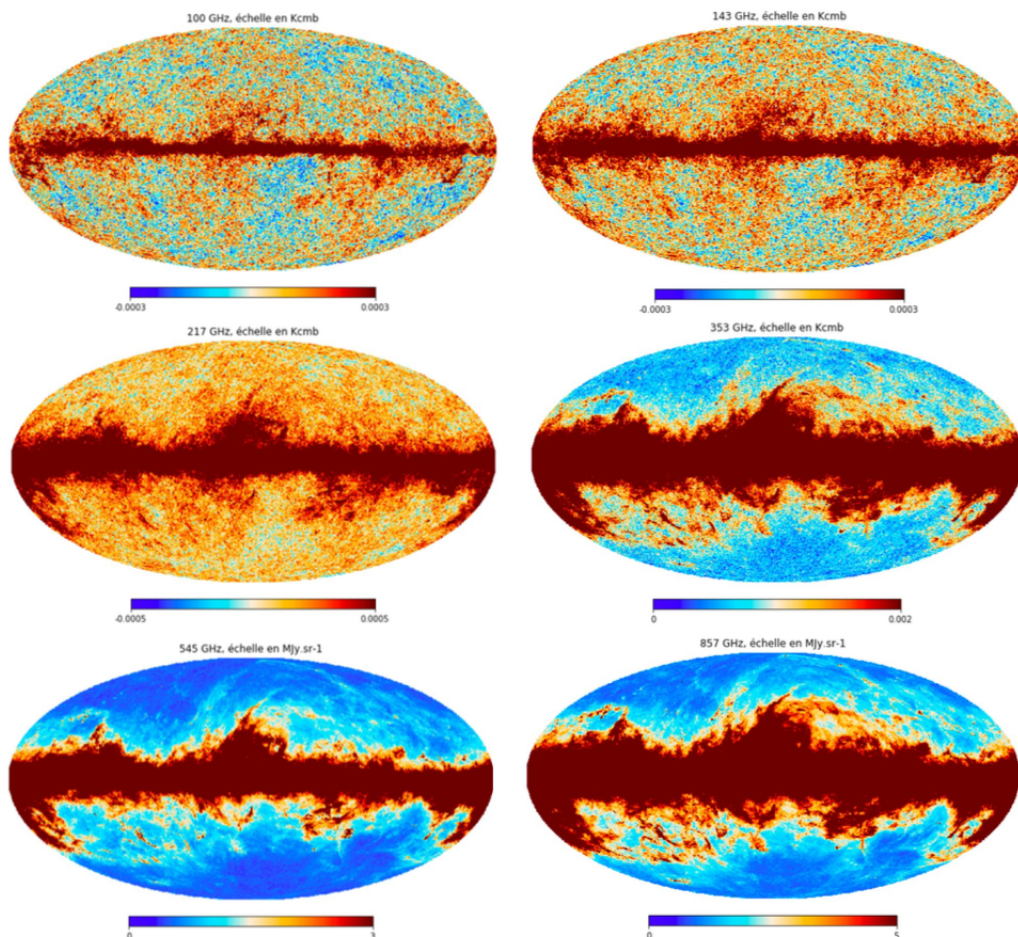

Modélisation et analyse des observations Planck, prévisions pour la mission LiteBird



Représentations de Mollweide du ciel observés sous les six canaux de fréquence Planck - de haut en bas et de gauche à droite : 100 GHz, 143 GHz, 217 GHz, 353 GHz, 545 GHz, 857 GHz.

Mélissa FAUR
Promo 2019 - Juillet 2020

Tuteur : M. François LEVRIER
Laboratoire : LPENS - AXE
ASTROPHYSIQUE

Résumé : Le satellite Planck avait pour mission de cartographier le rayonnement submillimétrique et infrarouge de l'ensemble du ciel, afin d'affiner l'étude du rayonnement primordial du Big Bang : le fond diffus cosmologique. Au cours de mon stage, j'ai simulé les résultats attendus du satellite pour une observation d'un nuage de poussières, émettant dans la même plage de fréquences que le fond diffus, que j'ai comparé aux cartes du nuage publiées par le collectif Planck. Cela m'a permis de comprendre les travaux d'analyse et traitement des données mis en place par le collectif scientifique. Par suite, j'ai comparé les performances du satellite Planck à celles de la future mission LiteBird, qui partage le même objectif que son prédécesseur : une étude plus précise du fond diffus.

Abstract : The Planck probe was launched to map the entire sky in the submillimetric and infrared domains, in order to study the Universe first light after the Big Bang : the Cosmic Microwave Background (CMB). During my internship, I simulated Planck's observational results of a gas cloud, in order to compare it to the official maps published by the researchers of the Planck Collaboration. Thanks to this simulation, I understood the necessity of data analysis and correctional algorithms devised by the collaboration to exploit Planck's raw data. To conclude my work, I compared the performances of the Planck mission and the future LiteBird mission which also aims at studying more precisely the Cosmic Microwave Background.

Table des matières

1	Présentation du sujet - Inscription dans le travail du laboratoire	3
2	Travaux effectués	3
2.1	Familiarisation avec le sujet d'étude	3
2.1.1	Création du signal d'un nuage de poussières typique	4
2.1.2	Loi de déplacement de Wien	4
2.1.3	Etude du nuage à partir des points transmis	4
2.2	Correction de couleur	5
2.2.1	Correction de couleur sur un paramètre - exemple du fond diffus cosmologique	5
2.2.2	Correction de couleur sur deux paramètres - exemple du nuage de poussières	5
2.3	Bruits et ajout de points	5
2.3.1	Ajout de bruit sur un signal de fond diffus cosmologique	6
2.3.2	Ajout de bruit sur un signal de nuage de poussières	6
2.3.3	Ajout d'un 7ème point aux données Planck	7
2.4	Etude de Chamaeleon-Musca	8
2.5	Comparaison LiteBird - Planck	9
2.5.1	Interpolation des signaux typiques	9
2.5.2	χ^2	10
3	Conclusion	11
4	Remerciements	11
5	Références bibliographiques	12
6	Annexe	13
6.1	Récupération des données	13
6.2	HealPix	13
6.3	Projection gnomonique	14
6.4	Ajout des signaux free-free et synchrotron	14
6.5	Cartes de température, indice spectral et profondeur optique	15

1 Présentation du sujet - Inscription dans le travail du laboratoire

Le sujet du stage m'a été proposé par François Levrier, chercheur au LPENS en astrophysique. Il a fait partie de la collaboration internationale qui a étudié les données du satellite Planck pendant une dizaine d'années, et travaille aujourd'hui sur la future mission LiteBird, qui cherche à sonder plus profondément le rayonnement du fond diffus cosmologique.

Ce stage a été réalisé en télétravail durant quatre semaines en juillet, au cours desquelles je n'ai pas eu l'occasion de me rendre au LPENS. Néanmoins, j'ai eu un contact régulier avec M. Levrier qui m'a permis de me familiariser rapidement avec mon sujet d'étude, et d'étudier en autonomie.

M. Levrier m'a confié comme objectifs principaux de modéliser sous *python* les données que devrait renvoyer le satellite Planck s'il observait des sources connues, comme le fond diffus cosmologique, ou un nuage de poussières de température et composition connues. Par suite, l'objectif était de comparer ces données théoriques d'observation aux véritables données publiées par le collectif Planck, de manière à saisir les stratégies de traitement de données mises en place par les chercheurs. Ses pour supprimer les bruits et biais de mesure. Enfin, cette étude du satellite Planck devait me permettre de comprendre et modéliser les résultats de la prochaine mission, LiteBird.

