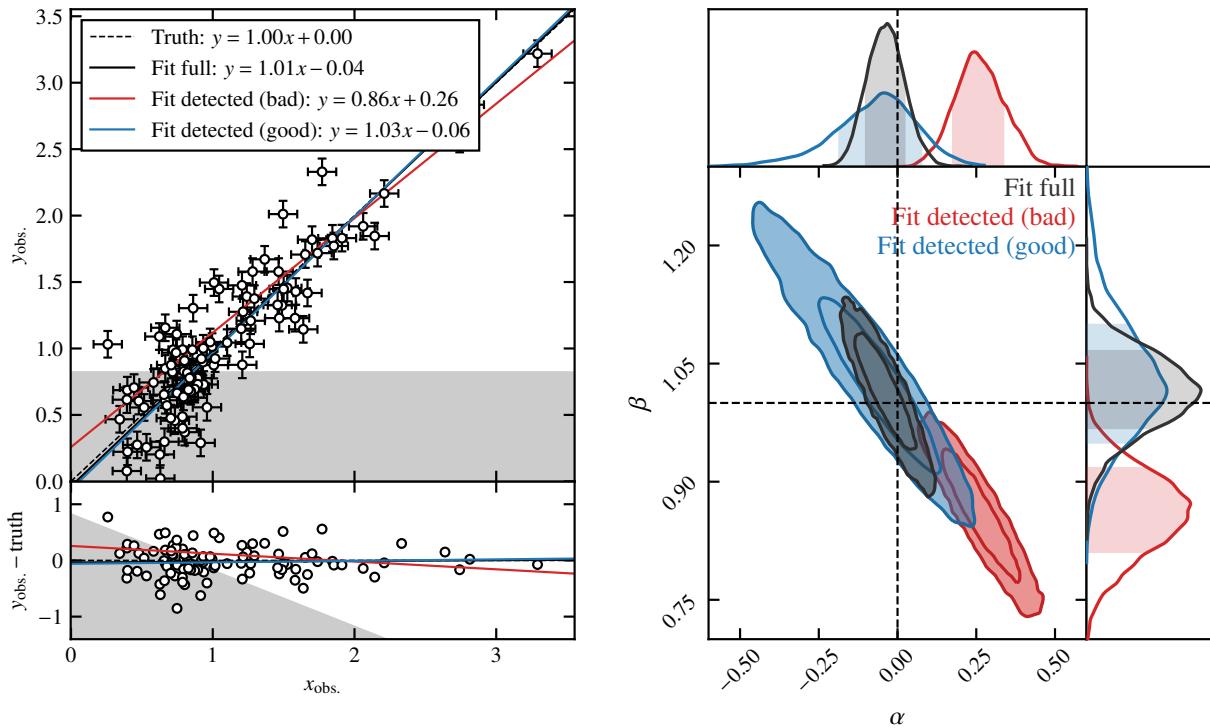


Figure 7.2 : Illustration du biais de Malmquist dans l'ajustement d'une relation d'échelle. **Gauche :** représentation des données et des relations d'échelle ajustées dans le plan \$(x, y)\$ (haut) et résidus autour de la relation d'échelle réelle (bas). **Droite :** intervalles de confiance à \$1\sigma\$ et \$2\sigma\$ dans le plan \$(\alpha, \beta)\$ obtenus par régression. Voir texte pour le détail des différentes régressions.

dans cet exemple, les points satisfaisant \$y > 0.88\$. Les 40 points non considérés sont donc les points de valeur de \$y\$ plus faible que ce seuil, correspondant à la région grisée. En répétant le même ajustement que celui réalisé précédemment – qui avait donné la droite noire – sur les points au-dessus de ce seuil, on obtient une relation d'échelle biaisée, représentée en rouge. Le biais sur les paramètres \$\alpha\$ et \$\beta\$ est clairement visible sur le panneau droit (en rouge également). La raison de ce biais est la surreprésentation des points au dessus du seuil en \$y\$. En effet, du fait de l'existence d'une dispersion intrinsèque, des points de faible valeur de \$x\$, dont l'observable \$y\$ aurait été inférieure au seuil en l'absence de dispersion intrinsèque, se retrouvent au dessus du seuil par une fluctuation statistique positive. À l'inverse, des points de plus grande valeur de \$x\$ se retrouvent sous le seuil, et donc non-détectés, du fait d'une fluctuation négative. Par conséquent, l'échantillon détecté comporte une surpopulation de points au dessus de la relation d'échelle réelle, introduisant un biais. Ce biais peut être compris visuellement grâce à la partie inférieure du panneau gauche de la figure 7.2. Celui-ci représente la différence entre les points de données et relation d'échelle du panneau supérieur et la relation réelle. On voit clairement la surreprésentation des points au dessus de la relation d'échelle dans la région du seuil en \$y\$.

Afin de réaliser un ajustement non-affecté par le biais de Malmquist, il est donc nécessaire de tenir compte de la présence d'une sélection en \$y\$. Dans le cadre de la modélisation *errors in variable*, cette prise en compte peut être réalisée en tronquant la distribution de probabilité associée à chaque point de données. Pour une coupure à une valeur de seuil \$y_{\text{th.}}\$, l'équation (7.22) peut être réécrite comme :

$$P(x_i, y_i | X_i, Y_i) \propto \mathcal{N}_{2D}(\{X_i, Y_i\}, V_i) \times H(y_i - y_{\text{th.}}), \quad (7.24)$$

où $H(x)$ est la fonction de Heaviside, définie comme

$$H(x) = \begin{cases} 1 & \text{si } x \geq 0, \\ 0 & \text{sinon.} \end{cases} \quad (7.25)$$

Cette troncature permet de supprimer toute l'information apportée par les données en dessous du seuil de détection, et donc la surreprésentation des points au dessus de la relation d'échelle réelle. Les résultats de l'ajustement des points détectés en tenant compte de la troncature (équation 7.24) sont représentés en bleu sur la figure 7.2. La droite obtenue est très proche de la relation d'échelle réelle utilisée pour générer les données. Cette similarité montre que la troncature permet de corriger le biais de Malmquist.

Distribution de masse

Parmi les effets de sélection pouvant causer des biais dans l'ajustement d'une relation linéaire, on compte également la modélisation de la distribution de la variable latente Z . Les échantillons de données utilisés pour l'ajustement de relations d'échelle⁴ ne sont en général pas représentatifs de la distribution d'objets dans l'Univers. Dans le cas des amas de galaxies, les échantillons considérés représentent souvent les amas les plus massifs de l'Univers (puisque ces derniers sont en général plus lumineux), et peuvent surreprésenter les amas possédant certaines propriétés physiques (par exemple les amas à cœur froid en X ; voir section 3.4.3, page 67). La méconnaissance de la distribution intrinsèque de la variable latente peut induire des biais sur les paramètres de la relation d'échelle, du fait de la non-représentativité de l'échantillon [199]. Ces effets sont intimement liés aux biais induits par l'utilisation d'estimateurs dispersés de la variable latente (la masse des amas dans ce contexte), et font donc partie des biais dits d'Eddington discutés en 7.2.1 (page 148) [196, 197].

Afin de tenir compte de la différence entre les distributions en masse de l'échantillon et de la population parente (c'est-à-dire tous les amas de l'Univers), il est possible d'inclure la distribution de Z dans la modélisation statistique. Cette approche a été introduite par Kelly [199] et reprise pour de nombreuses analyses de relations d'échelles masse-observable [50, 185, 200]. Dans cette approche, la distribution de la variable latente est modélisée comme une somme de n_{mix} gaussiennes pondérées :

$$P(Z) = \frac{1}{n_{\text{mix}}} \sum_{k=1}^{n_{\text{mix}}} \pi_k \mathcal{N}(\mu_k, \sigma_k^2), \quad (7.26)$$

où μ_k et σ_k^2 représentent les moyennes et variances de la gaussienne k , et π_k son poids⁵. Ce dernier peut être interprété comme la probabilité qu'un amas considéré dans un échantillon soit issu de la gaussienne k , de sorte à ce que $\sum_k \pi_k = 1$.

Les paramètres des gaussiennes π_k , μ_k et σ_k sont traités comme des paramètres de nuisance du modèle probabiliste de relation d'échelle ; les résultats finaux sont marginalisés sur ces paramètres. L'équation (7.26) représente le dernier sous-modèle du modèle bayésien hiérarchique de la relation d'échelle.

⁴Et en général en astronomie.

⁵Notons que les paramètres π_k , μ_k et σ_k pourraient évoluer avec le redshift, mais que nous ne considérerons pas de telles évolutions.

7.3 Ajustement de la relation d'échelle par MCMC

La section précédente a permis de détailler la modélisation probabiliste de la relation d'échelle adoptée pour les études menées au cours de cette thèse. L'objectif de l'ajustement de cette relation est l'obtention de la distribution de probabilité des paramètres d'intérêt de la relation d'échelle conditionnée par les données, $P(\vartheta|D)$. De même que pour l'ajustement des profils de pression avec PANC02 (voir chapitre 5), cette distribution peut être obtenue à partir du théorème de Bayes :

$$P(\vartheta|D) = \frac{P(D|\vartheta) P(\vartheta)}{P(D)}, \quad (7.27)$$

où ϑ est un vecteur représentant les paramètres de l'analyse, D représente les données, c'est-à-dire les valeurs de (x_i, y_i) et leurs covariances V_i . $P(D|\vartheta)$ représente la fonction de vraisemblance comparant le modèle aux données, obtenu par combinaison des équations (7.22), (7.20), (7.21) et (7.26). $P(\vartheta)$ représente la distribution *a priori* des paramètres, et $P(D)$ est une constante de normalisation.

Comme nous l'avons vu, le modèle statistique de relation d'échelle est complexe ; nous verrons de plus que son ajustement sur des données l'est également. Par conséquent, les études réalisées au cours de cette thèse se basent sur une librairie existante nommée LIRA (*LInear Regression in Astronomy*, [185]). Ce logiciel est disponible publiquement sous forme d'une librairie dans le langage R, dédié aux études statistiques et d'analyse de données. LIRA permet d'ajuster le modèle probabiliste détaillé dans la section précédente en utilisant un échantillonnage par Monte Carlo à Chaînes de Markov (MCMC). Cette section détaille la construction de la distribution *a priori* $P(\vartheta)$, et l'échantillonnage de la distribution postérieure $P(\vartheta|D)$ par MCMC.

7.3.1 Distribution de probabilité *a priori* des paramètres

L'approche bayésienne employée dans l'étude de la relation d'échelle nécessite la construction d'une distribution de probabilité *a priori* pour l'ensemble des paramètres de l'analyse. Les paramètres sont considérés comme indépendants, de sorte à ce que la distribution *a priori* puisse être écrite comme le produit des distributions *a priori* sur chaque paramètre individuel.

La table 7.1 présente un résumé des paramètres du modèle et de leur distribution *a priori*. Les distributions choisies sont volontairement peu informatives, de sorte à ce que les résultats de l'ajustement ne soient pas entachés d'erreurs dues à l'inclusion d'une connaissance *a priori* des paramètres erronée. Pour cela, la distribution uniforme \mathcal{U} est utilisée ; elle est définie comme :

$$\mathcal{U}(x_1, x_2) = \begin{cases} 1/|x_1 - x_2| & \text{entre } x_1 \text{ et } x_2, \\ 0 & \text{ailleurs.} \end{cases} \quad (7.28)$$

Les bornes de cette distribution doivent être finies, sans quoi la distribution *a priori* des paramètres ne peut plus être intégrable à 1, et ne représente donc plus une distribution de probabilité. Le choix des intervalles $[-10^4, 10^4]$ (ou $[0, 10^4]$ dans le cas des valeurs strictement positives) est arbitraire. Puisque les grandeurs en jeu dans l'ajustement sont mises à l'échelle (équation 7.18), ces intervalles sont suffisamment larges pour ne pas exclure de région de l'espace des paramètres pouvant fournir une bonne description des données.

Certains paramètres ne se voient pas attribuer une distribution *a priori* uniforme. Pour le paramètre

Paramètre	Description	Distribution <i>a priori</i>
<i>Paramètres d'intérêt : relation $Y = f(Z)$</i>		
α_{YZ}	Ordonnée à l'origine	$\mathcal{U}(-10^4, 10^4)$
β_{YZ}	Pente	Student t_1
σ_{YZ}	Dispersion intrinsèque	$\mathcal{U}(0, 10^4)$
γ	Évolution avec le redshift de la normalisation	$\delta(2/3)$
δ	Évolution avec le redshift de la pente	$\delta(0)$
<i>Paramètres de nuisance : relation $X = f(Z)$</i>		
$\alpha_{X Z}$	Ordonnée à l'origine	$\delta(0)$
$\beta_{X Z}$	Pente	$\delta(1)$
$\sigma_{X Z}$	Dispersion intrinsèque	$\delta(0)$
<i>Paramètres de nuisance : distribution de la variable latente</i>		
μ_k	Moyennes des gaussiennes	$\mathcal{U}(-10^4, 10^4)$
σ_k	Dispersions des gaussiennes	$\mathcal{U}(0, 10^4)$
π_k	Poids des gaussiennes	Dirichlet(n_{mix})

Table 7.1 : Liste des paramètres du modèle statistique de relation d'échelle masse-observable. La dernière colonne indique les distributions *a priori* par défaut pour chaque paramètre.

de pente β_{YZ} , la distribution de Student à un degré de liberté est utilisée :

$$t_1(x) = \frac{1}{\pi} (1 + x^2)^{-1}. \quad (7.29)$$

Une distribution $t_1(\beta)$ sur la pente d'une relation linéaire correspond à une distribution uniforme sur l'angle entre cette droite et l'axe des abscisses, $\theta = \arctan(\beta)$. L'utilisation de la distribution de Student revient donc à l'utilisation d'une distribution *a priori* non-informative en coordonnées polaires. Ce choix peut donner lieu à des valeurs de pentes moins biaisées vers les grandes valeurs de pente qu'un *a priori* uniforme sur β [105, 185].

Un autre groupe de paramètres pour lesquels la distribution *a priori* n'est pas uniforme est l'ensemble des poids des gaussiennes considérées pour la modélisation de la distribution intrinsèque de la variable latente, π_k , $k \in [0 \dots n_{\text{mix}}]$. La distribution adoptée est la distribution de Dirichlet, qui permet à la somme des poids d'être égale à 1.

Enfin, les paramètres quantifiant l'évolution de la relation d'échelle avec le redshift sont par défaut fixés à leur valeur autosimilaire, $\gamma = 2/3$ et $\delta = 0$ (équation 7.18).

Notons que la distribution *a priori* décrite dans cette section est la distribution par défaut pour un ajustement de relation d'échelle. Il est possible de spécifier une autre distribution pour chaque paramètre, selon les objectifs de l'analyse. Lors de la présentation de résultats dans les sections 7.4 et 7.5, les distributions choisies pour chaque paramètre seront spécifiées lorsqu'elles seront différentes des options par défaut.

7.3.2 Analyse MCMC par échantillonnage de Gibbs

Nous avons vu dans le chapitre 5 qu'une analyse par Monte Carlo à Chaînes de Markov (MCMC) permettait d'obtenir un échantillon de la distribution postérieure de probabilité $P(\vartheta|D)$ dans l'espace

des paramètres. Cette approche y est illustrée au travers de l'algorithme de Metropolis-Hastings, représentant l'algorithme de MCMC le plus simple à définir. Nous avons vu cependant que d'autres algorithmes existaient et fournissaient de meilleures performances pour l'échantillonnage de distributions de probabilités dans un grand nombre de dimensions. Par exemple, dans PANC02, l'échantillonneur choisi utilise une variation de l'algorithme de Metropolis-Hastings dont la fonction de proposition évolue avec les itérations.

Pour l'ajustement d'une relation d'échelle, il est utile de faire appel à un autre type d'échantillonnage, plus performant pour un très grand nombre de dimensions. En effet, l'utilisation de la modélisation *errors in variable* implique le traitement de chacune des grandeurs X_i et Y_i comme des paramètres de nuisance. Ainsi, pour un échantillon de N amas, l'analyse comporte $2N$ paramètres de nuisance, en plus de ceux du modèle résumés dans la table 7.1. Pour le grand programme SZ de NIKA2, cela représente un nombre de paramètres de l'analyse de l'ordre de la centaine, alors que l'ajustement des propriétés thermodynamiques d'un amas avec PANC02 n'utilise qu'une dizaine de paramètres.

Afin de répondre à ce besoin, un grand nombre de logiciels dédiés à l'ajustement de relations d'échelle (voir par exemple [185, 200]) font appel à l'échantillonnage de Gibbs (*Gibbs sampling*, voir par exemple le chapitre 11.3 de [201]). La principale différence entre cette technique et l'algorithme de Metropolis-Hastings est le fait que le mouvement dans l'espace des paramètres n'est pas fait pour tous les paramètres simultanément, mais par groupe de paramètres. Un vecteur de paramètres ϑ à N dimensions peut être divisé en d groupes :

$$\vartheta = (\vartheta_1, \dots, \vartheta_N) \rightarrow \theta = (\theta_1, \dots, \theta_d) = (\{\vartheta_1, \dots, \vartheta_{n_1}\}, \{\vartheta_{n_1+1} \dots \vartheta_{n_1+n_2}\}, \dots, \{\vartheta_{N-n_d}, \dots, \vartheta_N\}), \quad (7.30)$$

où n_i est la dimension du groupe i .

La procédure itérative générant une chaîne de Markov, détaillée en section 5.2.3 (page 107), est alors appliquée. La particularité de l'échantillonnage de Gibbs réside dans le fait qu'à chaque itération, l'échantillonneur boucle sur les d groupes pour proposer un mouvement dans l'espace des paramètres groupe par groupe, plutôt que pour tous les paramètres simultanément. De plus, le mouvement d'un groupe de paramètres est conditionné par celui des autres paramètres, afin de tenir compte des corrélations entre les paramètres de groupes différents. Dans le cas du modèle bayésien hiérarchique utilisé ici, les groupes de paramètres sont définis en fonction du sous-modèle auquel ils appartiennent. Par exemple, un groupe est constitué par les paramètres d'intérêt de la relation d'échelle (équation 7.20); un autre est consacré aux caractéristiques de l'estimateur de masse (équation 7.21), etc. La procédure employée pour générer un pas du MCMC dans le cas d'une régression linéaire utilisant la modélisation *errors in variable* est décrite dans la section 6.2.1 de [199], et comporte 18 étapes qui ne seront pas détaillées ici.

L'échantillonnage de Gibbs est un algorithme MCMC qui produit des chaînes de Markov, c'est-à-dire un échantillonnage de la distribution postérieure de probabilités dans l'espace des paramètres. Il offre une grande performance au vu du nombre important de paramètres dans l'analyse. L'échantillonnage réalisé dans LIRA utilise le logiciel JAGS⁶, dédié à l'échantillonnage de Gibbs pour des modèles bayésiens structurels. La marginalisation sur les paramètres de nuisance permet d'obtenir la distribution de probabilité des paramètres d'intérêt au vu des données. Cette distribution permet alors de calculer la probabilité $P(Y|Z)$ qu'un amas de masse Z soit détectable avec une valeur d'observable Y . Comme nous l'avons vu en section 2.3 (page 40), cette probabilité peut être utilisée pour l'étalonnage en masse d'un échantillon d'amas de galaxies en vue de son exploitation cosmologique (équations

⁶Just Another Gibbs Sampler, <https://mcmc-jags.sourceforge.io/>

??).

7.4 Validation sur données simulées

Nous avons vu dans les sections précédentes que le modèle bayésien hiérarchique utilisé dans cette étude pouvait tenir compte de plusieurs spécificités des jeux de données d'astronomie, et que l'ajustement pouvait être fait à l'aide de la librairie LIRA. Afin de tester la capacité de LIRA à livrer des mesures justes des paramètres d'une relation linéaire, nous commençons par générer des jeux de données simples, comportant les différentes caractéristiques développées en section 7.2 (page 146). Ces jeux de données n'ont pas vocation à être des réalisations réalistes pour des relations d'échelle masse-observable ; de telles analyses seront présentées par la suite.

7.4.1 Incertitudes de mesure et dispersion intrinsèque

Nous commençons par générer des données simples, en considérant la présence d'incertitudes de mesures décorrélées sur les deux axes, et l'existence d'une dispersion intrinsèque. Un total de cent points i sont générés de la façon suivante :

$$\begin{aligned} Z_i &\sim \mathcal{N}(0, 1); \\ X_i &\sim \mathcal{N}(\alpha_{X|Z} + Z_i, \sigma_{X|Z}^2) \quad (\text{soit } \beta_{X|Z} = 1); \\ Y_i &\sim \mathcal{N}(\alpha_{Y|Z} + \beta_{Y|Z} Z_i, \sigma_{Y|Z}^2); \\ x_i &\sim \mathcal{N}(X_i, 0.01); \\ y_i &\sim \mathcal{N}(Y_i, 0.01), \end{aligned} \tag{7.31}$$

où $A \sim B$ signifie « A est distribué suivant B », et les valeurs utilisées pour les paramètres sont consignées en table 7.2.

L'échantillon des (x_i, y_i) est ensuite ajusté avec LIRA, en considérant les distributions *a priori* par défaut (voir table 7.1), sauf pour les paramètres de la relation $X = f(Z)$ qui sont fixés. La liste des valeurs réelles des paramètres et de leurs distributions *a priori* pour cette analyse est donnée en table 7.2.

L'objectif de cette analyse est de pouvoir estimer si les paramètres d'intérêt de la relation d'échelle sont correctement reconstruits par LIRA. Pour cela, nous définissons deux grandeurs permettant de quantifier le biais des estimateurs de paramètres donnés par LIRA : pour un paramètre ϑ ,

$$\begin{aligned} \xi_\vartheta &\equiv \frac{\text{Med}[\vartheta_i]_i - \hat{\vartheta}}{|\hat{\vartheta}|}, \\ \zeta_\vartheta &\equiv \frac{\text{Med}[\vartheta_i]_i - \hat{\vartheta}}{\sqrt{\text{Var}[\vartheta_i]_i}}, \end{aligned} \tag{7.32}$$

où ϑ_i représente l'échantillon i de la chaîne de Markov sur le paramètre ϑ , et $\hat{\vartheta}$ est la valeur réelle du paramètre, utilisée pour générer l'analyse. Les opérations $\text{Med}[\dots]_i$ et $\text{Var}[\dots]_i$ représentent la médiane et la variance des échantillons.

Ces deux grandeurs permettent de comparer la valeur de la médiane des chaînes, utilisée comme estimateur des paramètres, à la valeur réelle du paramètre. ξ permet d'exprimer le biais de l'estimateur en fraction de la valeur réelle, tandis que ζ mesure le biais par rapport à l'écart-type des chaînes, et

Paramètre ϑ	Paramètres de l'analyse		Résultats		
	Valeur d'entrée	Distribution <i>a priori</i>	ξ_ϑ [%]	ζ_ϑ [σ]	η_ϑ [%]
<i>Dispersion intrinsèque seulement</i> (section 7.4.1)					
$\alpha_{Y Z}$	1	$\mathcal{U}(-10^4, 10^4)$	-0.03	-0.02	1.74
$\beta_{Y Z}$	1	Student t_1	0.03	0.01	1.76
$\sigma_{Y Z}$	0.5	$\mathcal{U}(0, 10^4)$	-4.88	-0.13	23.87
<i>Estimateur de masse dispersé</i> (section 7.4.2)					
$\alpha_{Y Z}$	1	$\mathcal{U}(-10^4, 10^4)$	0.27	0.00	2.11
$\beta_{Y Z}$	1	Student t_1	-0.99	-0.01	2.39
$\sigma_{Y Z}$	0.5	$\mathcal{U}(0, 10^4)$	-4.94	-0.13	43.70
$\sigma_{X Z}$	0.1	$\mathcal{U}(0, 10^4)$	-	-	-
<i>Estimateur de masse biaisé</i> (section 7.4.3)					
$\alpha_{Y Z}$	1	$\mathcal{U}(-10^4, 10^4)$	0.02	0.00	5.88
$\beta_{Y Z}$	1	Student t_1	0.04	0.02	1.76
$\sigma_{Y Z}$	0.5	$\mathcal{U}(0, 10^4)$	-4.91	-0.13	23.86
$\alpha_{X Z}$	0.8	$\mathcal{U}(0.6, 1)$	-	-	-
<i>Données tronquées</i> (section 7.4.4)					
$\alpha_{Y Z}$	1	$\mathcal{U}(-10^4, 10^4)$	-2.08	-0.56	3.47
$\beta_{Y Z}$	1	Student t_1	1.79	0.53	3.26
$\sigma_{Y Z}$	0.5	$\mathcal{U}(0, 10^4)$	-14.15	-0.58	24.84

Table 7.2 : Liste des valeurs d'entrée des paramètres, de leurs distributions *a priori* et des résultats pour les analyses de validation présentées en section 7.4. La distribution *a priori* des paramètres non-répertoriés dans cette table est celle par défaut (table 7.1). Lorsqu'elles ne sont pas précisées, les valeurs d'entrée des paramètres $\alpha_{X|Z}$ et $\sigma_{X|Z}$ sont nulles. Les biais et dispersions des estimateurs des paramètres $\alpha_{X|Z}$ et $\sigma_{X|Z}$ ne sont pas répertoriés car ceux-ci sont des paramètres de nuisance, dont les contraintes ne sont pas l'objectif de l'analyse.

peut donc être interprété comme la significativité du biais. Enfin, la variance des chaînes est également importante, puisqu'elle permet de quantifier l'incertitude sur ce paramètre. On définit alors l'écart-type relatif sur un paramètre comme :

$$\eta_\vartheta \equiv \frac{\sqrt{\text{Var}[\vartheta_i]}_i}{|\hat{\vartheta}|}. \quad (7.33)$$

La procédure de génération d'un échantillon aléatoire et d'ajustement d'une relation linéaire avec LIRA est répété 5000 fois. L'objectif est la mesure des performances de LIRA à travers l'étude de la distribution des paramètres ξ_ϑ , ζ_ϑ et η_ϑ pour chaque paramètre d'intérêt ϑ . Pour chaque réalisation, ces trois grandeurs sont calculées pour chacun des paramètres d'intérêt $\alpha_{Y|Z}$, $\beta_{Y|Z}$, et $\sigma_{Y|Z}$. Les valeurs moyennes pour chacun des paramètres d'intérêt de la relation d'échelle sont présentées dans la partie droite de la table 7.2. On voit que l'analyse permet d'obtenir des résultats très peu biaisés pour tous les paramètres, soit $\xi \ll 100\%$ et $\zeta \ll 1\sigma$ pour tous les paramètres d'intérêt. La dispersion des estimateurs des paramètres η est également calculée. Les données utilisées pour la validation n'ayant pas pour but d'être réalistes, la valeur de η ne présente pas d'intérêt majeur, si ce n'est d'être comparée entre les différentes analyses. En effet, nous verrons par la suite que l'ajout de paramètres de nuisance dans l'analyse augmente la dispersion des estimateurs des paramètres d'intérêt.

7.4.2 Estimateur de masse dispersé

La procédure présentée en 7.4.1 a permis de confirmer que l'estimation des paramètres d'intérêt de la relation d'échelle n'était pas biaisée dans le cas d'une relation d'échelle comportant des incertitudes sur la variable latente et sur la réponse, et une dispersion intrinsèque non-nulle. Nous cherchons ensuite à généraliser la validation aux cas de données plus complexes. Nous commençons par considérer un estimateur de la variable latente dispersé, c'est-à-dire $\sigma_{X|Z} \neq 0$. Cette dispersion correspond, dans le cas de l'étude d'une relation masse-observable, à l'utilisation d'un estimateur dispersé de la masse d'un amas de galaxies, comme la masse mesurée par lentillage gravitationnel. Le paramètre $\sigma_{X|Z}$ devient donc un paramètre de nuisance dans l'ajustement, plutôt que d'être fixé comme nul dans la section 7.4.1. Si cette dispersion n'est pas prise en compte dans l'analyse, elle génère un biais de type Eddington [196, 197].

Les paramètres de l'analyse, leur distribution *a priori* et les résultats pour les estimateurs de biais ξ et ζ et de dispersion η sont présentées en table 7.2. Tout comme dans la section précédente, on voit que les estimateurs pour les paramètres d'intérêt de la relation d'échelle sont très peu biaisés. En revanche, on voit que la dispersion des paramètres η augmente avec l'ajout de $\sigma_{X|Z}$ comme paramètre de nuisance. Cette augmentation de la dispersion des estimateurs correspond à la propagation de l'incertitude sur le paramètre ajouté. Elle montre donc l'effet de la prise en compte d'incertitudes systématiques dans l'analyse. L'ajout de paramètres de nuisance permet donc d'augmenter la justesse des estimateurs – dans ce cas, en corrigeant le biais d'Eddington – au détriment de leur précision. Il est donc important de procéder avec précaution lors de l'ajout de paramètres de nuisances dans l'analyse, au risque d'augmenter artificiellement les incertitudes sur les paramètres d'intérêt.

7.4.3 Estimateur de masse biaisé

Nous nous intéressons ensuite à l'ajustement d'une relation d'échelle utilisant un estimateur biaisé pour la variable latente. Ce cas correspond à une valeur non-nulle du paramètre $\alpha_{X|Z}$, qui devient donc un paramètre de l'analyse. Dans le cas de l'étude d'une relation masse-observable, ce biais correspond à l'utilisation d'un estimateur de masse biaisé, par exemple la masse hydrostatique, dont le biais a été décrit en 2.1.3 (page 28).

Nous avons choisi d'adopter un *a priori* informatif sur le biais de l'estimateur de variable latente, avec une distribution uniforme étroite centrée sur la valeur réelle du biais utilisé. En effet, il est possible de mesurer le biais hydrostatique, et les analyses de données et de simulations indiquent que sa valeur est comprise entre 0.6 et 1 (voir figure 2.1). Ainsi, l'utilisation d'un *a priori* informatif est justifiée, et permet de réduire l'espace des paramètres sur la base d'arguments physiques. Or, comme nous l'avons vu dans la section précédente, l'ajout de paramètres de nuisance diminue la précision des estimateurs de paramètres obtenus par l'ajustement de relations d'échelle.

Les résultats de l'analyse sont consignés dans la troisième partie de la table 7.2. Les valeurs moyennes des biais ξ et ζ nous montrent que les estimateurs des paramètres d'intérêt ne sont pas biaisés. Les valeurs de dispersions η montrent également que les estimateurs sont aussi précis que ceux pour des estimateurs de masse non-biaisés et non-dispersés (section 7.4.1), sauf pour le paramètre $\alpha_{Y|Z}$, dont l'estimateur est trois fois moins précis. Ainsi, une méconnaissance du biais sur les estimateurs de masse entraîne une imprécision sur la valeur de l'ordonnée à l'origine de la relation masse-observable.

7.4.4 Données tronquées

La procédure de correction du biais de Malmquist présentée en section 7.2.2 (page 149) est également testée. Comme l'a montré la figure 7.2, l'ajustement d'une relation sur des données tronquées à une valeur d'observable mène à un biais sur les estimateurs des paramètres si cette troncature n'est pas prise en compte. Afin de s'assurer que les estimateurs des paramètres d'intérêt de la relation d'échelle sont effectivement non-biaisés lorsque la coupure est prise en compte, nous répétons l'analyse réalisée en section 7.4.1 avec des échantillons tronqués. Pour cela, des échantillons de 200 points sont créés en utilisant la procédure décrite dans l'équation (7.31), dont seuls sont conservés les points i dont la valeur d'observable mesurée satisfait $y_i > 1$. Cette coupure permet de créer des échantillons d'environ 100 points, comparables à ceux utilisés dans les tests précédents. L'ajustement est ensuite réalisé en considérant la coupure au travers de l'équation (7.24).

Les résultats sont consignés dans la table 7.2. Les estimateurs des paramètres d'intérêt sont légèrement plus biaisés que lors de l'ajustement d'une relation sur des échantillons non-affectés par une sélection. Toutefois, ces biais restent faibles et peu significatifs, avec des valeurs de ζ de l'ordre de 0.5σ . La dispersion η des estimateurs de la pente et de l'ordonnée à l'origine de la relation d'échelle est également augmentée, par un facteur d'environ deux.

Conclusions

Les différentes caractéristiques des jeux de données utilisés pour l'ajustement de relations d'échelle masse-observable peuvent être pris en compte par le modèle probabiliste présenté en section 7.2. L'ajustement de ce modèle avec LIRA permet d'obtenir des estimateurs peu biaisés des paramètres d'intérêt de la relation d'échelle. La dispersion de ces estimateurs comprend la propagation des effets systématiques, qui sont considérés par l'ajout de paramètres de nuisance dans l'ajustement. Dans une analyse cosmologique, cette dispersion implique des contraintes moins fortes sur les paramètres cosmologiques. Il est donc utile de s'interroger en amont sur la pertinence physique des différents paramètres de nuisance utilisés, afin de pouvoir éventuellement diminuer l'espace des paramètres accessible et ainsi augmenter la précision des analyses, sans pour autant compromettre leur justesse.

7.5 Application au grand programme SZ de NIKA2

L'estimation des paramètres de la relation d'échelle liant l'observable SZ Y_{500} et la masse M_{500} est l'un des enjeux majeurs du grand programme SZ de NIKA2. La procédure d'ajustement développée et testée dans ce chapitre a pour objectif de fournir une estimation juste et précise de ces paramètres à partir des mesures de Y_{500} et M_{500} effectuées à partir des observations NIKA2, et plus particulièrement grâce au logiciel PANC02, comme décrit en section 5.2.4 (page 111). Dans cette section, nous quantifions le biais et la précision des estimateurs des paramètres de la relation d'échelle au moyen de simulations d'échantillons d'amas réalistes, similaires à celui du grand programme SZ.

7.5.1 Simulation d'échantillons

La fonction de sélection du grand programme SZ est complexe. Elle consiste en une couverture homogène du plan masse-redshift dans un intervalle donné, en sélectionnant des amas de galaxies détectés par leur effet SZ dans les données du satellite *Planck* et de l'Atacama Cosmology Telescope (ACT). L'homogénéité est assurée par la distribution des amas dans des intervalles de masse et de redshift, nommés « boîtes », résultant en l'échantillon représenté en figure 3.10.

Afin d'étudier la relation d'échelle du grand programme SZ, nous utilisons une analyse Monte-Carlo basée sur la génération d'échantillons d'amas similaires au LPSZ. Ces échantillons sont générés en suivant la procédure suivante, schématisée en figure 7.3 :

1. Un échantillon de halos est tiré aléatoirement à partir d'une fonction de masse donnée. Nous choisissons une fonction de masse de Tinker *et al.* [32], avec les paramètres cosmologiques mesurés par *Planck* [13]. Comme nous l'avons vu en section 2.3 (page 40), le produit de la fonction de masse et du volume comobile donne la distribution d'amas en masse et en redshift attendue dans l'Univers. Le résultat du tirage est donc un ensemble de points dans le plan (M_{500}, z) dont la distribution dans ce plan est réaliste.
2. Une relation d'échelle fiduciale est utilisée pour associer des valeurs d'observable SZ à chaque amas simulé. Nous choisissons les valeurs des paramètres de la relation mesurés par la collaboration *Planck* [148, 202], donné par l'équation (7.16) :

$$\alpha_{Y|Z} = -0.19 ; \quad \beta_{Y|Z} = 1.79 ; \quad \sigma_{Y|Z} = 0.075. \quad (7.34)$$

Les valeurs des paramètres d'évolution avec le redshift sont fixées à leur valeur par défaut, $\gamma = 2/3$ et $\delta = 0$. Ainsi, chaque amas se voit attribuer une valeur d'observable SZ selon sa masse et son redshift. Puisque nous n'étudions pas l'évolution de la relation avec le redshift au-delà de la prédiction autosimilaire, nous définissons l'observable SZ comme :

$$Y \equiv \log_{10} \left[E^{-2/3}(z) \frac{Y_{500}}{10^{-4} \text{ Mpc}^2} \right], \quad (7.35)$$

et la relation d'échelle de l'équation (7.20) devient :

$$P(Y|Z) = \mathcal{N}\left(\alpha_{Y|Z} + \beta_{Y|Z}Z, \sigma_{Y|Z}^2\right). \quad (7.36)$$

3. La sélection en boîtes du grand programme SZ est appliquée à la population d'amas simulée. Cinq amas sont sélectionnés pour chaque boîte, selon leur valeur de redshift et d'observable SZ. L'échantillon est alors réduit à un ensemble de 50 amas, dont la distribution en masse et en redshift est similaire à celle de l'échantillon du grand programme SZ.
4. Des incertitudes de mesures sont ajoutées aux valeurs de masse et d'observable. Afin d'utiliser des incertitudes réalistes, nous réalisons l'ajustement des propriétés thermodynamiques d'un amas simulé idéal avec PANCO2. En considérant un bruit blanc d'amplitude réaliste pour le grand programme SZ et un filtrage réaliste également, nous trouvons des incertitudes relatives de 12% et 13% sur Y_{500} et M_{500} respectivement, et un coefficient de corrélation de 84%.