2 Travaux effectués

2.1 Familiarisation avec le sujet d'étude

Les satellites Planck et LiteBird ont pour objectif de cartographier le rayonnement du fond diffus cosmologique sur tout le ciel. Pour réaliser cet objectif, il leur faut détecter toutes les autres sources de rayonnement qui viennent cacher ce signal (Cf. [8]), notamment les nuages moléculaires émettant dans la plage de fréquence du fond diffus. C'est le cas des nuages moléculaires froids, se trouvant dans une phase stable entre 15K et 30K (Cf. carte de température p.11 [5]). Un très bon modèle de spectre d'émission de ces poussières à la température T est celui du corps noir modifié noté I_M : $I_M(\nu, \beta, \tau_0, T) = \tau_0(\frac{\nu}{\nu_0})^\beta B_\nu(T)$, où $B_\nu(T)$ correspond au spectre d'émission d'un corps noir à la température T , i.e. $B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}$.

Ici, β est l'indice spectral des poussières, qui dépend de leur nature, τ_0 est la profondeur optique du nuage à la fréquence ν_0 , quantité qui dépend du coefficient d'absorption des poussières à fréquence ν_0 donnée κ_0 (en $m^2.kg^{-1}$) et de la densité de colonne massique de poussière N_m (en $kg.m^{-2}$) sur la ligne de visée selon la formule : $\tau_0 = \kappa_0 N_m$. Ce modèle est détaillé en référence [4].

Je prendrai à plusieurs reprises par la suite comme exemple un nuage de poussières avec une profondeur optique de 10^{-4} , valeur typique que l'on peut trouver dans des nuages galactiques (Cf. carte de profondeur optique p.11 [5]), une température de 20K et un indice spectral de 2 qui correspond à un nuage poussières silicatées (Cf. p.64 [12]). Pour le fond diffus cosmologique, je modéliserai le spectre d'émission comme étant isotrope et correspondant à celui d'un corps noir de température $T_{CMB} = 2.725K$.

Les satellites Planck et LiteBird sont équipés de plusieurs capteurs d'intensité, chacun mesurant un point de données centré sur une fréquence. Pour Planck, il y a six canaux en haute fréquence qui sont centrés sur : 100 GHz, 143 GHz, 217 GHz, 353 GHz, 545 GHz et 857 GHz. Chacune de ces fréquences correspond à un capteur de bande passante entre 50 et 200 GHz centrée sur la fréquence du canal. Il y a aussi trois canaux à basse fréquence, que nous n'utiliserons pas dans cette étude. Pour un capteur donné, à chaque fréquence de sa bande passante ν est associée une valeur de transmission t_ν (Cf. Figure 1 - A). Ainsi, la mesure d'intensité transmise par le capteur centré en ν_0 vaut : $I_t(\nu_0) = \frac{\sum_{\nu \in BP} I(\nu) \cdot t_\nu}{\sum_{\nu \in BP} t_\nu}$, BP représentant les fréquences de la bande passante du capteur en ν_0 . Pour LiteBird, il y a neuf canaux haute fréquence, détaillés en partie 2.5.

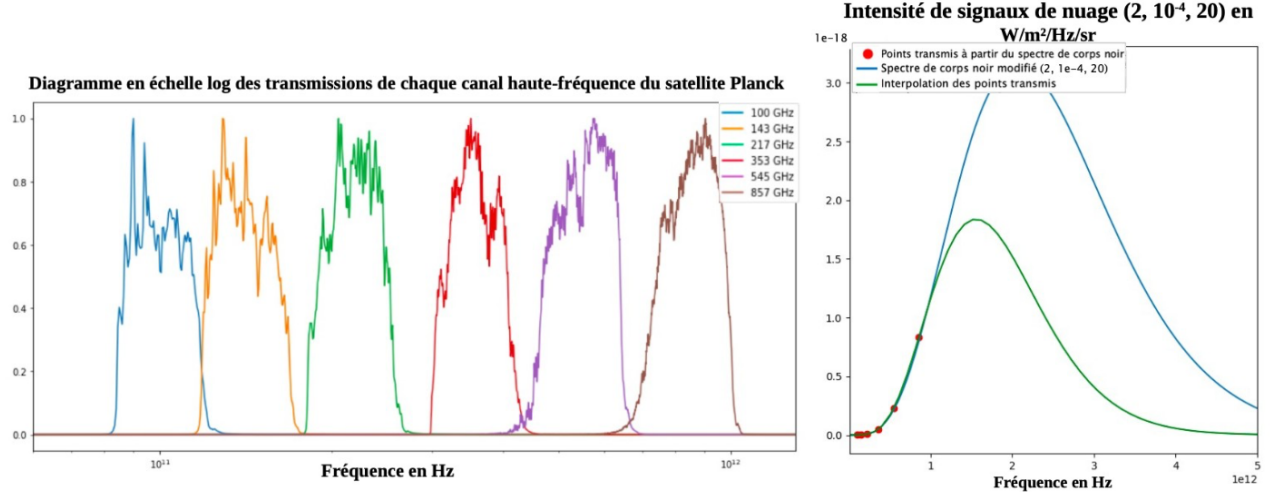


FIGURE 1 – A/ Transmissions pour chaque canal en fréquence du satellite Planck en fonction de la fréquence en Hz - B/ Spectre d'émission d'un nuage $(2, 10^{-4}, 20K)$, points transmis et interpolation des points en fonction de la fréquence en Hz

2.1.1 Création du signal d'un nuage de poussières typique

Reproduisons numériquement sous Python les signaux reçus par Planck lorsque qu'il observe un nuage moléculaire de paramètres $(\beta, \tau_0, T) = (2, 10^{-4}, 20K)$. J'ai commencé par créer une fonction $I_M(\nu, \beta, \tau_0, T)$ qui permet de calculer un spectre de corps noir modifié d'un nuage de poussières à partir d'un tableau de fréquences et de la donnée des trois paramètres (β, τ_0, T) . C'est la courbe en bleue sur le graphe 1 - B. A partir du tableau du spectre d'émission et des tables des transmissions que j'ai récupérées des fichiers officiels disponibles sur le site [9], j'ai recréé les six points de données renvoyés par les six capteurs de Planck (en orange sur la figure 1 - B).

2.1.2 Loi de déplacement de Wien

Pour compléter cette petite étude théorique du spectre d'émission des poussières, j'ai retrouvé numériquement la loi de déplacement de Wien $\nu_{max} = W_\beta T$ pour ces corps noirs modifiés à T . La fréquence maximale d'émission est une fonction de l'indice spectral et de la température, que l'on peut retrouver à indice spectral donné par une méthode de Newton, en résolvant l'équation suivante : $\frac{\partial I(\nu, \beta, \tau_0, T)}{\partial \nu} = 0 \Rightarrow e^x = \frac{3+\beta}{3+\beta-x}$ avec $x = \frac{h\nu}{k_B T}$. Le terme $3 + \beta$ est d'ordre unité, alors que le terme e^{x_0} peut prendre des valeurs plus importantes (par exemple pour $\nu = 2\text{THz}$ et $T = 20K$ ce terme vaut 123). On fait l'hypothèse que l'on peut écrire $x_0 = 3 + \beta + \epsilon$ avec $\epsilon \ll 1$ de manière à minimiser le dénominateur de la fraction. En utilisant $e^\epsilon \sim 1$ on obtient la loi de déplacement suivante : $\nu_{max} = \frac{k_B(3+\beta)[1+e^{-(3+\beta)}]}{h} \times T$. A $\beta = 2$, j'ai trouvé, à la fois numériquement et à partir de la formule ci-dessus, $W_\beta = 1,03 \times 10^{11} \text{ Hz/K}$. Ainsi, pour le nuage de paramètres $(2, 10^{-4}, 20K)$, on retrouve une fréquence d'émission du maximum d'intensité en $\nu_{max} = 2.07 \text{ THz}$, ce que l'on peut vérifier sur la figure 1 - B.

2.1.3 Etude du nuage à partir des points transmis

Je cherche alors à estimer la composition, la quantité de poussières et la température du nuage en interpolant les six points obtenus par la fonction $I_M(\nu, \beta, \tau_0, T)$, grâce à `scipy.optimize.curve_fit`. On obtient l'estimation suivante des paramètres : $(2.59 \pm 0.20, 1.8 \times 10^{-4} \pm 0.3 \times 10^{-4}, 13.4 \pm 1.3K)$ (Cf. courbe verte sur la figure 1 - B). On ne retrouve pas les valeurs des paramètres initiaux dans les intervalles à 68%. Je comprends donc un premier problème associé à l'analyse des données de nuage moléculaires du satellite :

une simple interpolation par le spectre d'émission d'un corps noir modifié ne permet pas de remonter aux valeurs de (β, τ_0, T) car il manque des points pour contraindre l'interpolation. En effet, je vois que les six points que j'obtiens ne se trouvent que d'un côté de la courbe en cloche. L'ajout d'un septième point à l'opposé permettrait à l'interpolation d'être plus précise. De plus, les données ne sont pas corrigées : observons dès maintenant ce qu'une correction peut apporter à l'estimation des paramètres.

2.2 Correction de couleur

Une solution pour corriger les données du satellite Planck de manière à ce que les interpolations retournent des paramètres plus proches de ceux attendus est la correction de couleur.

2.2.1 Correction de couleur sur un paramètre - exemple du fond diffus cosmologique

J'ai construit ma correction de couleur de la façon suivante : d'abord j'ai créé un tableau t_T contenant N_T points de température dans l'intervalle $[2.7, 2.8]K$. Par suite, pour chaque canal de fréquence ν_0 , j'ai calculé le point de correction ainsi : $C_{\nu_0} = \frac{\sum_{i \in [0, N_T-1]} B_{\nu_0}(t_T[i]) - B_t(\nu_0, t_T[i])}{N_T}$ où ici la quantité $B_t(\nu_0, t_T[i])$ correspond au point transmis par Planck à ν_0 d'un spectre de corps noir à la température $t_T[i]$. En effectuant cela pour chaque canal, il ne reste plus qu'à additionner aux signaux transmis ces points de correction avant d'interpoler. Sans correction de couleur, les six points transmis par le satellite Planck puis interpolés retournent cette estimation de la température du fond diffus : $T_1 = 2.73 \pm 0.02K$, ce qui est déjà correct. Si on y ajoute la correction de couleur décrite ci-dessus, on obtient : $T_2 = 2.725 \pm 0.001K$

En résumé, ne serait-ce qu'une connaissance grossière, "entre 2.7K et 2.8K", de la température du fond diffus apporte une précision d'un ordre de grandeurs sur l'écart-type de l'interpolation, et réduit le biais entre la valeur interpolée et la valeur attendue.

2.2.2 Correction de couleur sur deux paramètres - exemple du nuage de poussières

On peut mettre en place une correction de couleur à deux paramètres pour le nuage de paramètres $(2, 10^{-4}, 20K)$, c'est-à-dire une correction à la fois pour un intervalle de température, ici $[19, 21]K$, et un intervalle d'indice spectral, $[1.8, 2.2]$. On fixe $\tau_0 = 10^{-4}$. Le point de correction à ν_0 s'exprime alors de la façon qui suit : $C_{\nu_0} = \frac{\sum_{i \in [0, N_\beta-1], j \in [0, N_T-1]} I_M(\nu_0, t_\beta[i], 10^{-4}, t_T[j]) - I_t(\nu_0)}{N_T N_\beta}$, en définissant t_β et N_β comme pour la température, et en considérant des spectres de corps noir modifié. Après correction et interpolation, les paramètres retournés sont $(2.000 \pm 2 \times 10^{-3}, 1.0002 \times 10^{-4} \pm 2 \times 10^{-8}, 20.000K \pm 3 \times 10^{-3}K)$. On retrouve avec une excellente précision les paramètres attendus !

Néanmoins, il n'est pas très réaliste de considérer que l'on connaît précisément dans quels intervalles vont se trouver la température et l'indice spectral du nuage étudié pour configurer la correction de couleur. Par exemple, si on prend un intervalle non centré sur 20K pour la correction en température, comme $[19, 20]K$, le signal corrigé interpolé présente les paramètres : $(2.03 \pm 0.01, 1.03 \times 10^{-4} \pm 0.01 \times 10^{-4}, 19.6 \pm 0.1K)$. Pour les deux premiers paramètres, on retrouve les valeurs initiales dans les intervalles à 68 %. Pour les valeurs de température, on ne retrouve pas la température de 20K dans l'intervalle à 68 % car la température tend vers celle du milieu de l'intervalle de correction $[19, 20]K$ soit 19.5K. Malgré tout, on a réduit le biais sur la température d'un ordre de grandeur, passant de 6.6 K sans correction à 0.4 K avec.

En bref, la correction de couleur, même décentrée, apporte un gain de précision et réduit le biais sur les paramètres de l'interpolation.

2.3 Bruits et ajout de points

Les données récupérées par les satellites sont emplies de divers bruits, comme les bruits des instruments, modélisés ici par un bruit gaussien centré en 0. J'ai donc cherché à comprendre comment les interpolations

Résultats partie 2.3				
Signaux type Planck		Indice spectral	Profondeur optique	Température en K
Bruit et Monte-Carlo	fond diffus corrigé sans bruit			2.725 ± 0.001
	fond diffus corrigé avec bruit en $10^{-20}W/m^2/Hz/sr$			2.723 ± 0.002
	fond diffus corrigé avec bruit en $10^{-19}W/m^2/Hz/sr$			2.714 ± 0.008
	fond diffus corrigé avec Monte-Carlo, bruit en $10^{-19}W/m^2/Hz/sr$			2.73 ± 0.02
	nuage $(2, 10^{-4}, 20K)$ corrigé sans bruit	2.000 ± 0.002	$1.0002 \times 10^{-4} \pm 2 \times 10^{-8}$	20.000 ± 0.003
	nuage $(2, 10^{-4}, 20K)$ corrigé avec bruit en $10^{-21}W/m^2/Hz/sr$	2.03 ± 0.11	$1.0 \times 10^{-4} \pm 0.1 \times 10^{-4}$	19.4 ± 1.8
	nuage $(2, 10^{-4}, 20K)$ corrigé avec Monte-Carlo, bruit en $10^{-21}W/m^2/Hz/sr$	2.00 ± 0.12		20.1 ± 2.0
Ajout du septième point	nuage $(2, 10^{-4}, 20K)$ sans correction	1.97 ± 0.05	$1.07 \times 10^{-4} \pm 0.01 \times 10^{-4}$	$20.0 \pm 0.2K$
	nuage $(2, 10^{-4}, 20K)$ corrigé	1.997 ± 0.001	$1.000 \times 10^{-4} \pm 0.001 \times 10^{-4}$	20.02 ± 0.01
	nuage $(2, 10^{-4}, 20K)$ corrigé avec Monte-Carlo, bruit en $10^{-21}W/m^2/Hz/sr$	1.97 ± 0.01		20.0 ± 0.1
	nuage $(2, 10^{-4}, 20K)$ corrigé avec Monte-Carlo, bruit en $10^{-20}W/m^2/Hz/sr$	1.98 ± 0.12		19.9 ± 0.6

FIGURE 2 – Résultats de la partie 2.3

des signaux transmis sont influencées par l'ajout de bruit. L'ensemble des résultats de cette partie est compilé dans le tableau de la figure 2.