Notons que cette procédure ignore un élément important de la fonction de sélection du grand programme SZ, qui est la sélection dans les catalogues *Planck* et ACT. En ne tenant pas compte des fonctions de sélection de ces deux relevés, nous émettons l'hypothèse que les catalogues d'amas qu'ils ont permis de construire sont représentatifs de la population d'amas dans l'Univers. Nous ignorons ainsi les mauvaises estimations de masse des amas dues à la présence de bruit corrélé dans les cartes. Une étude complète se devra de tenir compte de ces effets, afin de ne pas sous-estimer les incertitudes sur les paramètres de la relation d'échelle.

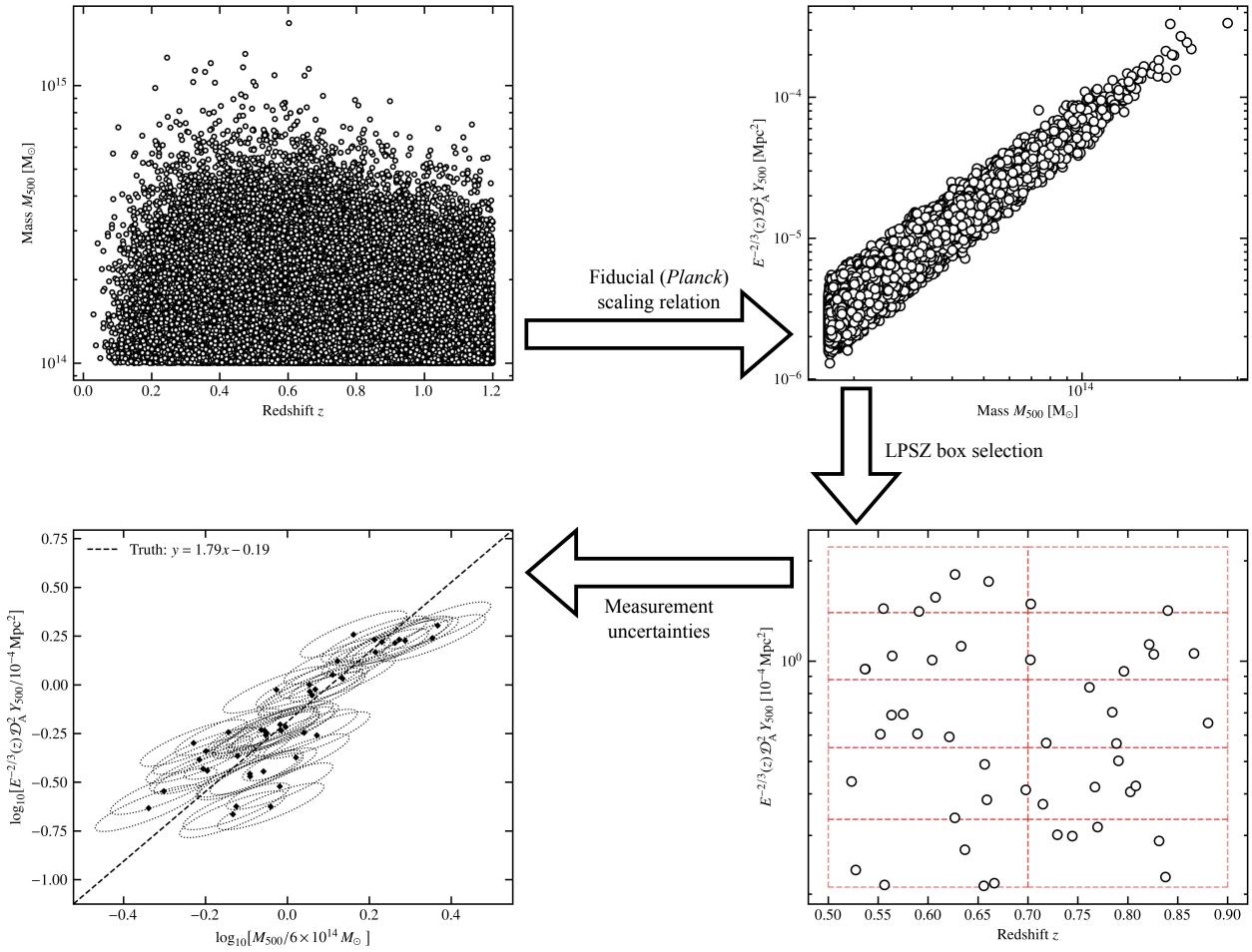


Figure 7.3 : Procédure de génération d'un échantillon d'amas similaire à celui du LPSZ. Un échantillon de halos est généré aléatoirement dans le plan (M_{500}, z) à partir d'une fonction de masse donnée (haut gauche). Une relation d'échelle fiducielle est utilisée pour obtenir des valeurs d'observable SZ correspondantes (haut droit). Une sélection similaire à celle du LPSZ sur les valeurs d'observable SZ est réalisée pour obtenir un échantillon de 50 amas (bas droit). Des incertitudes de mesure corrélées sur la masse et sur l'observable sont ensuite générées (bas gauche). L'ellipse autour de chaque point représente l'intervalle 1σ de la gaussienne à deux dimensions utilisée comme incertitude.

7.5.2 Modélisation et impact de la sélection

La fonction de sélection du grand programme SZ, reproduite pour générer des échantillons aléatoires, peut avoir un impact sur l'ajustement de la relation d'échelle. Toutefois, il est difficile de l'inclure dans le modèle bayésien hiérarchique mis en place par LIRA. Une première idée pourrait être, pour chaque amas, de définir une valeur de seuil en observable correspondant à la limite inférieure de la boîte dans laquelle il a été sélectionné. La troncature utilisée pour corriger du biais de Malmquist, décrite en section 7.2.2, pourrait alors être utilisée indépendamment pour chaque boîte. Cette procédure n'est toutefois pas correcte : il est faux de dire qu'un amas détecté dans une boîte i n'aurait pas pu faire partie de l'échantillon si sa valeur d'observable était inférieure à la limite basse de la boîte, alors qu'il aurait de fait pu être sélectionné dans la boîte inférieure $i - 1$.

Une façon de tenir compte de la sélection est donc de considérer une troncature des données à la valeur d'observable définie par la limite inférieure Y_{low} de la boîte de plus basse masse de l'échantillon. Afin d'évaluer la validité de cette modélisation, la procédure détaillée en 7.5.1 est utilisée pour générer 5000 échantillons fictifs. La relation d'échelle est ajustée deux fois pour chaque échantillon. Dans un

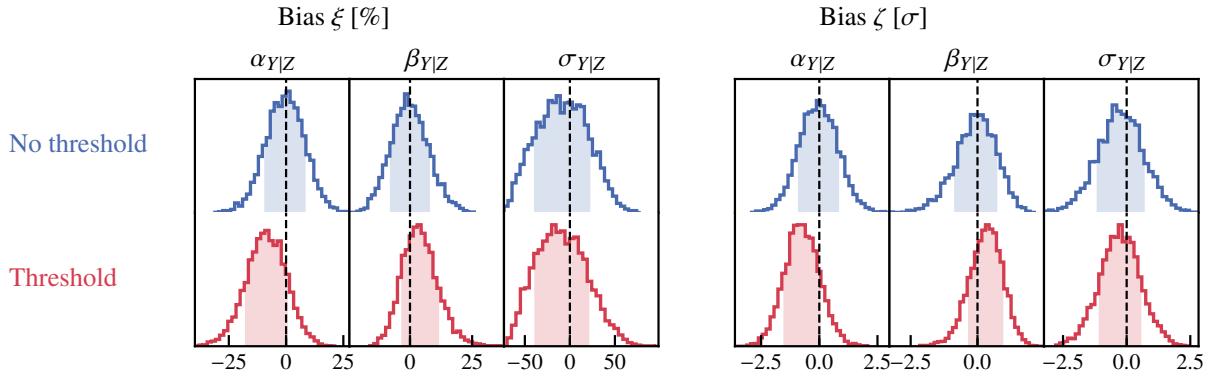


Figure 7.4 : Distribution des biais ξ (gauche) et ζ (droite) sur les trois paramètres d'intérêt de la relation d'échelle pour les 5000 réalisations d'échantillons. Les ajustements sont réalisés sans tenir compte de coupure en observable (bleu) et en tenant compte d'une coupure à Y_{low} (rouge). Les régions colorées correspondent aux intervalles de confiance à 1σ .

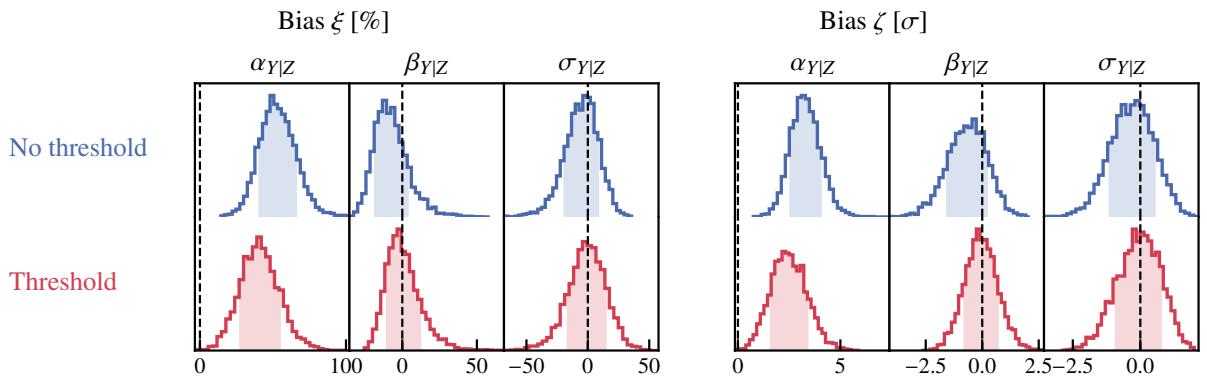


Figure 7.5 : Biais ξ et ζ sur les paramètres d'intérêt de la relation d'échelle pour une dispersion intrinsèque deux fois supérieure à celle supposée par *Planck* ($\sigma_{Y|Z} = 0.15$). La légende est identique à celle de la figure 7.4.

premier temps, nous ne considérons pas de coupure en observable, ce qui revient à ignorer la sélection en boîtes. L'ajustement est ensuite réalisé en tenant compte d'une coupure inférieure en observable à la valeur de Y_{low} .

Les distributions des valeurs de biais ξ et ζ pour ces ajustements sont présentées en figure 7.4. On observe que les résultats ne sont pas biaisés, avec des valeurs moyennes de ζ de l'ordre de 0.05σ . Lors de la prise en compte d'une coupure en observable à une valeur de Y_{low} , les résultats sont légèrement biaisés ($\zeta \sim 0.5\sigma$). Ce résultat indique que la fonction de sélection du grand programme SZ n'est pas bien décrite par une coupure en observable à la valeur la plus basse pouvant être sélectionnée pour la construction de l'échantillon.

L'absence de biais pour l'analyse ignorant les effets de sélection est surprenante, car elle semble indiquer que la fonction de sélection complexe du grand programme SZ n'induit aucun biais sur la mesure de la relation d'échelle masse-observable. Afin de comprendre ce résultat, nous avons répété l'analyse précédente en considérant une valeur de dispersion intrinsèque deux fois plus grande ($\sigma_{Y|Z} = 0.15$). Comme nous l'avons vu en 7.2.2 (page 149), le biais de Malmquist est conditionné par l'existence d'une dispersion intrinsèque. En effet, c'est cette dispersion qui permet à certains amas d'être détectés avec une valeur d'observable plus élevée que celle prédictive par la relation linéaire masse-observable, ce qui induit une surreprésentation des points au dessus de la relation. Ainsi, en augmentant la dispersion intrinsèque de la relation, les effets de sélection sont amplifiés, et pourront être plus facilement

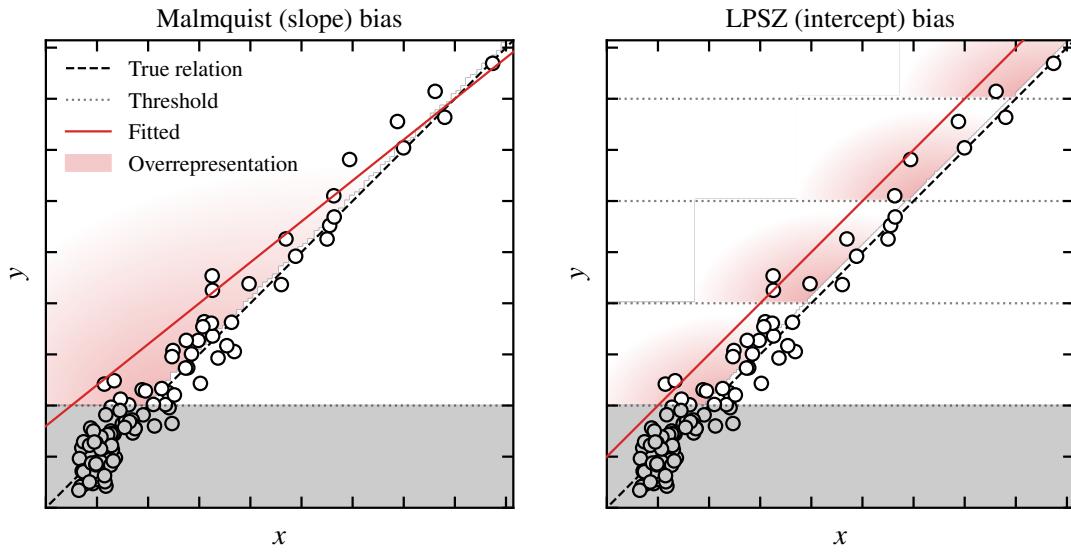


Figure 7.6 : Illustration de la cause du biais sur l’ordonnée à l’origine. Pour une coupure en observable (*gauche*), une surreprésentation de points au dessus du seuil (région rouge) entraîne une rotation de la relation d’échelle ajustée (droite rouge), et donc un biais sur la pente. Pour une sélection en boîtes comme celle du LPSZ (*droite*), des coupures correspondant aux seuils des boîtes créent des surreprésentations au dessus de ces seuils. Le résultat est un décalage de la relation d’échelle ajustée, correspondant à un biais sur l’ordonnée à l’origine.

détectables.

Les résultats de cette analyse sont présentés en figure 7.5. Les effets induits par la sélection de l’échantillon du grand programme SZ y apparaissent très clairement pour cette valeur de dispersion. Ceux-ci se manifestent sous la forme d’un biais significatif sur l’ordonnée à l’origine de la relation $\alpha_{\gamma Z}$, avec une valeur moyenne de $\zeta \simeq 3\sigma$. En revanche, à l’inverse du biais de Malmquist habituel (voir figure 7.2), aucun biais significatif n’est détecté sur la pente de la relation $\beta_{\gamma Z}$. Le biais sur l’ajustement de la relation d’échelle induit par la fonction de sélection du grand programme SZ est donc singulier. De plus, la prise en compte des effets de sélection par une coupure en observable ne permet pas de corriger ce biais.

Afin de comprendre l’origine du biais sur l’ordonnée à l’origine de la relation d’échelle, nous pouvons appliquer la même réflexion que celle permettant d’expliquer le biais de Malmquist. Comme nous l’avons vu en section 7.2.2 (page 149), le biais de Malmquist est dû à une surreprésentation des points au dessus d’un seuil en observable, qui entraîne une inclinaison de la pente de la relation. Dans le cas du grand programme SZ, les données sont composées de points sélectionnés en boîtes, créant des surreprésentations régulièrement espacées le long de la relation d’échelle. Par conséquent, une droite ajustée passant par les régions surreprésentées sera décalée vers le haut, correspondant bien à un biais sur l’ordonnée à l’origine. Cette interprétation de l’origine du biais est illustrée en figure 7.6, comparant les surreprésentations des points dans le cas d’une coupure en observable (*gauche*) et d’une sélection en boîtes (*droite*).

La sélection du LPSZ entraîne donc un biais sur l’ordonnée à l’origine de la relation d’échelle. La valeur de ce biais dépend de la dispersion intrinsèque autour de la relation d’échelle. Si la valeur de $\sigma_{\gamma Z}$ supposée par la collaboration *Planck* [148] est réaliste, le biais est complètement négligeable. Si cette valeur est sous-estimée, le biais peut être significatif. Afin de pouvoir arriver à des résultats non-

Paramètre ϑ	Distribution <i>a priori</i>	$\eta_\vartheta [\%]$
<i>Analyse de base</i>		
$\alpha_{Y Z}$	$\mathcal{U}(-10^4, 10^4)$	10.2
$\beta_{Y Z}$	Student t_1	10.1
$\sigma_{Y Z}$	$\mathcal{U}(0, 10^4)$	34.0
<i>Analyse de base + biais hydrostatique</i>		
$\alpha_{Y Z}$	$\mathcal{U}(-10^4, 10^4)$	232.0
$\beta_{Y Z}$	Student t_1	11.8
$\sigma_{Y Z}$	$\mathcal{U}(0, 10^4)$	35.2
$\alpha_{X Z}$	$\mathcal{U}(-0.4, 0.4)$	–
<i>Analyse de base + estimateur de masse dispersé</i>		
$\alpha_{Y Z}$	$\mathcal{U}(-10^4, 10^4)$	14.7
$\beta_{Y Z}$	Student t_1	16.8
$\sigma_{Y Z}$	$\mathcal{U}(0, 10^4)$	45.7
$\sigma_{X Z}$	$\mathcal{U}(0, 10^4)$	–

Table 7.3 : Liste des distributions *a priori* et des incertitudes relatives moyennes η pour les paramètres d'intérêt de la relation d'échelle dans le cas du LPSZ. Les dispersions des estimateurs des paramètres $\alpha_{X|Z}$ et $\sigma_{X|Z}$ ne sont pas répertoriées car ceux-ci sont des paramètres de nuisance, dont les contraintes ne sont pas l'objectif de l'analyse.

biaisés, il sera possible d'étudier l'évolution de la valeur de ce biais avec la dispersion intrinsèque grâce à des simulations, grâce à la répétition des études présentées dans cette section pour plusieurs valeurs de dispersion intrinsèque. Comme le montrent les figures 7.4 et 7.5, les estimations de la dispersion intrinsèque de la relation d'échelle ne sont pas biaisées pour les deux valeurs testées. Par conséquent, il sera possible de connaître la valeur de la dispersion intrinsèque à l'issue du grand programme SZ, et de corriger la valeur de l'ordonnée à l'origine *ad hoc* grâce aux résultats obtenus sur des simulations.