2.3.1 Ajout de bruit sur un signal de fond diffus cosmologique

J'ai considéré tout d'abord le signal transmis par Planck du fond diffus, corrigé avec la correction présentée en partie précédente, et auquel on ajoute un bruit gaussien de moyenne nulle et d'écart type $10^{-20}W/m^2/Hz/sr$ (Cf. Figure 3 - A). On obtient un écart-type du même ordre que celui trouvé sans bruit. Je passe donc à un bruit en $10^{-19}W/m^2/Hz/sr$: on remarque une augmentation du biais sur la valeur de la température interpolée, et une multiplication par dix de l'écart-type sur celle-ci (voir le décalage des courbes sur la figure 3 - A). J'ai alors cherché une méthode pour réduire cette influence du bruit sur les interpolations de signaux.

Je me suis intéressée à la méthode statistique dite de « Monte-Carlo ». Il s'agit d'effectuer N fois l'opération précédente, à savoir ajouter un bruit gaussien au signal corrigé et l'interpoler, et de stocker les valeurs de température trouvées dans un tableau t_{MC} . La moyenne de ce tableau et sa variance, $Var(t_{MC}) = \frac{\sum_i (t_{MC}[i] - T_{MC})^2}{N}$ permettent d'estimer la température interpolée moyenne et son incertitude. J'ai effectué ces calculs pour $N = 100$ et un bruit d'écart-type $10^{-19}W/m^2/Hz/sr$. Ainsi, la mise en place de cette méthode statistique permet de réduire le biais sur la température interpolée des signaux bruités, qui passe de 0.011 K à 0.001K, mais ne permet pas d'améliorer l'ordre de grandeur de la précision sur ce paramètre, qui reste le même que sans méthode statistique.

2.3.2 Ajout de bruit sur un signal de nuage de poussières

J'ai considéré ensuite un signal transmis par Planck à partir du nuage $(2, 10^{-4}, 20K)$, que l'on corrige avec une correction centrée sur les paramètres, et auquel on ajoute un bruit gaussien de moyenne nulle et d'écart type $10^{-21}W/m^2/Hz/sr$. Ce spectre présente une intensité maximale à la fréquence 2.07 THz qui se trouve être en dehors des canaux de fréquence de Planck, à la différence d'un signal de fond diffus. Je m'attends donc à une plus grande sensibilité au bruit que pour le signal de fond diffus. D'après la table, on retrouve encore les valeurs attendues dans les intervalles à 68%, mais avec des écarts-types d'au moins un ordre de grandeur plus grands que ceux sans bruit.

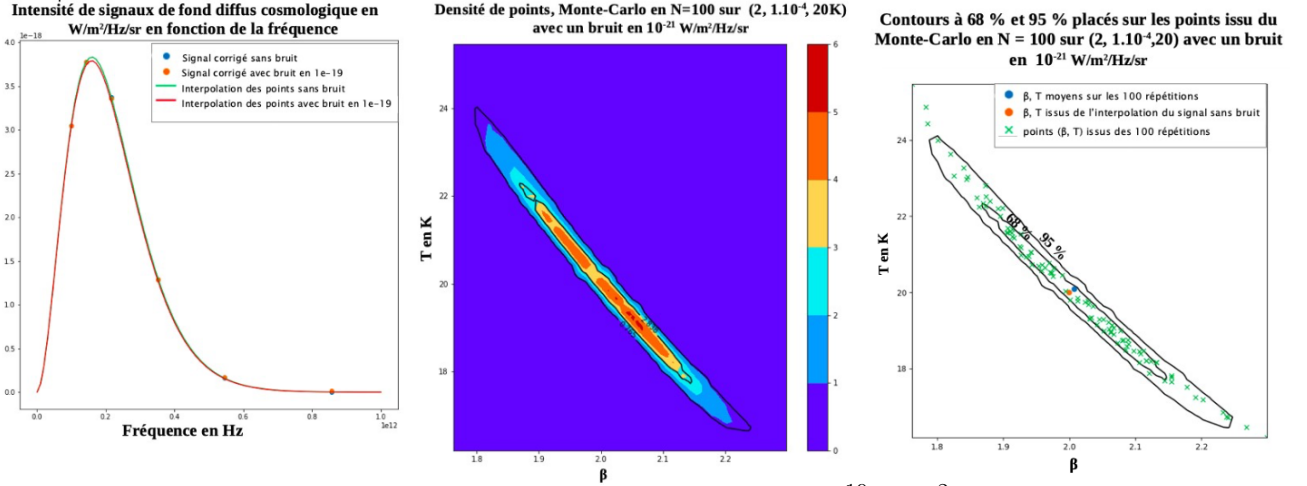


FIGURE 3 – A/ Points transmis du fond diffus cosmologique bruités en $10^{-19} \text{ W/m}^2/\text{Hz}/\text{sr}$ et interpolation - B/ Densité de points issus du Monte-Carlo sur le nuage $(2, 10^{-4}, 20\text{K})$ en fonction de l'indice spectral et de la température en K - C/ Contours à 68% et 95% de cette densité de points

J'applique par la suite une méthode de « Monte-Carlo » à deux paramètres, indice spectral et température, avec $\sigma_{\text{bruit}} = 10^{-21} \text{ W/m}^2/\text{Hz}/\text{sr}$. On obtient alors la carte centrale de la figure 3, qui représente la densité de points (β, T) issus de 100 répétitions, et on détermine un indice spectral et une température moyens ainsi que leurs écarts-types. Je remarque que la température et l'indice spectral moyens issus de cet algorithme sont très proches des valeurs des paramètres issus du signal corrigé interpolé sans bruit, comme le montre la superposition des deux points sur la figure 3 - C, mais les écarts-types restent du même ordre de grandeur que ceux du signal bruité seul.

En conclusion, l'algorithme de « Monte-Carlo » permet d'optimiser la détermination des paramètres d'un signal bruité, mais n'influe pas sur les incertitudes.

2.3.3 Ajout d'un 7ème point aux données Planck

Pour endiguer la vulnérabilité au bruit des données Planck, d'autant plus présente pour des signaux de nuage, nous pouvons ajouter un point de mesure fictif de l'autre côté de la cloche, à 3 THz comme expliqué plus haut. La transmission de ce nouveau canal est celle du canal 217 GHz que j'ai recentrée sur 3 THz. Il est intéressant de rajouter ce septième point sachant que les cartes officielles de température, indice spectral et de profondeur optique publiées par le collectif Planck correspondent à des interpolations de signaux de sept points, constitués en ajoutant un point vers 3 THz issu des données de la mission IRAS [2].

En premier lieu, avec l'ajout de ce septième point sans aucun bruit ni correction de couleur, on obtient pour l'indice spectral et la température des quantités qui contiennent les valeurs attendues dans l'intervalle à 68 %. Ce n'est pas le cas pour la profondeur optique, qui n'est pas retrouvée dans l'intervalle à 68 % mais qui est beaucoup moins biaisée que sans septième point ni correction (0.07 plutôt que 0.8). En y rajoutant une correction de couleur centrée, les valeurs sont encore moins biaisées que précédemment, et les écarts-types sont plus petits d'un ordre de grandeur. Testons maintenant la sensibilité au bruit pour un signal de nuage $(2, 10^{-4}, 20\text{K})$ sans correction par l'algorithme de Monte-Carlo. Puisque j'ai obtenu de très bons résultats pour un bruit en 10^{-21} , à savoir des valeurs attendues contenues dans les intervalles à 68% et des écarts-types de l'ordre de ceux obtenus pour des signaux non corrigés, j'ai fait l'étude pour un bruit en 10^{-20} . Alors, les valeurs de paramètres récupérées sont peu biaisées, et ont des écarts-types du même ordre de grandeur que ceux trouvés pour des bruits en 10^{-21} sur les signaux à six points corrigés.