7.5.3 Incertitudes sur les paramètres d'intérêt

En plus du biais des estimateurs des paramètres de la relation d'échelle, nous pouvons étudier la dispersion de ces derniers, donnée par la grandeur η définie dans l'équation (7.33). Les valeurs obtenues sont une incertitude relative moyenne de 10% pour les paramètres $\alpha_{Y|Z}$ et $\beta_{Y|Z}$, et de 34% pour la dispersion intrinsèque $\sigma_{Y|Z}$. Ces valeurs sont reportées en table 7.3. De telles incertitudes sont comparables avec celles obtenues pour la même relation d'échelle par la collaboration *Planck* [148]. Cependant, comme nous l'avons discuté en section 7.1.2 (page 144), les incertitudes obtenues dans cette étude ne sont pas directement comparables à celles obtenues ici, car l'analyse bayésienne hiérarchique utilisée propage des incertitudes systématiques qui ne sont pas prises en compte dans l'analyse de [148]. Il est plus juste de comparer ces résultats avec ceux de CoMaLit [50], utilisant le même échantillon d'amas que la collaboration *Planck* et une régression similaire à celle utilisée ici. Les incertitudes de 10% sur les paramètres d'intérêt représentent une grande amélioration, due à la qualité des contraintes sur Y_{500} et M_{500} attendues dans le cadre du grand programme SZ de NIKA2 et à la couverture en masse de son échantillon.

L'analyse présentée précédemment fixe certains des paramètres de nuisance inclus dans le modèle hiérarchique utilisé pour la régression. Comme nous l'avons vu en section 7.4 (page 155), la présence de ces paramètres dans le modèle augmente la dispersion des estimateurs des paramètres d'intérêt de

la relation d'échelle. Nous cherchons donc à quantifier cette augmentation.

Nous considérons dans un premier temps un biais dans l'estimateur de masse $\alpha_{X|Z}$ (équation 7.21). Il s'agit là d'une considération réaliste, puisque les masses utilisées dans le cadre du grand programme SZ sont des masses hydrostatiques, et que celles-ci sont en moyenne affectés d'un biais (voir section 2.1.3). L'objectif de cette analyse est d'évaluer la perte en précision des estimateurs des paramètres d'intérêt de la relation d'échelle due à la propagation de l'incertitude sur le biais hydrostatique. Nous répétons donc notre analyse en considérant une distribution *a priori* non-informative sur $\alpha_{X|Z}$ correspondant à des valeurs sensées du biais hydrostatique b , soit une distribution uniforme entre -0.4 et 0.4. Les résultats sont également présentés en table 7.3. Les incertitudes sur la pente et la dispersion intrinsèque de la relation d'échelle sont presque inchangées. Cependant, l'incertitude sur l'ordonnée à l'origine est augmentée d'un facteur supérieur à 20. Les données du grand programme SZ n'apportent pas de contraintes sur le biais hydrostatique ; par conséquent, c'est la distribution *a priori* sur ce paramètre qui domine. Il sera donc nécessaire d'adopter un *a priori* informatif pour l'ajustement de la relation d'échelle du grand programme SZ, sans quoi l'un des paramètres d'intérêt ne sera pas contraint.

Enfin, nous répétons l'analyse en considérant un estimateur de masse non-biaisé, mais dispersé, soit $\sigma_{X|Z} \neq 0$. L'objectif est ici d'évaluer la perte de précision due à la propagation de l'incertitude sur la dispersion de l'estimateur de masse hydrostatique. Les résultats sont présentés en table 7.3. Les incertitudes sur les paramètres d'intérêt sont augmentées, de $\sim 10\%$ à $\sim 15\%$ pour l'ordonnée à l'origine et la pente, et de 34% à 46% pour la dispersion intrinsèque. L'augmentation est significative, mais les incertitudes obtenues restent toutefois plus faibles que celles obtenues par les travaux précédents sur les données *Planck*. Là encore, un *a priori* informatif sur la dispersion de l'estimateur de masse permettrait de réduire le volume de l'espace des paramètres, et donc les incertitudes sur les paramètres d'intérêt. Une telle connaissance peut être obtenue par l'étude d'amas simulés, en comparant la masse réelle des amas et celle obtenue par combinaison de leurs profils de pression et de densité (voir par exemple [46]).

7.5.4 Extensions du grand programme SZ

Comme nous l'avons vu en section 3.4 (page 63), le grand programme SZ de NIKA2 dispose de 300 heures d'observations avec NIKA2, attribuées à la collaboration pour la construction de l'instrument. La caméra NIKA2 étant aujourd'hui à disposition de la communauté scientifique, il est possible de soumettre des propositions pour observer d'autres amas que ceux présents au sein de l'échantillon du LPSZ. La possibilité d'observer un échantillon d'amas complémentaire au LPSZ est donc envisagée. Dans cette section, nous nous intéressons à l'apport de telles extensions au programme, du point de vue des contraintes sur la relation d'échelle masse-observable. Pour cela, nous répétons l'analyse présentée précédemment pour des échantillons combinant le LPSZ et d'autres amas, en considérant différentes sélections.

Augmentation de la taille d'échantillon

La première extension considérée est une augmentation du nombre d'amas par boîte dans l'échantillon du LPSZ, en considérant la même sélection que pour l'échantillon réel. Nous avons vu que le LPSZ comportait cinq amas par boîte ; nous répétons donc l'analyse pour un échantillon de taille multipliée par deux, trois, puis quatre. L'incertitude relative η sur les paramètres d'intérêt, définie par l'équation (7.33), est calculée et comparée à celle pour l'analyse de base, afin d'évaluer le gain associé

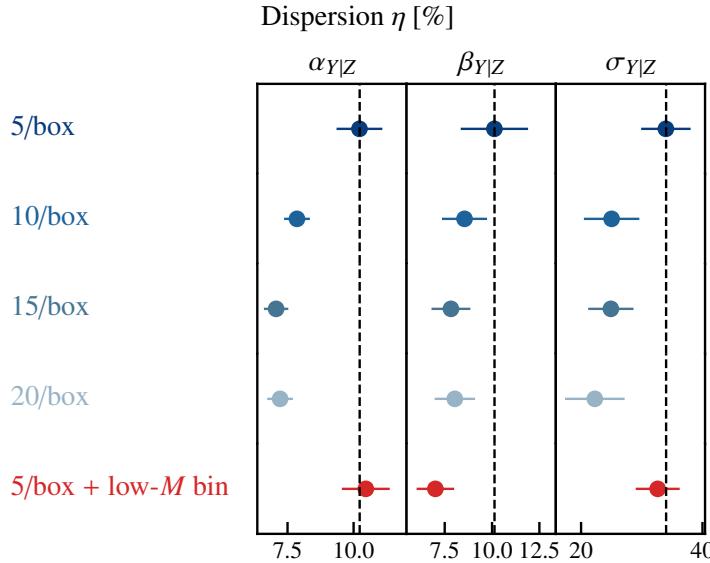


Figure 7.7 : Évolution de l’incertitude relative sur les paramètres d’intérêt avec les extensions du grand programme SZ : ajout d’amas avec la même fonction de sélection (bleu) et ajout d’amas de 10 amas de basse masse (rouge).

à ces augmentations d’échantillon.

Les résultats sont présentés en bleu sur la figure 7.7. On observe une diminution de l’incertitude sur les paramètres avec l’augmentation du nombre d’amas, qui correspond à une diminution de l’erreur statistique. En revanche, on remarque que cette diminution est faible, avec une erreur relative passant de 10% pour le LPSZ nominal (cinq amas par boîtes) à ~7.5% pour une multiplication par quatre de la taille de l’échantillon. Ce gain est très faible en regard du temps d’observation nécessaire à son obtention, puisqu’il exigerait une proposition de temps ouvert de 900 heures d’observation. Nous en déduisons qu’une extension réaliste du programme en considérant la même sélection ne serait pas bénéfique du point de vue de la précision de l’estimation de la relation d’échelle.

Extension à basse masse

S’il n’est pas réaliste d’améliorer les contraintes mises par le LPSZ sur la relation d’échelle en augmentant simplement le nombre d’amas présents dans ce dernier, il reste possible de considérer une sélection différente. Par exemple, il est possible d’observer des amas en dehors de la gamme de masse et de redshift de l’échantillon de base. Une extension en redshift n’apporterait que peu de contraintes supplémentaires mesurables dans cette étude. En effet, nous n’avons pas considéré la mesure de l’évolution avec le redshift de la relation d’échelle. Ainsi, ajouter des amas à bas (ou haut) redshift sur la même gamme de masse que celle couverte par l’échantillon du LPSZ reviendrait simplement à un ajout d’amas dans l’échantillon, et nous avons vu précédemment que le gain en précision serait minime.

En revanche, il est possible d’étendre la gamme de masse couverte. Étant donné le peu d’amas de masse supérieure à $10^{15} M_\odot$ dans l’Univers (voir figure 2.8), une extension à haute masse semble irréaliste. Nous considérons donc une extension à basse masse de l’échantillon du LPSZ, en ajoutant un intervalle en observable SZ. Cette extension est illustrée en figure 7.8. Elle correspond à l’ajout de deux boîtes à faibles valeurs de Y , une pour chaque intervalle en redshift de l’échantillon, comportant cinq amas chacune. L’extension représente alors seulement dix amas, soit bien moins que les 150 nécessaires à multiplier la taille de l’échantillon par quatre.

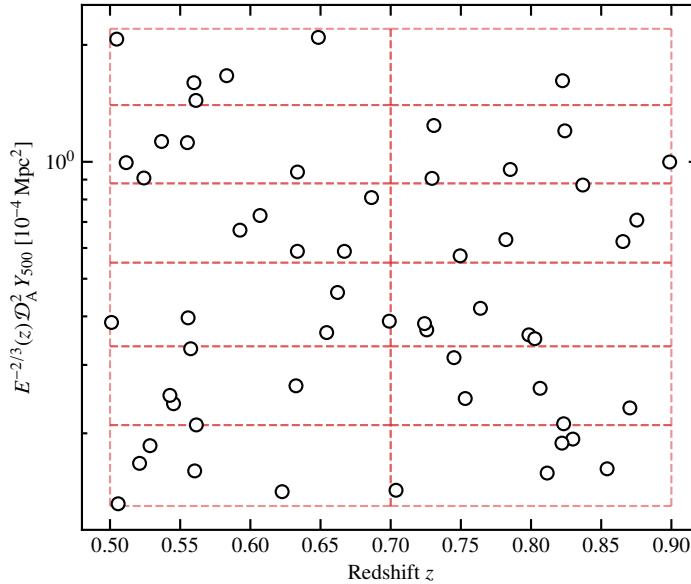


Figure 7.8 : Exemple d'échantillon combinant celui du grand programme SZ (voir figure 3.10) et une extension à basse masse. Deux boîtes (bas redshift et haut redshift) sont ajoutées correspondant à des faibles valeurs d'observable Y .

Les incertitudes relatives sur les paramètres de la relation d'échelle pour 5000 réalisations d'échantillons comportant cette extension sont présentées en rouge sur la figure 7.7. On voit que les incertitudes sur l'ordonnée à l'origine et la dispersion intrinsèque de la relation sont inchangées par rapport à l'analyse de base (ligne noire pointillée). En revanche, l'incertitude sur la pente β est diminuée, atteignant une valeur moyenne de $\sim 7\%$. Ce gain correspond à une couverture étendue en masse, permettant de mieux ajuster la pente de la relation. Le suivi de dix amas de basse masse pourrait donc améliorer les contraintes mises par le LPSZ sur la relation d'échelle, en plus d'études intéressantes sur les propriétés physiques de ces objets [138]. Il est toutefois à noter que si de tels amas existent (par exemple dans le dernier catalogue ACT [2], voir figure 2.8), le temps nécessaire pour observer des amas de basse masse est important. Il sera donc nécessaire de calculer le temps d'observation requis, pour éventuellement construire un échantillon de moins de dix amas.

7.6 Conclusions et perspectives

Dans ce chapitre, nous avons présenté la relation d'échelle liant le paramètre de Compton intégré Y_{500} et la masse M_{500} , son origine physique, et sa modélisation statistique. Nous avons mis en place une analyse basée sur un modèle bayésien hiérarchique, consistant en une combinaison structurelle de plusieurs sous-modèles répondant chacun à un effet systématique pouvant affecter l'ajustement de cette relation. La validation de l'algorithme a montré que ce modèle pouvait prendre en compte ces systématiques afin de livrer des estimations non-biaisées des paramètres d'intérêt de la relation d'échelle. Enfin, l'application à des jeux de données simulés réalistes similaires au grand programme SZ de NIKA2 nous a permis d'identifier un biais dû à la fonction de sélection de ce dernier, analogue au biais de Malmquist. Nous avons vu que ce biais pouvait être significatif dans le cas d'une dispersion intrinsèque importante autour de la relation d'échelle. Enfin, nous avons pu évaluer les incertitudes sur les paramètres d'intérêt de la relation. Si les observations du grand programme SZ permettent d'atteindre la qualité de données attendue, celui-ci permettra un gain en précision sur la connaissance de la relation d'échelle par rapport aux analyses des données *Planck*, avec une prise en compte plus

complète des systématiques affectant l’analyse.

Afin de renforcer ces conclusions, certains axes d’approfondissement peuvent être envisagés pour cette étude. Lors de la génération d’échantillons simulés, nous n’avons pas tenu compte des diverses systématiques associées aux catalogues d’amas d’entrée. En effet, l’échantillon du grand programme SZ a été construit à partir des catalogues *Planck* et ACT d’amas de galaxies, et ceux-ci ne représentent pas une mesure parfaite de l’Univers. Par exemple, plusieurs contaminations peuvent affecter la mesure du paramètre de Compton intégré d’un amas avec *Planck* : des sources ponctuelles non-résolues, une anisotropie primaire du CMB, ou simplement du bruit instrumental. Ces sources d’incertitude n’ont pas été prises en compte au cours de cette étude, et pourraient impacter les contraintes sur la relation d’échelle avec le LPSZ.

D’autre part, comme nous l’avons vu au cours de ce chapitre, un biais de sélection peut affecter l’estimation de la relation d’échelle avec le grand programme SZ. Une piste pour corriger ce biais serait d’étudier sa dépendance avec les paramètres de la relation. Répéter l’étude présentée dans ce chapitre pour différentes valeurs de dispersion intrinsèque permettrait de quantifier l’évolution du biais avec ce paramètre, et donc de pouvoir mettre en place une correction *ad-hoc* pour les résultats finaux du grand programme SZ, en fonction de la valeur de dispersion mesurée lors de l’ajustement des données réelles.

Nous avons vu que la procédure d’ajustement du modèle bayésien hiérarchique permettait de tenir compte d’un grand nombre d’effets systématiques en introduisant des paramètres de nuisance. Cependant, si cette procédure permet d’obtenir une propagation complète des incertitudes sur les résultats finaux, elle implique par construction une augmentation de ces incertitudes. La question des *a priori* sur les paramètres de nuisance est donc centrale. Afin de pouvoir concilier une propagation complète de toutes les incertitudes systématiques et une mesure précise de la relation d’échelle, il sera nécessaire d’imposer des *a priori* informatifs sur les paramètres de nuisance comme le biais et la dispersion de l’estimateur de masse hydrostatique. Ces *a priori* pourront être mesurés en combinant le LPSZ avec des données externes (par exemple en optique avec des mesures de masse dynamique ou par dispersion de vitesses), ou en utilisant des simulations hydrodynamiques.

Enfin, l’étude développée a été utilisée pour commencer à prévoir une extension du grand programme SZ pouvant améliorer le pouvoir de contrainte de ce dernier sur la relation d’échelle. Nous avons vu que multiplier la taille de l’échantillon par deux, trois ou quatre demandait des temps d’observations beaucoup trop importants pour les améliorations atteignables. En revanche, nous avons aussi montré que le suivi avec NIKA2 d’un échantillon d’amas de faible masse permettrait une amélioration conséquente des contraintes sur la pente de la relation avec seulement dix amas. Ce résultat pourra être utilisé pour justifier une demande de temps d’observation avec NIKA2.

Conclusions

Ce manuscrit résume le travail effectué au cours de mes trois années de thèse, s'inscrivant dans le cadre général de la cosmologie observationnelle avec des amas de galaxies. Plus particulièrement, je me suis intéressé à l'utilisation de suivis d'amas pour mieux contraindre les outils nécessaires aux études cosmologiques, et mieux comprendre les systématiques affectant ces outils. Mon travail s'est concentré sur les axes de recherche suivants :

- L'analyse des données brutes de NIKA2 pour construire les cartes de l'effet SZ ;
- L'extraction des propriétés physiques d'un amas distant de faible masse à partir d'observations NIKA2 et *XMM-Newton* ;
- Le développement et la publication d'un logiciel permettant la mesure des propriétés thermodynamiques du milieu intra-amas à partir d'observations SZ ;
- L'étude de la relation d'échelle masse-observable dans le cas du grand programme SZ de NIKA2.