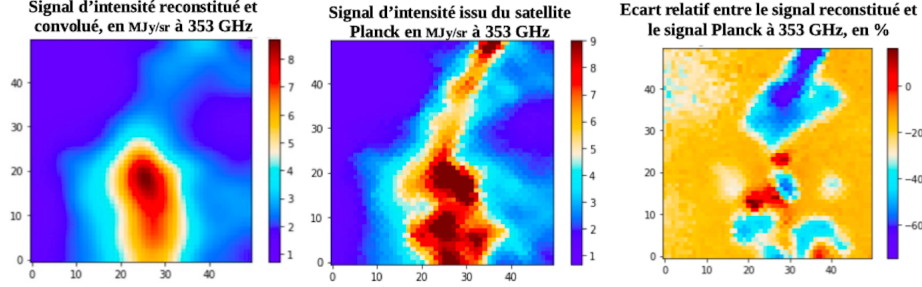


FIGURE 4 – A/ Carte d'intensité à 353 GHz reconstituée et convoluée - B/ Carte d'intensité à 353 GHz mesurée par le satellite Planck - C/ Ecart relatif de ces deux cartes en %

Pour résumé, l'ajout du septième point aux données du satellite Planck diminue d'un facteur dix la sensibilité au bruit du satellite, et offre la possibilité de s'affranchir de la correction de couleur qui peut être difficile à calibrer.

2.4 Etude de Chamaeleon-Musca

Dans cette partie, je me suis concentrée sur l'étude du nuage Chamaeleon-Musca. A partir des cartes officielles ouvertes sous *python*, j'ai effectué six projections gnomoniques de taille en pixels 50x50 centrée sur le point de longitude et latitude (297.3°, -15.5°) avec une résolution de 1.5'/pixel (Cf. [6] pour la position du nuage). Pour plus de détails sur le traitement des fichiers officiels sous *python* et la création de projection, je vous invite à lire l'annexe. Pour les canaux en dessous de 353 GHz inclus, les cartes sont renvoyées en K_{CMB} , une unité de mesure d'intensité lumineuse qui se fonde sur le rayonnement de corps noir du fond diffus cosmologique (Cf. [5]). C'est dans la référence [7] que j'ai pu trouver les coefficients calculés par le collectif Planck pour effectuer la conversion de K_{CMB} en MJy/sr pour chacun de ces 4 canaux de fréquence. De cette façon, j'ai obtenu six cartes en 50x50 pixels, représentant l'émission du nuage Chamaeleon-Musca à six fréquences différentes, en MJy/sr , soit en unité SI en $10^{-20}W/m^2/Hz/sr$.

Nous rencontrons tout d'abord sur ces cartes une première difficulté : l'ordre de grandeur de l'intensité émise est de $10MJy/sr$ au plus (Cf. figure 4 - B), alors que l'ordre de grandeur de l'émission du fond diffus cosmologique est de $100MJy/sr$ sur les canaux de fréquence Planck (Cf. figure 3 - A). J'ai donc considéré que les cartes que j'avais récupérées sur le site en référence [9] avaient déjà été privées du rayonnement du fond diffus. J'ai voulu vérifier cette hypothèse en recréant des cartes d'émission aux différents canaux à partir des cartes de température, indice spectral et profondeur optique publiées par le collectif Planck (Cf. [7]). J'ai donc récupéré ces cartes, effectué une projection gnomonique centrée sur Chamaeleon-Musca comme précédemment, et ensuite reconstituée à partir du triplet (β, τ_0, T) en chaque pixel les six points transmis par le satellite Planck. Pour obtenir des cartes comparables avec les cartes de données 'brutes' Planck, il me restait à prendre en compte l'influence optique de la taille finie du télescope sur les données. En effet, le télescope Planck est composé d'un lobe de taille 5' : j'ai donc convolué les cartes reconstituées avec un lobe de taille 5' pour obtenir des cartes comparables aux cartes 'brutes'. Observons le résultat sur la figure 4 - A ; nous voyons bien que cela ne colle toujours pas aux données Planck 'brutes' (Cf. figure 4 - B), puisqu'il y a des zones d'écart relatif supérieur à 50% d'après la figure 4 - C.

Une autre piste que j'ai décidé d'explorer pour retrouver ces cartes 'brutes' a été d'ajouter les autres rayonnements à mes cartes reconstituées d'émission de nuages, en me concentrant sur les deux principaux : le synchrotron et le free-free. Malheureusement, l'ajout de ces deux cartes à mes cartes reconstituées n'arrangent pas les écarts avec les cartes 'brutes' (Cf. [8], annexe). Ma dernière piste a été de considérer que le problème venait de mes conversions d'unité. Je détaille cette petite étude en annexe.

Ce bout d'étude m'a permis de me confronter aux soucis d'unité des données : toutes ne sont pas être

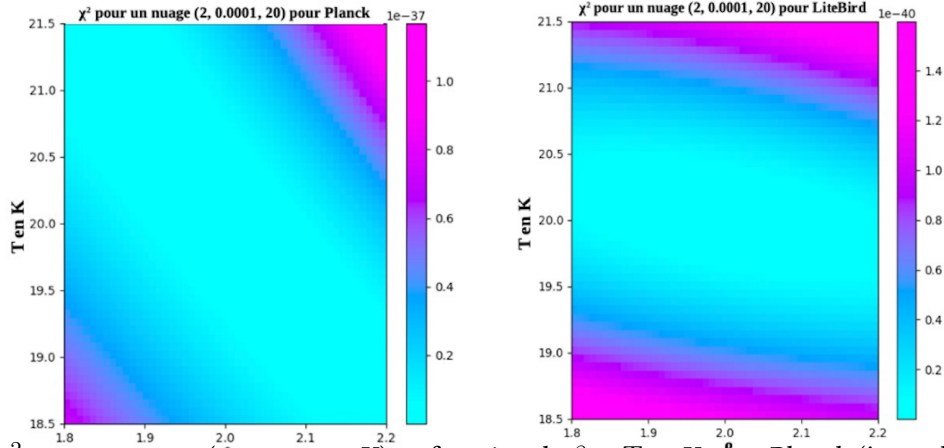


FIGURE 5 – χ^2 pour un nuage $(2, 0.0001, 20 \text{ K})$ en fonction de β et T en K pour Planck (à gauche) et LiteBird (à droite)

mesurées dans des unités SI, pour des raisons premières de stockage de données, mais aussi de pratique (les astrophysiciens sont plus habitués à utiliser des unités de température pour mesurer les intensités). Il ne me manquait que de temps et de lectures plus poussées des différents articles de la collaboration Planck pour éclairer ces zones d'ombres et reconstituer fidèlement les cartes brutes.

2.5 Comparaison LiteBird - Planck

La mission LiteBird a été sélectionnée en mai 2019 parmi d'autres projets internationaux pour étudier la polarisation du fond diffus cosmologique lors d'une mission de trois ans qui sera lancée dans la décennie 2020. C'est une mission japonaise à laquelle ont été conviés des membres de la collaboration Planck pour leur expérience. Le futur satellite LiteBird disposera de 15 canaux (Cf. [3]), dont neuf en haute fréquence (supérieure à 100 GHz) : 100 GHz, 119 GHz, 140 GHz, 166 GHz, 195 GHz, 235 GHz, 280 GHz, 337 GHz et 402 GHz. A la différence de Planck, il n'y a pas de canal de fréquence supérieure à 500 GHz. Voyons en quoi cela va influencer sur les interpolations de signaux typiques.

2.5.1 Interpolation des signaux typiques

Pour créer les signaux "transmis" par LiteBird, qui n'a pas encore été conçu, j'ai utilisé la transmission à 217 GHz du satellite Planck que j'ai centrée sur les bonnes fréquences. Pour le signal de fond diffus, sans correction de couleur, une interpolation des neuf points transmis par LiteBird donne une température de $2.72 \pm 0.01 \text{ K}$. De son côté, Planck donnait : $2.73 \pm 0.15 \text{ K}$. L'interpolation de LiteBird donne un résultat plus biaisé que Planck, mais avec une plus grande précision. Ajoutons-y une correction de couleur; on obtient : $T = 2.73 \pm 0.02 \text{ K}$. Planck, de son côté, pour un signal de fond diffus corrigé présentait un biais de $5 \times 10^{-4} \text{ K}$ et un écart-type de $5 \times 10^{-4} \text{ K}$. Ainsi, avec LiteBird, on gagne en écart sur le biais, mais on perd un ordre de grandeur en précision.

Le satellite LiteBird est aussi amené à étudier des signaux d'extrema atteints à des fréquences au delà de 1 THz. En effet, comme Planck, LiteBird va cartographier les signaux occultants le fond diffus, comme les signaux issus de nuages moléculaires qui émettent typiquement autour du THz. De fait, passons maintenant au nuage $(2, 10^{-4}, 20)$. Une interpolation des neuf points sans correction de couleur donne : $(1.51 \pm 0.24, 1.10^{-5} \pm 5.10^{-5}, 118 \pm 409 \text{ K})$. Ces résultats ne sont même pas exploitables car les écarts-types sur la profondeur optique et la température sont de l'ordre de grandeur des valeurs interpolées ! Avec une correction de couleur, on obtient : $(2.0009 \pm 0.0003, 1.003 \times 10^{-4} \pm 0.001 \times 10^{-4}, 19.96 \pm 0.01 \text{ K})$. Le biais sur les valeurs interpolées est fortement diminué, et les écarts-types sur celles-ci sont aussi plus faibles.

que pour les interpolations sur les données Planck. Le fait que LiteBird n'ait pas de point à une fréquence au dessus de 500 GHz fait que les interpolations sur les signaux de maxima de l'ordre du THz sont plus biaisées que celles de Planck, mais restent plus précises si l'on y associe une correction de couleur bien centrée.

2.5.2 χ^2

A propos du bruit, d'après les prévisions publiées par le collectif LiteBird, le satellite sera 10 fois plus sensible que Planck. De fait, on peut refaire les mêmes études de bruit que pour Planck, en prenant la précaution de diviser les bruits appliqués par dix pour chaque signal.

Enfin, j'ai effectué une dernière comparaison des précisions des deux satellites en effectuant une étude du coefficient χ^2 . Commençons par le définir (Cf. [11]). On cherche à mesurer l'écart entre le modèle f , évalués en les points $\{x_i\}$, et les points de mesures correspondants $\{y_i\}$. Dans le cas présent où l'incertitude sur chaque point de mesure est la même, le χ^2 vaut : $\chi^2 = \sum_i (f(x_i) - y_i)^2$. Dans mon étude, $\{f(x_i)\}$ correspond à un signal transmis par le satellite de paramètres connus et y_i correspond à un signal transmis de paramètres à déterminer, appelé 'signal test'. La somme sur i s'apparentera à une somme sur les canaux de fréquence. En premier lieu, pour le satellite Planck, on se place à un canal de fréquence donné : 857 GHz. On prend comme signal test le signal transmis par Planck (non corrigé) avec les paramètres $(2, 10^{-4}, 20K)$. A profondeur optique fixe de $\tau_0 = 10^{-4}$, je fais varier la température sur l'intervalle $[18.5, 21.5]K$ et l'indice spectral sur l'intervalle $[1.8, 2.2]$. En chaque point de ce plan (β, T) , je calcule le point transmis I_t par Planck pour un nuage (β, τ_0, T) , sans correction. J'obtiens de cette manière une carte d'intensité transmise à 857 GHz en fonction de la température et de l'indice spectral. J'effectue cela pour chacun des canaux et je calcule le χ^2 en chaque point du plan (β, T) (Cf. Figure 5). On s'attend à trouver un minimum de cette carte au point $(\beta, T) = (2, 20K)$. Il s'avère qu'il vaut $\chi_{Planck}^2 = 5 \times 10^{-44} (W/m^2/Hz/sr)^2$, atteint au point $(2.00, 19.97K)$. Les biais sur l'indice spectral et la température sont 100 fois plus petits que ceux trouvés par interpolation avec *scipy* sans correction. Ainsi, cette comparaison des points transmis par rapport à un point test permet de déterminer la composition, quantité, et température du nuage correspondant aux points test de manière plus précise que *scipy*, et sans correction de couleur. Cependant, ce n'est pas si simple. Ici, le τ_0 est fixé à la bonne valeur, mais si je regarde les coordonnées du minimum obtenues si je prends une profondeur optique de 1.8×10^{-4} j'ai $(2.8, 12.5K)$. De la même façon, l'algorithme de *scipy* explore à la fois des intervalles sur β , τ_0 et T tombe sûrement sur un minimum local dans une vallée (comme les vallées bleues visible sur les cartes en figure 5) à un τ_0 différent de 10^{-4} sans explorer le reste des degrés de liberté, ce qui explique les biais importants des valeurs interpolées.

On effectue le même calcul de χ^2 pour LiteBird et on obtient $\chi_{LB}^2 = 4 \times 10^{-44} (W/m^2/Hz/sr)^2$ au point $(2.01, 19.97K)$. On peut comparer ces deux quantités à partir du coefficient χ^2 réduit qui se trouve avec la formule suivante, avec n le nombre de points et p le nombre de paramètres : $\chi_r^2(n, p) = \frac{\chi^2}{n-p}$. Après calcul, on a $\chi_{r, LB}^2(9, 3) < \chi_{r, Planck}^2(6, 3)$. En résumé, l'algorithme de χ^2 pour des données de type LiteBird retourne un minimum plus petit qu'à partir des données de type Planck : l'estimation des paramètres du nuage test avec LiteBird minimise davantage l'écart entre le signal minimal et le signal test que Planck.

Pour conclure, le satellite LiteBird est moins précis sur des analyses de signaux du fond diffus malgré la correction de couleur, mais les interpolations de ses signaux de nuage moléculaire avec correction sont plus précises que celles de Planck. De plus, les données type LiteBird permettent de déterminer plus précisément que les données type Planck les paramètres des nuages test d'après la comparaison des χ_r^2 .

3 Conclusion

Au cours de ce stage, j'ai pu découvrir en détail en quoi consistait le travail de traitement de données d'une mission satellite. La première étape a été de choisir un modèle de source d'émission, de le modéliser et de simuler le signal de sortie du satellite. Ces premières prises en main faites, j'ai pu rapidement saisir les difficultés associées à l'interpolation de si (six) peu de points, sujets au bruit, et aux aléas de la configuration des fonctions *python*. Néanmoins, j'ai exploré la piste de la correction de couleur qui permet d'améliorer significativement les interpolations de paramètres, moyennant une connaissance grossière de leur valeur. Après cela, j'ai pu me confronter à la difficulté de l'analyse des cartes, qui doivent subir des conversions et ajustements avant d'être exprimées en unités SI. Enfin, ce savoir-faire acquis m'a permis de comparer sur différents volets les satellites Planck et LiteBird, en ne partant que de la liste des canaux en fréquence de LiteBird, de sa sensibilité et de sa taille. Ce petit nombre de données a suffi pour démontrer le gain de précision apporté par l'utilisation du satellite LiteBird.