Je me suis tout d'abord intéressé au traitement des données brutes de NIKA2. Comme nous l'avons vu au chapitre 4, la construction de cartes SZ à partir des données en temps n'est pas triviale, du fait de la structure complexe du bruit. En participant aux campagnes d'observation avec NIKA2, j'ai pu acquérir une compréhension des différentes sources de bruit dans les données brutes, de leur origine et de leurs caractéristiques. Cette compréhension m'a permis de maîtriser le processus de décorrélation et de pouvoir traiter des données NIKA2 de manière autonome. J'ai également pu participer à des études sur l'amélioration de ce processus, dont les résultats sont à l'heure actuelle très préliminaires, et qui ne sont pas présentés dans ce manuscrit. Enfin, j'ai développé le programme permettant d'analyser les données NIKA2 des amas de galaxies du grand programme SZ. Ce programme utilise le *pipeline* de la collaboration NIKA2 pour réaliser la décorrélation d'observations d'amas, et permet de calculer le spectre de puissance du bruit résiduel et la fonction de transfert. Ces deux grandeurs quantifient respectivement la structure et l'amplitude du bruit n'ayant pas été retiré par la décorrélation, et le filtrage subi par le signal au cours de cette procédure. Elles servent donc à la fois de critères pour juger la qualité de la décorrélation, mais aussi d'outils permettant de propager les biais et systématiques lors de l'extraction des propriétés physiques d'amas à partir des cartes de l'effet SZ. Ce programme est disponible pour la collaboration NIKA2, et est aujourd'hui systématiquement utilisé pour la réduction des données associées aux amas de galaxies du grand programme SZ.

J'ai ensuite réalisé l'étude des propriétés physiques de l'un des amas de galaxies du grand programme SZ, constituant la deuxième étude d'un amas individuel du programme, et présentée dans le chapitre 6. J'ai choisi pour cette étude de combiner plusieurs difficultés, afin d'évaluer la qualité des contraintes pouvant être apportées sur la thermodynamique d'un amas à partir d'observations NIKA2 dans le pire des cas. D'une part, la qualité des données NIKA2 de cet amas est représentative de

celle attendue pour les amas du grand programme SZ, et inférieure aux premières observations d'amas avec NIKA2. D'autre part, cet amas fait partie des plus basses masses et des plus hauts redshifts de l'échantillon, en faisant une source compacte et faible. Enfin, le faible signal SZ de l'amas est fortement contaminé par la présence de galaxies submillimétriques dans le champ. Après avoir identifié les sources à partir de données du satellite *Herschel*, j'ai développé le logiciel PSTools, permettant de calculer la distribution de probabilité du flux d'une source submillimétrique dans la bande à 150 GHz de NIKA2. Ce calcul repose sur un ajustement par MCMC du spectre de la source à partir de son flux mesuré par *Herschel* et par NIKA2 à 260 GHz. Le logiciel PSTools est aujourd'hui disponible pour la collaboration NIKA2, et utilisé pour l'analyse des sources ponctuelles contaminant les observations d'amas du grand programme SZ. J'ai ensuite développé une extension du logiciel PANCO, développé par F. Ruppin [162], afin de tenir compte de la présence de sources ponctuelles dans l'ajustement des propriétés thermodynamiques d'un amas. Cette approche permet d'ajuster jointement les propriétés thermodynamiques de l'amas et les sources ponctuelles, considérées comme des paramètres de nuisance, et donc de propager l'incertitude sur la contamination par ces sources aux résultats finaux. J'ai donc pu obtenir une estimation du profil de pression de l'amas et des incertitudes associées. En combinant les observations SZ et X, j'ai pu caractériser plus en détail les propriétés thermodynamiques de l'amas. Celui-ci se trouve dans un état dynamique perturbé, avec un cœur chaud, et possède une masse compatible avec celle estimée grâce à son signal SZ intégré et une relation d'échelle masse-observable. Les contraintes mises sur les propriétés thermodynamiques représentent une amélioration en précision par rapport aux contraintes antérieures, illustrant la puissance de la combinaison d'observations X et SZ à haute résolution, même dans le cas de cette source complexe. Cette étude est publiée dans Kéruzoré *et al.*, A&A 644, A93, 2020 [138].

Alors que les observations NIKA2 des amas du grand programme SZ progressent, il est devenu capital de pouvoir disposer d'un moyen d'analyser les données NIKA2 rapidement, afin de pouvoir analyser des échantillons d'amas dans un temps limité. Au début de ma thèse, l'ajustement du profil de pression, réalisé par PANCO, constituait le goulot d'étranglement de l'analyse complète des données NIKA2 d'un amas, des données brutes aux propriétés thermodynamiques. J'ai donc développé le logiciel PANCO2, dont le but premier était de fournir à la collaboration un logiciel optimisé pour mesurer le profil de pression d'amas de galaxies à partir de cartes SZ. Ce logiciel, présenté au chapitre 5, permet de tenir compte des différents biais et systématiques caractéristiques des observations SZ, comme le filtrage du signal, la présence bruit corrélé résiduel, ou la contamination par des sources ponctuelles. PANCO2 propose deux modélisations du profil de pression : le modèle gNFW, paramétrique et largement utilisé dans la littérature, et un modèle non-paramétrique. Ces modèles sont ajustés sur les données en utilisant un échantillonnage MCMC, qui permet ensuite de calculer des estimateurs et intervalles de confiance pour les propriétés physiques de l'amas. J'ai réalisé la validation de PANCO2 sur des simulations, aussi bien analytiques qu'hydrodynamiques, et ai pu montrer que les contraintes apportées sur le profil de pression n'étaient pas biaisées. Au vu des performances de PANCO2, une généralisation du code est en cours pour pouvoir l'utiliser avec des données provenant d'instruments autres que NIKA2. Cette généralisation sera rendue publique, et accompagnée d'une publication présentant l'algorithme de PANCO2 et sa validation.

Enfin, j'ai pu travailler sur la relation d'échelle masse-observable dans le cadre du grand programme SZ. L'ajustement d'une telle relation est complexe, du fait du grand nombre de systématiques intervenants dans l'analyse. J'ai proposé d'utiliser un modèle bayésien hiérarchique pour l'ajustement, en utilisant la librairie LIRA [185]. Cette approche permet de tenir compte des effets systématiques, comme les incertitudes corrélées sur les mesures de masse et d'observable, les effets de sélection, ou encore le biais et la dispersion de l'estimateur de masse utilisé. J'ai évalué la capacité de LIRA à te-

nir compte de ces sources d'erreur en utilisant des jeux de données simulés. J'ai ensuite procédé à la validation de l'utilisation de cette méthode pour le grand programme SZ en mettant en place une procédure Monte Carlo de simulation d'échantillons réalistes du programme. Pour cela, un échantillon est généré suivant une fonction de masse, qui décrit la distribution théorique des amas dans l'Univers, puis une relation d'échelle fiduciale lui est appliquée. Un sous-échantillon d'amas est alors sélectionné de la même façon que le vrai échantillon du grand programme SZ, et utilisé pour ajuster la relation d'échelle et comparer le résultat à la relation fiduciale utilisée. Cette étude a permis de dresser plusieurs conclusions. Dans un premier temps, j'ai pu mettre en évidence un biais sur l'ordonnée à l'origine de la relation d'échelle dû à la fonction de sélection de l'échantillon, singulier par rapport aux biais de sélection habituellement rencontrés. Ce biais pourrait être négligeable ou significatif selon la valeur de la dispersion intrinsèque de la relation. Dans un second temps, j'ai pu estimer les incertitudes attendues sur les paramètres de la relation d'échelle pour le grand programme SZ. Ces incertitudes sont complètes, et tiennent compte à la fois de la statistique et des systématiques intervenant dans l'analyse. Elles montrent que le LPSZ apportera une amélioration à la connaissance de la relation par rapport aux mesures précédentes. Enfin, j'ai utilisé la même méthodologie pour évaluer l'apport d'une extension du grand programme SZ. J'ai montré qu'une augmentation du nombre d'amas dans le programme n'améliorerait que peu les contraintes sur la relation d'échelle. Toutefois, j'ai aussi pu montrer qu'il était possible d'améliorer les contraintes en étendant l'échantillon du programme aux amas de basse masse. Une telle extension pourra être proposée comme un programme en temps ouvert avec NIKA2 dans le futur.

Perspectives

Cette section présente une vision personnelle non exhaustive des perspectives en cosmologie observationnelle, et plus particulièrement avec des amas de galaxies.

La cosmologie observationnelle moderne a pour objectif une mesure toujours plus précise (avec des contraintes fortes) et juste (non-biaisée) des paramètres cosmologiques, afin de répondre aux questions fondamentales sur l’Univers toujours en suspens. Avec l’augmentation du nombre et de la qualité des observations, les incertitudes statistiques sur ces paramètres ont chuté, jusqu’à révéler des différences entre les résultats obtenus avec différentes sondes cosmologiques. L’une des motivations ayant mené à la création du grand programme SZ de NIKA2 était la tension apparente entre les contraintes cosmologiques imposées par le comptage d’amas et le spectre de puissance des anisotropies primaires du CMB, notamment sur la combinaison de paramètres S_8 (voir figure 2.9). Cette tension s’est amoindrie avec la réanalyse du CMB et une évolution dans la valeur de la profondeur optique à la réionisation τ (voir [51], et [13], section 3.6). Il demeure néanmoins des désaccords, par exemple sur le taux d’expansion actuel de l’Univers H_0 mesuré à partir du CMB et de sondes plus récentes, soulevant un grand nombre de questions sur de potentielles systématiques non prises en compte dans les analyses [40, 41]. Il est donc primordial d’étudier ces systématiques afin de pouvoir quantifier la significativité des tensions.

Suivis dédiés d’échantillons d’amas

L’étude des systématiques liées aux propriétés physiques des amas requiert une caractérisation fine de ces dernières, qui peut être apportée par des suivis dédiés de petits échantillons d’amas. Le grand programme SZ de NIKA2 permettra très prochainement d’obtenir une connaissance précise des propriétés thermodynamiques du milieu intra-amas, en combinant des observations SZ à haute résolution et X. L’étude des propriétés de ces amas permettra de comparer les outils nécessaires aux analyses cosmologiques – le profil de pression moyen des amas et la relation d’échelle $Y_{500} - M_{500}$ – avec les mesures actuelles, à plus bas redshift et utilisant des observations X seulement ou SZ à basse résolution. Outre les implications cosmologiques directes liées aux possibles changements dans ces outils, comme l’évolution de la relation d’échelle avec le redshift, la caractérisation des propriétés thermodynamiques permettra d’étudier le lien entre la physique des amas et les outils cosmologiques. Par exemple, la combinaison de la densité du milieu intra-amas, mesurée en X, et de sa pression mesurée en SZ permet de calculer l’entropie du gaz, qui renferme une grande quantité d’information sur l’histoire thermique de la formation de l’amas. L’étude du lien entre l’entropie d’un amas et sa position sur la relation d’échelle masse-observable pourrait donc permettre de comprendre la dispersion intrinsèque de cette relation, et d’affiner l’étalonnage des masses d’amas.

Nous avons également évoqué la possibilité d’étendre le LPSZ avec l’observation d’amas de plus

faible masse, celui-ci étant composé d'amas de masses supérieures à $3.5 \times 10^{14} M_{\odot}$. Nous avons vu au chapitre 7 qu'une telle extension pourrait améliorer les contraintes sur la relation d'échelle. En plus de cette amélioration, les amas de basse masse sont des objets intéressants dont l'étude peut renseigner sur la physique de la formation des structures, et donc sur la cosmologie. D'une part, ces amas sont bien plus nombreux que les amas massifs, la fonction de masse étant très piquée vers les faibles masses. L'augmentation du nombre d'amas détectés dans un relevé repose donc sur la détection de ces objets ; le Simons Observatory prévoit de détecter en SZ des objets jusqu'à $\sim 10^{14} M_{\odot}$. Les amas de faible masse seront donc un élément incontournable de la cosmologie dans les années à venir, et l'impact scientifique du LPSZ pourrait être renforcé en offrant des contraintes sur les propriétés de ces objets. D'autre part, les processus brisant l'équilibre thermique des amas – fusions, injection d'énergie non-thermique par des noyaux actifs de galaxies ou explosions de supernovas – sont plus visibles dans les potentiels gravitationnels plus faibles des amas de faible masse. L'étude des propriétés thermodynamiques de ces objets pourrait donc livrer des informations sur ces processus physiques, et amener à des études sur leur impact sur la cosmologie.

Enfin, dans le futur proche, les progrès instrumentaux permettront de réaliser des suivis d'amas qui affineront notre compréhension des propriétés du milieu intra-amas. En SZ, des instruments seront capables de cartographier l'effet SZ avec une meilleure résolution angulaire et une plus grande *mapping speed* que NIKA2, comme les caméras MISTRAL au Sardinia Radio Telescope [203] ou TolTEC au Large Millimeter Telescope [204]. Ces instruments ouvriront la possibilité de réaliser des suivis d'échantillons plus nombreux et couvrant des domaines plus étendus en masse et en redshift, afin d'étudier avec une grande précision l'évolution des propriétés thermodynamiques du milieu intra-amas avec le redshift. De même, l'instrument CONCERTO [88] installé au télescope APEX permet de réaliser un suivi spectroscopique dans le domaine millimétrique. Les observations d'amas en cours permettront de séparer les différentes composantes de l'effet SZ – thermique, cinétique, et relativiste – à partir d'un seul instrument, ce qui offrira une nouvelle possibilité de mesure des propriétés thermodynamiques du milieu intra-amas indépendante des observations X. Enfin, toutes ces observations pourront être combinées et comparées à celles réalisées par les observatoires X à venir, comme *Athena* [72, 205] et *Lynx* [206], pour améliorer la compréhension de l'état thermodynamiques des amas et des possibles différences entre les observations X et SZ.

Simulations numériques

L'étude des propriétés physiques des amas à partir d'observations est malheureusement limitée par la qualité de ces observations et par les hypothèses utilisées dans les analyses. Par exemple, la comparaison de différents estimateurs de masse d'un amas est affectée par le fait qu'aucun estimateur n'est parfait. Les simulations numériques d'amas offrent alors un moyen de contourner ces défauts et de statuer sur la qualité des hypothèses et des analyses. D'une part, l'utilisation des propriétés réelles des amas simulés permet d'évaluer la qualité des estimateurs de ces propriétés. Dans l'exemple de la masse des amas, comparer la masse réelle des amas à la masse hydrostatique obtenue à partir des propriétés thermodynamiques du milieu intra-amas permet de mesurer le biais hydrostatique réel, là où une comparaison entre la masse hydrostatique d'un amas et de sa masse obtenue par lentillage est affectée par les biais et systématiques des deux estimateurs. De plus, la création d'observations simulées d'amas synthétiques permet d'évaluer la qualité des analyses par comparaison des propriétés observées des amas et de leurs propriétés réelles. Ainsi, en mesurant la masse hydrostatique d'un amas à partir des cartes SZ et X simulées, on peut comparer les résultats à la masse hydrostatique réelle de l'amas, ce qui permet d'identifier les biais et systématiques dans l'analyse des données. On peut égale-

lement les comparer à la masse réelle de l'amas, ce qui permet d'évaluer les systématiques associées aux hypothèses ; par exemple, le biais $\alpha_{X|Z}$ et la dispersion $\sigma_{X|Z}$ de l'estimateur de masse hydrostatique dans l'équation (7.21). L'évolution des simulations numériques permet aujourd'hui d'atteindre un grand niveau de complexité physique, avec des observations simulées de tout le ciel [207, 208] ou de régions abritant des amas de galaxies par *zoom-in*, comme la simulation Three Hundred [209]. Cette dernière comporte plus de 300 amas simulés avec un grand niveau de précision, qui peuvent être utilisés pour générer des observations simulées avec plusieurs types d'instruments. Ces données permettront de quantifier les sources d'erreur présentes dans les analyses cosmologiques basées sur des amas, augmentant ainsi la qualité des contraintes cosmologiques.

Notons également que la validité des résultats de ces études est limitée par le réalisme des simulations. La réalisation de simulations hydrodynamiques d'amas de galaxies repose sur plusieurs hypothèses, en particulier pour les processus physiques se déroulant à des échelles plus petites que leur résolution (rétroaction des AGN, formation stellaire, etc). Ces hypothèses sont en général justifiées par les observations : lorsqu'une simulation est réalisée, ses résultats sont comparés aux amas réellement observés dans l'Univers. L'origine des éventuelles différences est ensuite interprétée pour raffiner les hypothèses de base de la simulation. Grâce aux suivis dédiés à grande résolution comme le grand programme SZ de NIKA2 et aux autres études à venir, détaillés au paragraphe précédent, notre connaissance des propriétés physiques des amas sera améliorée, ce qui permettra d'offrir un retour aux simulations afin de les rendre de plus en plus réalistes.

Cosmologie multilongueur d'onde avec des amas

Le caractère multilongueur d'onde des amas de galaxies est un grand atout pour l'étude des systématiques affectant la cosmologie. En effet, les observations d'amas dans différentes longueurs d'onde sont sensibles à des propriétés physiques différentes, et sont donc associées à des sources d'incertitudes différentes. Combiner ces observations permettra donc de tirer parti de ces différences afin de caractériser les diverses systématiques propres à chaque domaine de longueurs d'onde.

Un exemple de bénéfice apporté par l'étude jointe de relevés d'amas est l'étalonnage des masses de ces objets. Comme nous l'avons déjà évoqué dans cette thèse, cet étalonnage représente l'une des sources majeures d'incertitude dans les analyses cosmologiques basées sur des catalogues d'amas, et ce dans toutes les longueurs d'onde d'intérêt. Les systématiques affectant les mesures de masse dans chaque longueur d'onde sont toutefois très différentes : biais hydrostatique pour les observations SZ et X, biais et dispersion dus à la modélisation des amas pour le lentillage de galaxies d'arrière-plan [62, 63] et pour le lentillage du fond diffus cosmologique [49]. Il est alors possible d'exploiter la différence entre ces systématiques pour mieux les comprendre, en combinant des relevés millimétriques (Simons Observatory, CMB-S4), X (eROSITA) et visibles (LSST, *Euclid*) [210, 211]. La comparaison des masses d'amas estimées à partir de ces relevés permettra d'étudier le biais et la dispersion de chaque estimateur, afin d'améliorer l'utilisation des amas de galaxies comme sondes cosmologiques.