Pour conclure sur l'organisation de mon stage, j'ai réussi à bien m'adapter au télétravail. Je suis restée efficace sur les quatre semaines de stage par des échanges journaliers avec mon tuteur de stage, et par l'intérêt que je portais au sujet d'étude. Je retiendrai de ce stage une première expérience de travail long sur un projet, mais aussi une découverte assez exhaustive du travail d'analyse mené par les collaborations scientifiques sur des missions comme la mission Planck.

4 Remerciements

Je remercie M. François Levrier de m'avoir accompagnée dans ces travaux à distance. J'ai pu toucher du doigt le travail d'analyse de données de satellites, et j'ai été très bien guidée pour gagner rapidement en autonomie et parvenir à prendre des initiatives dans mon travail.

5 Références bibliographiques

- [1] Hivon, E., *The Healpix primer*, astro-ph/9905275, 1999.
- [2] Beichman, C.A., *Infrared Astronomical Satellite (IRAS) Catalogs and Atlases. Volume 1 : Explanatory Supplement*, NASA reference publication 19890004823, 1988
- [3] Ghigna, T. and Matsumura, T. and Patanchon, G. and Ishino, H. and Hazumi, M., *Requirements for future CMB satellite missions : photometric and band-pass response calibration*, hal-02572507, May 2020
- [4] Planck Collaboration : Abergel, A. and Ade, P. A. R. and Aghanim, N. and Alves, M. I. R. and Aniano, G. and Armitage-Caplan, C. and Arnaud, M. and Ashdown, M. and Atrio-Barandela, F. and et al., "Planck 2013 results. XI. All-sky model of thermal dust emission", in *Astronomy & Astrophysics* 571, A11 (2014). DOI : 10.1051/0004-6361/201323195
- [5] Planck Collaboration : Ade, P. A. R. and Aghanim, N. and Armitage-Caplan, C. and Arnaud, M. and Ashdown, M. and Atrio-Barandela, F. and Aumont, J. and Baccigalupi, C. and Banday, A. J. and et al., "Planck 2013 results. IX. HFI spectral response", in *Astronomy & Astrophysics* 571, A9 (2014). DOI : 10.1051/0004-6361/201321531
- [6] Planck Collaboration : Ade, P. A. R. and Aghanim, N. and Alina, D. and Alves, M. I. R. and Aniano, G. and Armitage-Caplan, C. and Arnaud, M. and Ashdown, M. and Atrio-Barandela, F. and et al., "Planck intermediate results. XX. Comparison of polarized thermal emission from Galactic dust with simulations of MHD turbulence", in *Astronomy & Astrophysics* 576, A105 (2015). DOI : 10.1051/0004-6361/201424086
- [7] Planck Collaboration : N. Aghanim and Y. Akrami and M. Ashdown and J. Aumont and C. Baccigalupi and M. Ballardini and A. J. Banday and R. B. Barreiro and N. Bartolo and et al, "Planck 2018 results. III. High Frequency Instrument data processing and frequency maps", in *Astronomy & Astrophysics* (2018), arXiv :1807.06207
- [8] Planck Collaboration : Akrami,Y., and Ashdown, M. and Aumont, J. and Baccigalupi, C., and Ballardini, M. and Banday, A. J. and Barreiro, R. B. and Bartolo, N. and Basak, S. and et al., *Planck 2018 results. IV. Diffuse component separation*, arXiv :1807.06208, 2018.
- [9] Site compilant les données officielles obtenues par la mission Planck [Planck Legacy Archive](#)
- [10] Page wiki mise en place par le collectif Planck pour compléter le site de compilation de données [Explanatory supplement](#)
- [11] Le cours "Incertitudes expérimentales" de F. Levrier, *Master Astronomie, Astrophysique et Ingénierie Spatiale*, 2018.
- [12] Le cours "Processus électromagnétiques, transfert de rayonnement et interaction matière-rayonnement" de F. Levrier, *Master Astronomie, Astrophysique et Ingénierie Spatiale*, 2018.
- [13] Le site de Yann Olivier à propos des [projections gnomoniques](#)

6 Annexe

6.1 Récupération des données

Pour avancer dans mon étude des satellites, il m'a fallu accéder aux données envoyées par le satellite et analysées par les chercheurs, qui sont enregistrées dans des fichiers FITS disponibles sur le site [9]. Ceux-ci peuvent être lu par la bibliothèque *astropy*, qui offre aussi plein d'autres outils utiles à l'astrophysicien, comme la possibilité de donner des unités aux scalaires et tableaux.

Filename: HFI_RIMO_R3.00.fits

No.	Name	Ver	Type	Cards	Dimensions	Format
0	HFI_RIMO	1	PrimaryHDU	14	()	
1	MAP_PARAMS	1	BinTableHDU	50	6R x 10C	[3A, 1D, 1D, 1D, 1D, 1D, 1D, 1D, 1D, 1D]
2	BANDPASS_F100	1	BinTableHDU	28	12297R x 3C	[1E, 1E, 1A]
3	BANDPASS_F100	1	BinTableHDU	28	12297R x 3C	[1E, 1E, 1A]
4	BANDPASS_F143	1	BinTableHDU	28	12602R x 3C	[1E, 1E, 1A]
5	BANDPASS_F217	1	BinTableHDU	28	12600R x 3C	[1E, 1E, 1A]
6	BANDPASS_F353	1	BinTableHDU	28	12930R x 3C	[1E, 1E, 1A]
7	BANDPASS_F545	1	BinTableHDU	28	13930R x 3C	[1E, 1E, 1A]
8	BANDPASS_F857	1	BinTableHDU	28	16819R x 3C	[1E, 1E, 1A]
9	BANDPASS_PSB353	1	BinTableHDU	28	12930R x 3C	[1E, 1E, 1A]

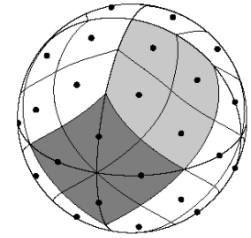
Le premier fichier important à analyser est le *RIMO*, pour "reduced instrumental model". Dans ce fichier sont stockées les transmissions de chaque capteur haute fréquence du satellite, et les fréquences associées. Cela m'a permis de tracer la figure 1 - A représentant les différentes bandes passantes du satellite Planck.

FIGURE 6 - 'header' du fichier RIMO de Planck

Afin de lire ce fichier, j'ai utilisé la bibliothèque *astropy*, qui récupère le fichier sous forme d'un objet avec deux attributs : "header" et "data". Les "headers" sont les en-têtes contenues dans le fichier, qui nous renseignent sur la nature et la taille des données accessibles, ce qu'illustre la figure 7. L'attribut "data" permet de récupérer les données décrites dans le "header". Le tout aide à facilement lire et récupérer des données de fichiers de plusieurs gigaoctets.

6.2 HealPix

Les cartes du ciel sont codées selon la méthode HealPix [1], c'est-à-dire que la sphère céleste est codée dans un tableau à une dimension où chaque indice correspond à une certaine zone du ciel. Tout d'abord, la sphère est découpée en *base resolution pixels*, des pixels de base qui correspondent aux zones grisées sur la figure 7. Il y en a quatre de centres situés au-dessus de l'équateur, quatre en-dessous et quatre sur l'équateur.