Les amas à l'ère de la combinaison de sondes

La combinaison des contraintes sur les paramètres cosmologiques obtenues à l'aide de plusieurs sondes permet différents types d'étude. D'une part, les différentes dégénérescences propres à chaque sonde font de la combinaison de sondes un moyen d'obtenir des contraintes plus précises sur les paramètres cosmologiques. De plus, la comparaison des paramètres obtenus grâce à des sondes différentes

permet de mettre en évidence d'éventuelles tensions entre les sondes.

Les amas de galaxies étant par essence des objets multilongueur d'onde, ils peuvent être détectés dans la grande majorité des relevés cosmologiques. Si ces objets sont rarement présentés comme l'objectif premier d'un relevé, ils constituent donc néanmoins une opportunité de combinaison de sondes pour chaque expérience visant à contraindre les paramètres cosmologiques. Par exemple, l'analyse des données *Planck* en considérant conjointement les amas de galaxies et les anisotropies primaires du CMB permet de préciser les contraintes sur S_8 , avec une diminution de l'incertitude d'environ 30% par rapport au CMB seul [51]. Les expériences millimétriques à venir, comme le Simons Observatory ou CMB-S4, pourront détecter un grand nombre d'amas par effet SZ, et l'analyse jointe de ces catalogues d'amas et des anisotropies primaires du CMB offrira des contraintes d'une grande qualité sur les paramètres cosmologiques. En optique, les relevés réalisés par le Vera Rubin Observatory et par la mission *Euclid* permettront également la création de grands catalogues d'amas, et les contraintes cosmologiques résultantes pourront être combinées avec les analyses de la distribution à grande échelle des galaxies, du diagramme de Hubble des supernovas de type Ia, etc. Les amas de galaxies s'imposent donc comme une sonde permettant de maximiser l'impact scientifique des expériences à venir.

Bibliographie

- [1] V. Springel *et al.* «Simulations of the Formation, Evolution and Clustering of Galaxies and Quasars», *Nature* 435 (2005), p. 629-636. [Lien](#) (cf. p. 2).
- [2] M. Hilton *et al.* «The Atacama Cosmology Telescope : A Catalog of >4000 Sunyaev–Zel'dovich Galaxy Clusters», *ApJS* 253 (2021), p. 3. [Lien](#) (cf. p. 7, 42, 44, 45, 51, 52, 68, 166).
- [3] L. E. Bleem *et al.* «Galaxy Clusters Discovered via the Sunyaev-Zeldovich Effect in the 2500-Square-Degree SPT-SZ Survey», *ApJS* 216.2 (2015), p. 27. [Lien](#) (cf. p. 7, 41, 45, 51, 52).
- [4] L. E. Bleem *et al.* «The SPTpol Extended Cluster Survey», *ApJS* 247.1 (2020), p. 25. [Lien](#) (cf. p. 7, 39, 44, 45, 51, 52).
- [5] Planck Collaboration *et al.* «Planck 2015 Results - XXVII. The Second Planck Catalogue of Sunyaev-Zeldovich Sources», *A&A* 594 (2016), A27. [Lien](#) (cf. p. 7, 66-68, 72, 106).
- [6] P. Ade *et al.* «The Simons Observatory : Science Goals and Forecasts», *JCAP* 2019.02 (2019), p. 056-056. [Lien](#) (cf. p. 7, 46).
- [7] K. N. Abazajian *et al.* «CMB-S4 Science Book, First Edition», *arXiv* (2016), p. 1610.02743. [Lien](#) (cf. p. 7, 46).
- [8] LSST Dark Energy Science Collaboration. «Large Synoptic Survey Telescope : Dark Energy Science Collaboration», *arXiv* 1211 (2012), p. 1211.0310. [Lien](#) (cf. p. 7, 31).
- [9] B. Sartoris *et al.* «Next Generation Cosmology : Constraints from the Euclid Galaxy Cluster Survey», *MNRAS* 459 (2016), p. 1764-1780. [Lien](#) (cf. p. 7, 31, 46).
- [10] A. Pillepich *et al.* «Forecasts on Dark Energy from the X-Ray Cluster Survey with eROSITA : Constraints from Counts and Clustering», *MNRAS* 481.1 (2018), p. 613-626. [Lien](#) (cf. p. 7).
- [11] L. Perotto *et al.* «Calibration and Performance of the NIKA2 Camera at the IRAM 30-m Telescope», *A&A* 637 (2020), A71. [Lien](#) (cf. p. 8, 47, 52, 59-62, 84, 95, 97, 126, 133).
- [12] A. Einstein. «Die Feldgleichungen Der Gravitation», *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, Seite 844-847. (1915). [Lien](#) (cf. p. 12).
- [13] Planck Collaboration *et al.* «Planck 2018 Results. VI. Cosmological Parameters», *A&A* 641 (2020), A6. [Lien](#) (cf. p. 13, 21, 22, 29, 46, 159, 173).
- [14] G. Lemaître. «Un Univers Homogène de Masse Constante et de Rayon Croissant Rendant Compte de La Vitesse Radiale Des Nébuleuses Extra-Galactiques», *Annales de la Société Scientifique de Bruxelles*, A47, p. 49-59 47 (1927), p. 49-59. [Lien](#) (cf. p. 14).
- [15] E. Hubble. «A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae», *Proceedings of the National Academy of Sciences of the United States of America, Volume 15, Issue 3, pp. 168-173* 15 (1929), p. 168-173. [Lien](#) (cf. p. 14).

- [16] A. Einstein et W. de Sitter. «On the Relation between the Expansion and the Mean Density of the Universe», *Proceedings of the National Academy of Sciences of the United States of America, Volume 18, Issue 3, 1932, pp.213-214* 18 (1932), p. 213-214. [Lien](#) (cf. p. 14).
- [17] A. G. Riess *et al.* «Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant», *The Astronomical Journal, Volume 116, Issue 3, pp. 1009-1038.* 116 (1998), p. 1009-1038. [Lien](#) (cf. p. 14).
- [18] S. Perlmutter *et al.* «Measurements of Ω and Λ from 42 High-Redshift Supernovae», *ApJ, Volume 517, Issue 2, pp. 565-586.* 517 (1999), p. 565-586. [Lien](#) (cf. p. 14).
- [19] M. Chevallier et D. Polarski. «Accelerating Universes with Scaling Dark Matter», *International Journal of Modern Physics D* 10 (2001), p. 213-223. [Lien](#) (cf. p. 14).
- [20] E. V. Linder. «Exploring the Expansion History of the Universe», *Phys. Rev. Letters* 90 (2003), p. 091301. [Lien](#) (cf. p. 14).
- [21] T. Clifton *et al.* «Modified Gravity and Cosmology», *Physics Reports* 513 (2012), p. 1-189. [Lien](#) (cf. p. 14).
- [22] A. Tamosiunas. «Testing and Emulating Modified Gravity on Cosmological Scales», *arXiv* 2011 (2020), p. 2011.08786. [Lien](#) (cf. p. 14).
- [23] J. Lesgourgues et S. Pastor. «Massive Neutrinos and Cosmology», *Physics Reports* 429 (2006), p. 307-379. [Lien](#) (cf. p. 15).
- [24] R. H. Cyburt *et al.* «Big Bang Nucleosynthesis : Present Status», *Reviews of Modern Physics* 88 (2016), p. 015004. [Lien](#) (cf. p. 15).
- [25] D. J. Fixsen *et al.* «The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set», *ApJ* 473 (1996), p. 576. [Lien](#) (cf. p. 15).
- [26] L. F. Abbott, E. Farhi et M. B. Wise. «Particle Production in the New Inflationary Cosmology», *Physics Letters B* 117 (1982), p. 29-33. [Lien](#) (cf. p. 18).
- [27] A. Albrecht *et al.* «Reheating an Inflationary Universe», *Phys. Rev. Letters* 48 (1982), p. 1437-1440. [Lien](#) (cf. p. 18).
- [28] D. Baumann et L. McAllister. «Inflation and String Theory», *Issues in Science and Technology* (2015). [Lien](#) (cf. p. 19).
- [29] E. R. Harrison. «Fluctuations at the Threshold of Classical Cosmology», *Phys. Rev. D* 1 (1970), p. 2726-2730. [Lien](#) (cf. p. 21).
- [30] Y. B. Zeldovich. «A Hypothesis, Unifying the Structure and the Entropy of the Universe», *MNRAS* 160 (1972), 1P. [Lien](#) (cf. p. 21).
- [31] W. H. Press et P. Schechter. «Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation», *ApJ* 187 (1974), p. 425-438. [Lien](#) (cf. p. 22).
- [32] J. Tinker *et al.* «Toward a Halo Mass Function for Precision Cosmology : The Limits of Universality», *ApJ* 688.2 (2008), p. 709. [Lien](#) (cf. p. 22, 23, 159).
- [33] S. Bocquet *et al.* «Halo Mass Function : Baryon Impact, Fitting Formulae, and Implications for Cluster Cosmology», *MNRAS* 456.3 (2016), p. 2361-2373. [Lien](#) (cf. p. 22).
- [34] S. Bocquet *et al.* «The Mira-Titan Universe. III. Emulation of the Halo Mass Function», *ApJ* 901 (2020), p. 5. [Lien](#) (cf. p. 22).

- [35] I. Harrison et S. Hotchkiss. «A Consistent Approach to Falsifying Λ CDM with Rare Galaxy Clusters», *JCAP* 07 (2013), p. 022. [Lien](#) (cf. p. 23).
- [36] L. Salvati, M. Douspis et N. Aghanim. «Impact of Systematics on Cosmological Parameters from Future Galaxy Cluster Surveys», *A&A* 643 (2020), A20. [Lien](#) (cf. p. 23, 46).
- [37] E. Artis *et al.* «Impact of the Calibration of the Halo Mass Function on Galaxy Cluster Number Count Cosmology», *A&A* 649 (2021), A47. [Lien](#) (cf. p. 23, 46).
- [38] A. G. Riess *et al.* «Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond Λ CDM», *ApJ* 876 (2019), p. 85. [Lien](#) (cf. p. 24, 45).
- [39] K. C. Wong *et al.* «H0LiCOW – XIII. A 2.4 per Cent Measurement of H0 from Lensed Quasars : 5.3σ Tension between Early- and Late-Universe Probes», *MNRAS* 498 (2020), p. 1420-1439. [Lien](#) (cf. p. 24, 45).
- [40] G. Efstathiou. «To H0 or Not to H0 ?», *MNRAS* 505 (2021), p. 3866-3872. [Lien](#) (cf. p. 24, 45, 173).
- [41] W. L. Freedman. «Measurements of the Hubble Constant : Tensions in Perspective», *arXiv* (2021), p. 2106.15656. [Lien](#) (cf. p. 24, 173).
- [42] L. Zhou *et al.* «GOODS-ALMA : Optically Dark ALMA Galaxies Shed Light on a Cluster in Formation at $z = 3.5$ », *A&A* 642 (2020), A155. [Lien](#) (cf. p. 26).
- [43] G. W. Pratt *et al.* «The Galaxy Cluster Mass Scale and Its Impact on Cosmological Constraints from the Cluster Population», *Space Science Reviews* 215 (2019), p. 25. [Lien](#) (cf. p. 26, 27, 30, 31, 148).
- [44] L. Salvati *et al.* «Mass Bias Evolution in tSZ Cluster Cosmology», *A&A* 626 (2019), A27. [Lien](#) (cf. p. 28).
- [45] D. J. Barnes *et al.* «Characterizing Hydrostatic Mass Bias with MOCK-X», *MNRAS* (2021). [Lien](#) (cf. p. 28).
- [46] G. Gianfagna *et al.* «Exploring the Hydrostatic Mass Bias in MUSIC Clusters : Application to the NIKA2 Mock Sample», *MNRAS* 502.4 (2021), p. 5115-5133. [Lien](#) (cf. p. 28, 164).
- [47] M. S. Madhavacheril *et al.* «The Atacama Cosmology Telescope : Weighing Distant Clusters with the Most Ancient Light», *ApJL* 903.1 (2020), p. L13. [Lien](#) (cf. p. 28).
- [48] T. Louis et D. Alonso. «Calibrating Cluster Number Counts with CMB Lensing», *Phys. Rev. D* 95.4 (2017), p. 043517. [Lien](#) (cf. p. 28, 46).
- [49] I. Zubeldia et A. Challinor. «Cosmological Constraints from Planck Galaxy Clusters with CMB Lensing Mass Bias Calibration», *MNRAS* 489.1 (2019), p. 401-419. [Lien](#) (cf. p. 28, 175).
- [50] M. Sereno et S. Ettori. «Comparing Masses in Literature (CoMaLit) - I. Bias and Scatter in Weak Lensing and X-Ray Mass Estimates of Clusters», *MNRAS* 450 (2015), p. 3633. [Lien](#) (cf. p. 28, 151, 163).
- [51] L. Salvati, M. Douspis et N. Aghanim. «Constraints from Thermal Sunyaev-Zel'dovich Cluster Counts and Power Spectrum Combined with CMB», *A&A* 614 (2018), A13. [Lien](#) (cf. p. 28, 29, 173, 176).

- [52] Planck Collaboration *et al.* «Planck Intermediate Results - V. Pressure Profiles of Galaxy Clusters from the Sunyaev-Zeldovich Effect», *A&A* 550 (2013), A131. [Lien](#) (cf. p. 29).
- [53] A. Biviano. «From Messier to Abell : 200 Years of Science with Galaxy Clusters», *arXiv* (2000), astro-ph/0010409. [Lien](#) (cf. p. 29).
- [54] E. S. Rykoff *et al.* «redMaPPer. I. Algorithm and SDSS DR8 Catalog», *ApJ* 785.2 (2014), p. 104. [Lien](#) (cf. p. 30, 44).
- [55] R. Adam *et al.* «Euclid Preparation - III. Galaxy Cluster Detection in the Wide Photometric Survey, Performance and Algorithm Selection», *A&A* 627 (2019), A23. [Lien](#) (cf. p. 30).
- [56] M. Costanzi *et al.* «Cosmological Constraints from DES Y1 Cluster Abundances and SPT Multiwavelength Data», *Phys. Rev. D* 103 (2021), p. 043522. [Lien](#) (cf. p. 30).
- [57] A. Phriksee *et al.* «Weak Lensing Analysis of Codex Clusters Using Dark Energy Camera Legacy Survey : Mass–Richness Relation», *MNRAS* 491.2 (2020), p. 1643-1655. [Lien](#) (cf. p. 30).
- [58] M. Postman *et al.* «The Cluster Lensing and Supernova Survey with Hubble : An Overview», *ApJS* 199 (2012), p. 25. [Lien](#) (cf. p. 30, 71).
- [59] K. Umetsu. «Cluster–Galaxy Weak Lensing», *Astron Astrophys Rev* 28.1 (2020), p. 7. [Lien](#) (cf. p. 30).
- [60] M. R. Becker et A. V. Kravtsov. «On the Accuracy of Weak-Lensing Cluster Mass Reconstructions», *ApJ* 740 (2011), p. 25. [Lien](#) (cf. p. 30, 71).
- [61] R. Mandelbaum. «Instrumental Systematics and Weak Gravitational Lensing», *Journal of Instrumentation* 10 (2015), p. C05017. [Lien](#) (cf. p. 30).
- [62] S. Grandis *et al.* «Calibration of Bias and Scatter Involved in Cluster Mass Measurements Using Optical Weak Gravitational Lensing», *arXiv* (2021), p. 2103.16212. [Lien](#) (cf. p. 30, 31, 71, 148, 175).
- [63] M. W. Sommer *et al.* «Weak Lensing Mass Modeling Bias and the Impact of Miscentring», *arXiv* 2105 (2021), p. 2105.08027. [Lien](#) (cf. p. 30, 31, 148, 175).
- [64] E. Munari *et al.* «The Relation between Velocity Dispersion and Mass in Simulated Clusters of Galaxies : Dependence on the Tracer and the Baryonic Physics», *MNRAS* 430 (2013), p. 2638-2649. [Lien](#) (cf. p. 31).
- [65] A. Ferragamo *et al.* «Biases in Galaxy Cluster Velocity Dispersion and Mass Estimates in the Small Ngal Regime», *A&A* 641 (2020), A41. [Lien](#) (cf. p. 31).
- [66] LSST Science Collaboration *et al.* «LSST Science Book, Version 2.0», *arXiv* 0912 (2009), p. 0912.0201. [Lien](#) (cf. p. 31, 46).
- [67] L. Amendola *et al.* «Cosmology and Fundamental Physics with the Euclid Satellite», *Living Reviews in Relativity* 16 (2013), p. 6. [Lien](#) (cf. p. 31).
- [68] H. Böhringer et N. Schartel. «X-Ray Galaxy Cluster Studies for Astrophysics and Cosmology», *Astronomische Nachrichten* 334 (2013), p. 482. [Lien](#) (cf. p. 31).
- [69] H. Böhringer et N. Werner. «X-Ray Spectroscopy of Galaxy Clusters : Studying Astrophysical Processes in the Largest Celestial Laboratories», *Astron Astrophys Rev* 18.1 (2010), p. 127-196. [Lien](#) (cf. p. 31, 68, 71).

- [70] F. Jansen *et al.* «XMM-Newton Observatory. I. The Spacecraft and Operations», *A&A* 365 (2001), p. L1-L6. [Lien](#) (cf. p. 32).
- [71] M. C. Weisskopf *et al.* «An Overview of the Performance and Scientific Results from the Chandra X-Ray Observatory», *Publications of the Astronomical Society of the Pacific* 114 (2002), p. 1-24. [Lien](#) (cf. p. 32).
- [72] K. Nandra *et al.* «The Hot and Energetic Universe : A White Paper Presenting the Science Theme Motivating the Athena+ Mission», *arXiv* 1306 (2013), p. 1306.2307. [Lien](#) (cf. p. 32, 174).
- [73] A. Merloni *et al.* «eROSITA Science Book : Mapping the Structure of the Energetic Universe», *arXiv* 1209 (2012), p. 1209.3114. [Lien](#) (cf. p. 32, 46).
- [74] Y. B. Zeldovich et R. A. Sunyaev. «The Interaction of Matter and Radiation in a Hot-Model Universe», *Astrophys Space Sci* 4.3 (1969), p. 301-316. [Lien](#) (cf. p. 32).
- [75] R. A. Sunyaev et Y. B. Zeldovich. «The Interaction of Matter and Radiation in the Hot Model of the Universe, II», *Astrophysics and Space Science* 7 (1970), p. 20-30. [Lien](#) (cf. p. 32).
- [76] R. A. Sunyaev et Y. B. Zeldovich. «The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies», *Comments Astro Space Phys.* 4 (1972), p. 173. [Lien](#) (cf. p. 32).
- [77] R. A. Sunyaev et Y. B. Zeldovich. «The Velocity of Clusters of Galaxies Relative to the Microwave Background - The Possibility of Its Measurement», *MNRAS* 190 (1980), p. 413-420. [Lien](#) (cf. p. 32, 35).
- [78] M. Birkinshaw. «The Sunyaev-Zel'dovich Effect», *Physics Reports* 310.2 (1999), p. 97-195. [Lien](#) (cf. p. 32).
- [79] J. E. Carlstrom, G. P. Holder et E. D. Reese. «Cosmology with the Sunyaev-Zel'dovich Effect», *Annu. Rev. A&A* 40.1 (2002), p. 643-680. arXiv : [astro-ph/0208192](#). [Lien](#) (cf. p. 32).
- [80] T. Mroczkowski *et al.* «Astrophysics with the Spatially and Spectrally Resolved Sunyaev-Zeldovich Effects. A Millimetre/Submillimetre Probe of the Warm and Hot Universe», *Space Sci. Rev.* 215.1 (2019), p. 17. [Lien](#) (cf. p. 32, 36, 37).
- [81] A. S. Kompaneets. «The Establishment of Thermal Equilibrium between Quanta and Electrons», *Soviet Journal of Experimental and Theoretical Physics* 4 (1957), p. 730-737. [Lien](#) (cf. p. 33).
- [82] G. E. Freire Oliveira, C. Maes et K. Meerts. «On the Derivation of the Kompaneets Equation», *arXiv* 2103 (2021), p. 2103.06654. [Lien](#) (cf. p. 33).
- [83] N. Hand *et al.* «Evidence of Galaxy Cluster Motions with the Kinematic Sunyaev-Zel'dovich Effect», *Phys. Rev. Letters* 109 (2012), p. 041101. [Lien](#) (cf. p. 36).
- [84] T. Mroczkowski *et al.* «A Multi-Wavelength Study of the Sunyaev-Zel'dovich Effect in the Triple-Merger Cluster MACS J0717.5+3745 with MUSTANG and Bolocam», *ApJ* 761 (2012), p. 47. [Lien](#) (cf. p. 36).
- [85] R. Adam *et al.* «Mapping the Kinetic Sunyaev-Zel'dovich Effect toward MACS J0717.5+3745 with NIKA», *A&A* 598 (2017), A115. [Lien](#) (cf. p. 36).
- [86] G. Hurier. «High Significance Detection of the tSZ Effect Relativistic Corrections», *A&A* 596 (2016), A61. [Lien](#) (cf. p. 37).

- [87] M. Remazeilles *et al.* «Can We Neglect Relativistic Temperature Corrections in the Planck Thermal SZ Analysis?», *MNRAS* 483.3 (2019), p. 3459-3464. [Lien](#) (cf. p. 37).
- [88] The CONCERTO collaboration *et al.* «A Wide Field-of-View Low-Resolution Spectrometer at APEX : Instrument Design and Science Forecast», *arXiv* 2007 (2020), p. 2007.14246. [Lien](#) (cf. p. 37, 174).
- [89] J.-B. Melin, J. G. Bartlett et J. Delabrouille. «Catalog Extraction in SZ Cluster Surveys : A Matched Filter Approach», *A&A* 459 (2006), p. 341-352. [Lien](#) (cf. p. 38).
- [90] J.-B. Melin *et al.* «A Comparison of Algorithms for the Construction of SZ Cluster Catalogues», *A&A* 548 (2012), A51. [Lien](#) (cf. p. 38).
- [91] I. Zubeldia *et al.* «Understanding Matched Filters for Precision Cosmology», *arXiv* 2106 (2021), p. 2106.03718. [Lien](#) (cf. p. 38).
- [92] J.-M. Lamarre *et al.* «Planck Pre-Launch Status : The HFI Instrument, from Specification to Actual Performance», *A&A* 520 (2010), A9. [Lien](#) (cf. p. 38).
- [93] Planck Collaboration *et al.* «Planck 2018 Results. I. Overview and the Cosmological Legacy of Planck», *A&A* 641 (2020), A1. [Lien](#) (cf. p. 38).
- [94] M. Arnaud *et al.* «The Universal Galaxy Cluster Pressure Profile from a Representative Sample of Nearby Systems (REXCESS) and the YSZ – M500 Relation», *A&A* 517 (2010), A92. [Lien](#) (cf. p. 39, 40, 44, 66, 68, 71, 119, 126, 138-140, 144, 145).
- [95] S. W. Allen, A. E. Evrard et A. B. Mantz. «Cosmological Parameters from Observations of Galaxy Clusters», *Annu. Rev. A&A* 49.1 (2011), p. 409-470. [Lien](#) (cf. p. 40).
- [96] A. Kozmányan *et al.* «Deriving the Hubble Constant Using Planck and XMM-Newton Observations of Galaxy Clusters», *A&A* 621 (2019), A34. [Lien](#) (cf. p. 40).
- [97] F. Ruppin *et al.* «Impact of the Mean Pressure Profile of Galaxy Clusters on the Cosmological Constraints from the Planck tSZ Power Spectrum», *MNRAS* 490.1 (2019), p. 784-796. [Lien](#) (cf. p. 40, 46, 47, 69).
- [98] A. B. Mantz *et al.* «Cosmology and Astrophysics from Relaxed Galaxy Clusters - II. Cosmological Constraints», *MNRAS* 440 (2014), p. 2077-2098. [Lien](#) (cf. p. 40).
- [99] S. Bocquet *et al.* «Cluster Cosmology Constraints from the 2500 Deg² SPT-SZ Survey : Inclusion of Weak Gravitational Lensing Data from Magellan and the Hubble Space Telescope», *ApJ* 878.1 (2019), p. 55. [Lien](#) (cf. p. 41, 42, 44, 46).
- [100] Planck Collaboration *et al.* «Planck 2015 Results - XXIV. Cosmology from Sunyaev-Zeldovich Cluster Counts», *A&A* 594 (2016), A24. [Lien](#) (cf. p. 42, 46, 47).
- [101] T. de Haan *et al.* «Cosmological Constraints from Galaxy Clusters in the 2500 Square-Degree SPT-SZ Survey», *ApJ* 832.1 (2016), p. 95. [Lien](#) (cf. p. 42).
- [102] F. Pacaud *et al.* «The XXL Survey - XXV. Cosmological Analysis of the C1 Cluster Number Counts», *A&A* 620 (2018), A10. [Lien](#) (cf. p. 42, 46).
- [103] DES Collaboration *et al.* «Dark Energy Survey Year 1 Results : Cosmological Constraints from Cluster Abundances and Weak Lensing», *Phys. Rev. D* 102.2 (2020), p. 023509. [Lien](#) (cf. p. 42, 44-46).
- [104] S. Andreon et J. Bergé. «Richness-Mass Relation Self-Calibration for Galaxy Clusters», *A&A* 547 (2012), A117. [Lien](#) (cf. p. 43).

- [105] S. Andreon et M. Hurn. «Measurement Errors and Scaling Relations in Astrophysics : A Review», *Stat. Anal. Data Min. ASA Data Sci. J.* 6.1 (2013), p. 15-33. [Lien](#) (cf. p. 43, 146, 148, 153).
- [106] M. Costanzi *et al.* «Methods for Cluster Cosmology and Application to the SDSS in Preparation for DES Year 1 Release», *MNRAS* 488.4 (2019), p. 4779-4800. [Lien](#) (cf. p. 44).
- [107] G. F. Lesci *et al.* «AMICO Galaxy Clusters in KiDS-DR3 : Cosmological Constraints from Counts and Stacked Weak-Lensing», *arXiv* 2012 (2020), p. 2012.12273. [Lien](#) (cf. p. 44, 46).
- [108] H. Böhringer *et al.* «The Northern ROSAT All-Sky (NORAS) Galaxy Cluster Survey. I. X-Ray Properties of Clusters Detected as Extended X-Ray Sources», *ApJS* 129 (2000), p. 435-474. [Lien](#) (cf. p. 44).
- [109] H. Böhringer *et al.* «The ROSAT-ESO Flux Limited X-Ray (REFLEX) Galaxy Cluster Survey. V. The Cluster Catalogue», *A&A* 425 (2004), p. 367-383. [Lien](#) (cf. p. 44).
- [110] G. W. Pratt *et al.* «Galaxy Cluster X-Ray Luminosity Scaling Relations from a Representative Local Sample (REXCESS)», *A&A* 498.2 (2009), p. 361-378. [Lien](#) (cf. p. 44).
- [111] G. W. Pratt *et al.* «Gas Entropy in a Representative Sample of Nearby X-Ray Galaxy Clusters (REXCESS) : Relationship to Gas Mass Fraction», *A&A* 511 (2010), A85. [Lien](#) (cf. p. 44, 70, 128, 130).
- [112] C. Adami *et al.* «The XXL Survey. XX. The 365 Cluster Catalogue», *A&A* 620 (2018), A5. [Lien](#) (cf. p. 44).
- [113] Planck Collaboration *et al.* «Planck 2015 Results - XIII. Cosmological Parameters», *A&A* 594 (2016), A13. [Lien](#) (cf. p. 44, 45).
- [114] N. Huang *et al.* «Galaxy Clusters Selected via the Sunyaev–Zel'dovich Effect in the SPTpol 100-Square-Degree Survey», *AJ* 159.3 (2020), p. 110. [Lien](#) (cf. p. 45).
- [115] H. Böhringer, G. Chon et C. A. Collins. «Observational Evidence for a Local Underdensity in the Universe and Its Effect on the Measurement of the Hubble Constant», *A&A* 633 (2020), A19. [Lien](#) (cf. p. 45).
- [116] E. Di Valentino *et al.* «Interacting Dark Energy in the Early 2020s : A Promising Solution to the H₀ and Cosmic Shear Tensions», *Phys. Dark Universe* 30 (2020), p. 100666. [Lien](#) (cf. p. 45).
- [117] B. Bolliet *et al.* «Including Massive Neutrinos in Thermal Sunyaev Zeldovich Power Spectrum and Cluster Counts Analyses», *MNRAS* 497 (2020), p. 1332-1347. [Lien](#) (cf. p. 45).
- [118] M. Cataneo et D. Rapetti. «Tests of Gravity with Galaxy Clusters», *Int. J. Mod. Phys. D* 27.15 (2018), p. 1848006. [Lien](#) (cf. p. 45).
- [119] L. Balkenhol *et al.* «Constraints on Λ CDM Extensions from the SPT-3G 2018 EE and TE Power Spectra», *arXiv* 2103 (2021), p. 2103.13618. [Lien](#) (cf. p. 46).
- [120] S. Aiola *et al.* «The Atacama Cosmology Telescope : DR4 Maps and Cosmological Parameters», *JCAP* 12 (2020), p. 047. [Lien](#) (cf. p. 46).
- [121] T. M. C. Abbott *et al.* «Dark Energy Survey Year 1 Results : Cosmological Constraints from Galaxy Clustering and Weak Lensing», *Phys. Rev. D* 98 (2018), p. 043526. [Lien](#) (cf. p. 46).
- [122] C. Heymans *et al.* «KiDS-1000 Cosmology : Multi-Probe Weak Gravitational Lensing and Spectroscopic Galaxy Clustering Constraints», *A&A* 646 (2021), A140. [Lien](#) (cf. p. 46).

- [123] R. Adam *et al.* «The NIKA2 Large-Field-of-View Millimetre Continuum Camera for the 30 m IRAM Telescope», *A&A* 609 (2018), A115. [Lien](#) (cf. p. 47, 52, 54).
- [124] S. R. Dicker *et al.* «MUSTANG2 : A Large Focal Plan Array for the 100 Meter Green Bank Telescope», *Proc. SPIE* 9153 (2014), 91530J. [Lien](#) (cf. p. 47).
- [125] F. Mayet *et al.* «Cluster Cosmology with the NIKA2 SZ Large Program», *EPJ Web Conf.* 228 (2020), p. 00017. [Lien](#) (cf. p. 47, 63).
- [126] J. W. M. Baars *et al.* «The IRAM 30-m Millimeter Radio Telescope on Pico Veleta, Spain», *A&A* 175 (1987), p. 319-326. [Lien](#) (cf. p. 50, 51).
- [127] J. R. Pardo, J. Cernicharo et E. Serabyn. «Atmospheric Transmission at Microwaves (ATM) : An Improved Model for Millimeter/Submillimeter Applications», *IEEE Transactions on Antennas and Propagation* 49 (2001), p. 1683-1694. [Lien](#) (cf. p. 50, 55).
- [128] J. W. M. Baars *et al.* «Thermal Control of the IRAM 30-m Millimeter Radio Telescope», *A&A* 195 (1988), p. 364-371. [Lien](#) (cf. p. 51).
- [129] P. K. Day *et al.* «A Broadband Superconducting Detector Suitable for Use in Large Arrays», *Nature* 425 (2003), p. 817-821. [Lien](#) (cf. p. 52, 53).
- [130] M. Roesch *et al.* «Development of Lumped Element Kinetic Inductance Detectors for NIKA», *arXiv* 1212 (2012), p. 1212.4585. [Lien](#) (cf. p. 52, 53).
- [131] C. J. Gorter et H. Casimir. «On Supraconductivity I», *Physica* 1 (1934), p. 306-320. [Lien](#) (cf. p. 52).
- [132] J. Bardeen, L. N. Cooper et J. R. Schrieffer. «Microscopic Theory of Superconductivity», *Phys. Rev.* 106 (1957), p. 162-164. [Lien](#) (cf. p. 53).
- [133] J. Bardeen, L. N. Cooper et J. R. Schrieffer. «Theory of Superconductivity», *Phys. Rev.* 108 (1957), p. 1175-1204. [Lien](#) (cf. p. 53).
- [134] Planck Collaboration *et al.* «Planck 2015 Results - XXII. A Map of the Thermal Sunyaev-Zeldovich Effect», *A&A* 594 (2016), A22. [Lien](#) (cf. p. 64, 69).
- [135] P. a. R. Ade *et al.* «Planck Intermediate Results - X. Physics of the Hot Gas in the Coma Cluster», *A&A* 554 (2013), A140. [Lien](#) (cf. p. 64).
- [136] M. Hasselfield *et al.* «The Atacama Cosmology Telescope : Sunyaev-Zel'dovich Selected Galaxy Clusters at 148 GHz from Three Seasons of Data», *JCAP* 2013.07 (2013), p. 008-008. [Lien](#) (cf. p. 66, 67, 106, 126, 134, 138, 139).
- [137] F. Ruppin *et al.* «First Sunyaev-Zel'dovich Mapping with the NIKA2 Camera : Implication of Cluster Substructures for the Pressure Profile and Mass Estimate», *A&A* 615 (2018), A112. [Lien](#) (cf. p. 67, 125).
- [138] F. Kéruzoré *et al.* «Exploiting NIKA2/XMM-Newton Imaging Synergy for Intermediate-Mass High-z Galaxy Clusters within the NIKA2 SZ Large Program - Observations of ACT-CL J0215.4+0030 at $z \sim 0.9$ », *A&A* 644 (2020), A93. [Lien](#) (cf. p. 67, 113, 126, 166, 170).
- [139] M. McDonald *et al.* «The Redshift Evolution of the Mean Temperature, Pressure, and Entropy Profiles in 80 SPT-Selected Galaxy Clusters», *ApJ* 794.1 (2014), p. 67. [Lien](#) (cf. p. 69).
- [140] V. Ghirardini *et al.* «Evolution of the Thermodynamic Properties of Clusters of Galaxies out to Redshift of 1.8», *ApJ* 910 (2021), p. 14. [Lien](#) (cf. p. 69).

- [141] R. Adam *et al.* «Mapping the Hot Gas Temperature in Galaxy Clusters Using X-Ray and Sunyaev-Zel'dovich Imaging», *A&A* 606 (2017), A64. [Lien](#) (cf. p. 69).
- [142] F. Ruppin *et al.* «Unveiling the Merger Dynamics of the Most Massive MaDCoWS Cluster at $z = 1.2$ from a Multiwavelength Mapping of Its Intracluster Medium Properties», *ApJ* 893.1 (2020), p. 74. [Lien](#) (cf. p. 69).
- [143] G. M. Voit *et al.* «Modified Entropy Models for the Intracluster Medium», *ApJ* 576 (2002), p. 601-624. [Lien](#) (cf. p. 69).
- [144] G. M. Voit. «Tracing Cosmic Evolution with Clusters of Galaxies», *Reviews of Modern Physics* 77 (2005), p. 207-258. [Lien](#) (cf. p. 70).
- [145] D. S. Hudson *et al.* «What Is a Cool-Core Cluster? A Detailed Analysis of the Cores of the X-Ray Flux-Limited HIFLUGCS Cluster Sample», *A&A* 513 (2010), A37. [Lien](#) (cf. p. 71, 138).
- [146] R. Barrena *et al.* «Optical Validation and Characterisation of Planck PSZ1 Sources at the Canary Islands Observatories. II. Second Year of ITP13 Observations», *A&A* 638 (2020), A146. [Lien](#) (cf. p. 71).
- [147] M. Muñoz Echevarria *et al.* «The LPSZ-CLASH Galaxy Cluster Sample : Combining Lensing and Hydrostatic Mass Estimates», *Proc. MmUniverse Conf.* in prep. (2021) (cf. p. 71).
- [148] Planck Collaboration *et al.* «Planck Early Results. XI. Calibration of the Local Galaxy Cluster Sunyaev-Zeldovich Scaling Relations», *A&A* 536 (2011), A11. [Lien](#) (cf. p. 71, 144, 145, 159, 162, 163).
- [149] J. Guy *et al.* «SALT2 : Using Distant Supernovae to Improve the Use of Type Ia Supernovae as Distance Indicators», *A&A* 466 (2007), p. 11-21. [Lien](#) (cf. p. 72).
- [150] M. Betoule *et al.* «Improved Cosmological Constraints from a Joint Analysis of the SDSS-II and SNLS Supernova Samples», *A&A* 568 (2014), A22. [Lien](#) (cf. p. 72).
- [151] S. Albrand, J. Fulachier et F. Lambert. «The ATLAS Metadata Interface», *J. Phys. : Conf. Ser.* 219.4 (2010), p. 042030. [Lien](#) (cf. p. 73).
- [152] R. Adam. «Observation des amas de galaxies par effet Sunyaev-Zel'dovich et de la polarisation du fond diffus cosmologique : de Planck à NIKA». Thèse de doctorat. Université Grenoble Alpes, 2015. [Lien](#) (cf. p. 79).
- [153] F. Kéruzoré *et al.* *PSTools : A Toolbox for the Estimation of the Contamination by Sub-Millimeter Galaxies in NIKA2 Galaxy Cluster Observations*. Version 1.0. 2019 (cf. p. 94).
- [154] M. J. Griffin *et al.* «The Herschel-SPIRE Instrument and Its in-Flight Performance», *A&A* 518 (2010), p. L3. [Lien](#) (cf. p. 95).
- [155] A. Poglitsch *et al.* «The Photodetector Array Camera and Spectrometer (PACS) on the Herschel Space Observatory», *A&A* 518 (2010), p. L2. [Lien](#) (cf. p. 95, 129).
- [156] H. Dembinski *et al.* *Scikit-Hep/Iminuit : V1.4.9*. Zenodo, 2020. [Lien](#) (cf. p. 95).
- [157] D. Foreman-Mackey *et al.* «Emcee v3 : A Python Ensemble Sampling Toolkit for Affine-Invariant MCMC», *J. Open Source Softw.* 4 (2019), p. 1864. [Lien](#) (cf. p. 96, 108, 134).
- [158] F.-X. Désert *et al.* «Submillimetre Point Sources from the Archeops Experiment : Very Cold Clumps in the Galactic Plane», *A&A* 481.2 (2008), p. 411-421. [Lien](#) (cf. p. 96).

- [159] B. Magnelli *et al.* «A Herschel View of the Far-Infrared Properties of Submillimetre Galaxies», *A&A* 539 (2012), A155. [Lien](#) (cf. p. 96).
- [160] D. J. B. Smith *et al.* «Isothermal Dust Models of Herschel-ATLAS Galaxies», *MNRAS* 436.3 (2013), p. 2435-2453. [Lien](#) (cf. p. 96).
- [161] M. Muñoz Echevarria *et al.* «Multi-Probe Analysis of the Galaxy Cluster CLJ1227 : Unveiling Systematics in Mass Estimation», *Proc. MmUniverse Conf.* in prep. (2021) (cf. p. 99).
- [162] F. Ruppin. «Cosmologie via les observations d'amas de galaxies par effet Sunyaev-Zel'dovich avec NIKA2». Thèse de doctorat. Université Grenoble Alpes, 2018. [Lien](#) (cf. p. 101, 131, 170).
- [163] H. Zhao. «Analytical Models for Galactic Nuclei», *MNRAS* 278 (1996), p. 488-496. [Lien](#) (cf. p. 103).
- [164] D. Nagai, A. V. Kravtsov et A. Vikhlinin. «Effects of Galaxy Formation on Thermodynamics of the Intracluster Medium», *ApJ* 668.1 (2007), p. 1-14. [Lien](#) (cf. p. 103, 133).
- [165] F. Ruppin *et al.* «Non-Parametric Deprojection of NIKA SZ Observations : Pressure Distribution in the Planck-Discovered Cluster PSZ1 G045.85+57.71», *A&A* 597 (2017), A110. [Lien](#) (cf. p. 104).
- [166] C. Romero *et al.* «A Multi-Instrument Non-Parametric Reconstruction of the Electron Pressure Profile in the Galaxy Cluster CLJ1226.9+3332», *A&A* 612 (2018), A39. [Lien](#) (cf. p. 104).
- [167] N. Metropolis et S. Ulam. «The Monte Carlo Method», *J. Am. Stat. Assoc.* 44.247 (1949), p. 335-341. JSTOR : [2280232](#) (cf. p. 108).
- [168] W. K. Hastings. «Monte Carlo Sampling Methods Using Markov Chains and Their Applications», *Biometrika* 57.1 (1970), p. 97-109. [Lien](#) (cf. p. 108).
- [169] J. Goodman et J. Weare. «Ensemble Samplers with Affine Invariance», *Commun. Appl. Math. Comput. Sci.* 5.1 (2010), p. 65-80. [Lien](#) (cf. p. 108).
- [170] B. Nelson, E. B. Ford et M. J. Payne. «Run DMC : An Efficient, Parallel Code for Analyzing Radial Velocity Observations Using N-Body Integrations and Differential Evolution Markov Chain Monte Carlo», *ApJS* 210.1 (2013), p. 11. [Lien](#) (cf. p. 108).
- [171] A. Gelman et D. B. Rubin. «Inference from Iterative Simulation Using Multiple Sequences», *Statistical Science* 7 (1992), p. 457-472. [Lien](#) (cf. p. 109, 110, 134).
- [172] F. Sembolini *et al.* «The MUSIC of Galaxy Clusters – I. Baryon Properties and Scaling Relations of the Thermal Sunyaev–Zel'dovich Effect», *MNRAS* 429.1 (2013), p. 323-343. [Lien](#) (cf. p. 112, 120).
- [173] A. Klypin *et al.* «Resolving the Structure of Cold Dark Matter Halos», *ApJ* 554.2 (2001), p. 903. [Lien](#) (cf. p. 120).
- [174] S. Gottlöber et G. Yepes. «Shape, Spin, and Baryon Fraction of Clusters in the MareNostrum Universe», *ApJ* 664.1 (2007), p. 117. [Lien](#) (cf. p. 121).
- [175] F. Prada *et al.* «Halo Concentrations in the Standard Λ Cold Dark Matter Cosmology», *MNRAS* 423.4 (2012), p. 3018-3030. [Lien](#) (cf. p. 121).
- [176] F. Ruppin *et al.* «Impact of ICM Disturbances on the Mean Pressure Profile of Galaxy Clusters : A Prospective Study of the NIKA2 SZ Large Program with MUSIC Synthetic Clusters», *A&A* 631 (2019), A21. [Lien](#) (cf. p. 121, 122).

- [177] F. Menanteau *et al.* «The Atacama Cosmology Telescope : Physical Properties of Sunyaev-Zeldovich Effect Clusters on the Celestial Equator», *ApJ* 765.1 (2013), p. 67. [Lien](#) (cf. p. 126).
- [178] I. Bartalucci *et al.* «Resolving Galaxy Cluster Gas Properties at $z \sim 1$ with XMM-Newton and Chandra», *A&A* 598 (2017), A61. [Lien](#) (cf. p. 128).
- [179] M. P. Viero *et al.* «The Herschel Stripe 82 Survey (HerS) : Maps and Early Catalog», *ApJS* 210.2 (2014), p. 22. [Lien](#) (cf. p. 129, 131, 133).
- [180] D. J. Helfand, R. L. White et R. H. Becker. «The Last of FIRST : The Final Catalog and Source Identifications», *ApJ* 801.1 (2015), p. 26. [Lien](#) (cf. p. 130).
- [181] J. J. Condon. «Cosmological Evolution of Radio Sources.», *ApJ* 287 (1984), p. 461. [Lien](#) (cf. p. 131).
- [182] M. Ricci *et al.* «The XXL Survey. XLIV. Sunyaev-Zel'dovich Mapping of a Low-Mass Cluster at $z \sim 1$: A Multi-Wavelength Approach», *A&A* 642 (2020), A126. [Lien](#) (cf. p. 138).
- [183] M. McDonald *et al.* «The Remarkable Similarity of Massive Galaxy Clusters from $Z \sim 0$ to $Z \sim 1.9$ », *ApJ* 843.1 (2017), p. 28. [Lien](#) (cf. p. 138).
- [184] F. Ruppin *et al.* «Stability of Cool Cores During Galaxy Cluster Growth : A Joint Chandra/SPT Analysis of 67 Galaxy Clusters Along a Common Evolutionary Track Spanning 9 Gyr», *arXiv* 2012 (2020), p. 2012.14669. [Lien](#) (cf. p. 138).
- [185] M. Sereno. «A Bayesian Approach to Linear Regression in Astronomy», *MNRAS* 455.2 (2016), p. 2149-2162. [Lien](#) (cf. p. 142, 146, 151-154, 170).
- [186] A. V. Kravtsov, A. Vikhlinin et D. Nagai. «A New Robust Low-Scatter X-Ray Mass Indicator for Clusters of Galaxies», *ApJ* 650.1 (2006), p. 128. [Lien](#) (cf. p. 142, 144).
- [187] N. Kaiser. «Evolution of Clusters of Galaxies», *ApJ* 383 (1991), p. 104-111. [Lien](#) (cf. p. 142).
- [188] M. G. Akritas et M. A. Bershady. «Linear Regression for Astronomical Data with Measurement Errors and Intrinsic Scatter», *ApJ* 470 (1996), p. 706. [Lien](#) (cf. p. 145, 146).
- [189] B. Efron. «Bootstrap Methods : Another Look at the Jackknife», *Ann. Statist.* 7.1 (1979), p. 1-26. [Lien](#) (cf. p. 145).
- [190] K. Andersson *et al.* «X-Ray Properties of the First Sunyaev-Zel'dovich Effect Selected Galaxy Cluster Sample from the South Pole Telescope», *ApJ* 738.1 (2011), p. 48. [Lien](#) (cf. p. 145).
- [191] M. Sereno, S. Ettori et L. Moscardini. «CoMaLit - II. The Scaling Relation between Mass and Sunyaev-Zel'dovich Signal for Planck Selected Galaxy Clusters», *MNRAS* 450 (2015), p. 3649. [Lien](#) (cf. p. 145).
- [192] P. Dellaportas et D. A. Stephens. «Bayesian Analysis of Errors-in-Variables Regression Models», *Biometrics* 51.3 (1995), p. 1085-1095. JSTOR : 2533007 (cf. p. 146).
- [193] G. D'Agostini. «Fits, and Especially Linear Fits, with Errors on Both Axes, Extra Variance of the Data Points and Other Complications», *arXiv* (2005), physics/0511182. [Lien](#) (cf. p. 146).
- [194] E. D. Feigelson et G. J. Babu. *Modern Statistical Methods for Astronomy : With R Applications*. Cambridge, 2012. ISBN : 978-0-521-76727-9. [Lien](#) (cf. p. 146).
- [195] J. M. Hilbe, R. S. de Souza et E. E. O. Ishida. *Bayesian Models for Astrophysical Data : Using R, JAGS, Python, and Stan*. Cambridge, 2017. ISBN : 978-1-107-13308-2. [Lien](#) (cf. p. 146).
- [196] A. S. Eddington. «On a Formula for Correcting Statistics for the Effects of a Known Error of Observation», *MNRAS* 73 (1913), p. 359-360. [Lien](#) (cf. p. 148, 151, 157).

- [197] A. S. Eddington Sir. «The Correction of Statistics for Accidental Error», *MNRAS* 100 (1940), p. 354. [Lien](#) (cf. p. 148, 151, 157).
- [198] K. G. Malmquist. «On Some Relations in Stellar Statistics», *Meddelanden fran Lunds Astronomiska Observatorium Serie I* 100 (1922), p. 1-52. [Lien](#) (cf. p. 149).
- [199] B. C. Kelly. «Some Aspects of Measurement Error in Linear Regression of Astronomical Data», *ApJ* 665.2 (2007), p. 1489. [Lien](#) (cf. p. 151, 154).
- [200] A. B. Mantz. «A Gibbs Sampler for Multivariate Linear Regression», *MNRAS* 457.2 (2016), p. 1279-1288. [Lien](#) (cf. p. 151, 154).
- [201] A. Gelman *et al.* *Bayesian Data Analysis, Third Edition*. 2013. 677 p. ISBN : 978-1-4398-4095-5. Google Books : [ZXL6AQAAQBAJ](#) (cf. p. 154).
- [202] Planck Collaboration *et al.* «Planck 2013 Results. XX. Cosmology from Sunyaev–Zeldovich Cluster Counts», *A&A* 571 (2014), A20. [Lien](#) (cf. p. 159).
- [203] P. Bolli *et al.* «Sardinia Radio Telescope», *J. Astron. Instrum.* 4 (2015), p. 1550008-880. [Lien](#) (cf. p. 174).
- [204] G. W. Wilson *et al.* «The TolTEC Camera : An Overview of the Instrument and in-Lab Testing Results», 11453 (2020), p. 1145302. [Lien](#) (cf. p. 174).
- [205] D. Barret *et al.* «The Athena Space X-Ray Observatory and the Astrophysics of Hot Plasma», *Astron. Nachrichten* 341 (2020), p. 224-235. [Lien](#) (cf. p. 174).
- [206] The Lynx Team. «The Lynx Mission Concept Study Interim Report», *arXiv* (2018), p. 1809.09642. [Lien](#) (cf. p. 174).
- [207] LSST Dark Energy Science Collaboration (LSST DESC) *et al.* «The LSST DESC DC2 Simulated Sky Survey», *ApJS* 253 (2021), p. 31. [Lien](#) (cf. p. 175).
- [208] K. Heitmann *et al.* «The Last Journey. I. An Extreme-Scale Simulation on the Mira Supercomputer», *ApJS* 252 (2021), p. 19. [Lien](#) (cf. p. 175).
- [209] W. Cui *et al.* «The Three Hundred Project : A Large Catalogue of Theoretically Modelled Galaxy Clusters for Cosmological and Astrophysical Applications», *MNRAS* 480.3 (2018), p. 2898-2915. [Lien](#) (cf. p. 175).
- [210] J. Rhodes *et al.* «Scientific Synergy between LSST and Euclid», *ApJS* 233.2 (2017), p. 21. [Lien](#) (cf. p. 175).
- [211] S. Dodelson *et al.* «Cosmic Visions Dark Energy : Science», *arXiv* (2016), p. 1604.07626. [Lien](#) (cf. p. 175).

Résumé

Les amas de galaxies représentent l'aboutissement du processus de formation des grandes structures, et forment les plus grands objets gravitationnellement liés de l'Univers. Leur abondance en masse et en redshift permet de tracer la distribution de matière et son évolution avec le temps, et constitue donc une sonde cosmologique. L'effet Sunyaev-Zeldovich (SZ) permet de détecter les amas de galaxies grâce à leur empreinte sur le fond diffus cosmologique. Des catalogues d'amas ont pu être dressés grâce à l'observation de cet effet, notamment par le satellite *Planck*. Cependant, leur utilisation pour contraindre les paramètres cosmologiques requiert la connaissance de la masse des amas, qui n'est pas une observable directe. Il est donc nécessaire de s'appuyer sur des relations d'échelle liant les observables à la masse. Ces relations sont étalonnées sur des échantillons d'amas dont les propriétés physiques sont caractérisées avec un grand niveau de détail, notamment grâce à des observations à haute résolution. Cette connaissance fine des propriétés des amas permet d'étudier les systématiques intervenant en cosmologie, comme l'impact de l'état de relaxation des amas sur la relation d'échelle.

Cette thèse s'inscrit dans le cadre du grand programme SZ de NIKA2, qui a pour but d'exploiter les observations à haute résolution de l'effet SZ avec la caméra NIKA2 au télescope de 30 mètres de l'IRAM pour étudier les propriétés d'un échantillon d'amas de galaxies distants. Le travail effectué utilise des techniques statistiques avancées pour étudier les systématiques devant être prise en compte en cosmologie avec des amas. Une partie de ce travail porte sur l'analyse des données de NIKA2 afin de caractériser les propriétés thermodynamiques des amas en tenant compte des systématiques associées aux observations SZ, aboutissant au développement d'un logiciel public permettant de contraindre les propriétés thermodynamiques des amas de galaxies à partir d'observations SZ. Les méthodes développées sont ensuite utilisées pour caractériser les propriétés de l'un des amas les plus complexes du programme, en exploitant notamment la synergie entre les observations NIKA2 et *XMM-Newton* de l'amas de résolutions comparables. Par la suite, nous présentons une modélisation statistique rigoureuse de la relation masse-observable dans le cas des observations SZ, en utilisant un modèle bayésien hiérarchique permettant de tenir compte des systématiques affectant les mesures de masse et d'observable. Nous appliquons cette analyse sur des simulations réalistes, permettant de mettre en évidence les biais et effets systématiques impliqués dans cette analyse, et de prédire les contraintes pouvant être apportées par les observations avec NIKA2 sur la relation masse-observable.

Abstract

Galaxy clusters are the culmination of large-scale structure formation, and represent the largest gravitationally bound objects in the Universe. Their abundance in mass and redshift is therefore a tracer of matter distribution across space and time, and thus constitutes a cosmological probe. The Sunyaev-Zeldovich (SZ) effect enables the detection of clusters through their imprint on the cosmic microwave background. Cluster catalogs have been compiled thanks to the observation of this effect, for example by the *Planck* satellite. However, their use to constrain cosmological parameters requires the knowledge of cluster masses, which cannot be directly observed. Cosmological analyses therefore rely on scaling relations linking observables to cluster masses. These relations are calibrated on samples of few clusters with well-characterized properties, for example thanks to high-resolution observations. This detailed knowledge of cluster properties enables studying systematic effects taking place in cluster cosmology, such as the impact of relaxation state on the scaling relation.

This thesis takes place within the NIKA2 SZ large program, which aims to exploit high angular resolution observations of the SZ effect using the NIKA2 camera at the IRAM 30 meter telescope to study the physical properties of a sample of distant galaxy clusters. The work presented uses advanced statistical methods to study the systematics that need to be taken into account in cluster cosmology. Part of this work is dedicated to the analysis of NIKA2 data in order to characterize the thermodynamical properties of galaxy clusters while taking into account the systematics associated with SZ observations, leading to the development of a public software to constrain thermodynamical properties of a galaxy cluster from SZ observations. The methods developed are then used to characterize the properties of one of the most challenging clusters in the program, exploiting the synergy between NIKA2 and *XMM-Newton* observations of the cluster with comparable resolution. Subsequently, we present a rigorous statistical modeling of the mass-observable scaling relation for SZ observations, using a bayesian hierarchical model that allows us to take into account systematics affecting mass and observable measurements. We apply this analysis to realistic mock simulations, that allow us to highlight biases and systematic effects involved in the analysis, and to forecast the constraints that will be set on this scaling relation using NIKA2 observations.