La résolution de la carte Healpix va être définie par le paramètre N_{side} qui correspond au nombre de pixels le long d'une interface entre deux pixels de base. Sur la figure 7, ce paramètre vaut 2. Pour accéder à la résolution correspondante en minute d'arc par pixel, on peut utiliser le module *healpy* de *python*. Pour $N_{side} = 2048$, ce qui correspond à la résolution des six cartes Planck, on a une résolution de $1.71'$ /pixel.

Il y a deux façons de faire correspondre indice du tableau 1D et portion de la sphère céleste : soit la méthode RING, en anneau, soit la méthode NESTED, en paquets. Ces deux méthodes sont illustrées sur les figures 8 - A et 8 - B ci-dessous. Le module *healpy* effectue la lecture de tels fichiers FITS, et les decode automatiquement en RING ou NESTED selon ce qui est renseigné dans les "headers".

De plus, *healpy* contient de multiples fonctions de visualisation qui affichent directement des cartes à partir des fichiers FITS : on obtient ainsi les six cartes en représentation de Mollweide de la page 1. Un autre avantage offert par l'utilisation de ce module : les "projecteurs". Ces objets sont définis autour d'un point du ciel donné en longitude et latitude, et permettent de récupérer des cartes de taille réglable autour du

FIGURE 7 - Deux pixels de base représentés sur une sphère de $N_{side} = 2$ - Hivon, E., *The Healpix primer*

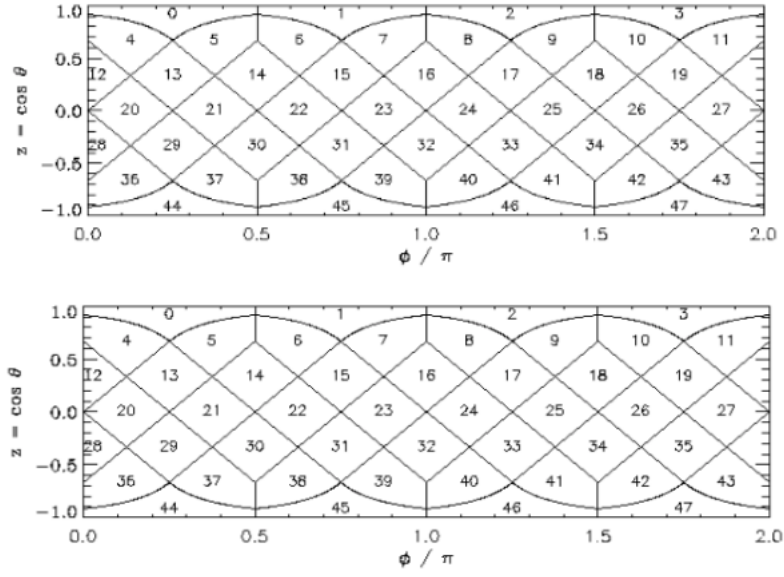


FIGURE 8 – A/ Représentation cylindrique de la sphère avec pixels codés en RING - B/ Pixels codés en NESTED - Hivon, E., *The Healpix primer*

point en question. C'est grâce à cette méthode de classe que j'ai pu créer les cartes du nuage Chamaeleon-Musca à chaque fréquence.

6.3 Projection gnomonique

Pour étudier Chamaeleon-Musca, j'ai effectué une projection gnomonique. La projection gnomonique, aussi appelée centrale, consiste en une représentation orthodromique d'une sphère à partir d'un de ses pôles. Cela signifie que chaque grand cercle (cercle sur la sphère de même rayon que la sphère) sera représenté par une droite, et réciproquement (Cf. [13]). METTRE PROJ GNOM CARTE HFI

J'ai utilisé le projecteur de *healpy* `healpy.projector.GnomonicProj` pour effectuer une projection gnomonique zoomée sur le point $(297.3^\circ, -15.5^\circ)$ de taille 50x50 en pixels. On peut par exemple trouver la projection gnomonique centrée de la carte d'intensité à 353 GHz sur la figure 4 - B.

Cet objet projecteur permet également de régler la résolution d'un pixel. Il nous faut alors prendre garde à satisfaire le critère de Nyquist, pour obtenir des cartes exploitables. Le lobe du satellite Planck a une taille de $5'$. A la même façon que le critère de Shannon, qui exige un échantillonnage de fréquence deux fois supérieure à la fréquence maximale du signal étudié, à savoir un pas temporel au moins deux fois plus petit que la période la plus fine contenue dans le signal, le critère de Nyquist stipule que la résolution du pixel doit être au moins deux fois plus petite que la taille de l'instrument. Ainsi, j'ai choisi ici une résolution de $1.5'$ par pixel. Dans tous les cas, la résolution est avant tout imposée par la résolution des cartes initiales, qui est de $1.7'$ et qui valide aussi le critère de Nyquist. Pour LiteBird, le lobe est de taille $18'$: la résolution est toujours adéquate.

6.4 Ajout des signaux free-free et synchrotron

Dans l'optique de retrouver les cartes 'brutes', j'ai ajouté les signaux synchrotron et free-free à mes cartes reconstituées à partir des données officielles en (β, τ_0, T) . Le synchrotron est un rayonnement de particules

Ecart relatif entre le signal reconstitué avec les bruits et le signal Planck à 353 GHz, en %

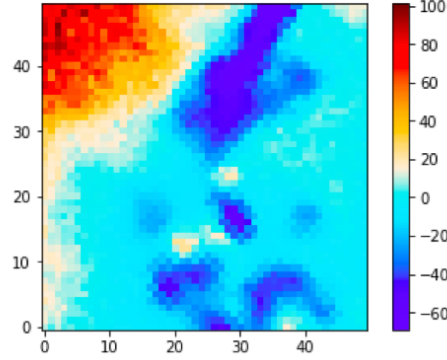


FIGURE 9 – Ecart relatif entre la carte reconstituée avec bruits et la carte Planck à 353 GHz, en %

chargées, dont les trajectoires sont courbées lorsqu'elles traversent des zones de fort champ magnétique, ce qui résulte en une émission électromagnétique. Le free-free, ou Bremsstrahlung thermique, correspond à l'émission des électrons libres se déplaçant dans des gaz, et qui sont déviés lorsqu'ils rencontrent des ions et noyaux.

Concentrons-nous sur le canal 353 GHz comme dans la partie 2.4. J'ai récupéré les deux cartes supplémentaires, que j'ai projetées gnomoniquement sur Chamaeleon-Musca, et que j'ai converties en unités SI pour 353 GHz. Cependant, à la différence de ce que j'ai fait dans la partie 2.4, je n'avais pas de coefficients de conversion explicites pour ces cartes dans la littérature que j'ai parcourue. Pour combler ce manque, j'ai mis en place un algorithme de minimisation pour explorer les coefficients de pondération à placer devant les cartes free-free et synchrotron de manière à minimiser la moyenne de l'écart-relatif entre les cartes $\{signal\ reconstitue + convolution + free - free + synchrotron\}$ et la carte Planck 'brute'.

Malgré tout, nous obtenons avec cette carte des écarts-relatifs encore plus important que sans les signaux free-free et synchrotron (jusqu'à 100%). Une piste à explorer pour prolonger cette étude serait d'abord de dissiper le flou quant à la présence du signal du fond diffus dans les signaux Planck ou non, et donc quant aux conversions à effectuer sur ces signaux. Par suite, j'avais mis en place une recherche des coefficients par minimisation de la moyenne des écarts-types de la carte en figure 9; une meilleure démarche serait de minimiser le maximum des écarts-types.

6.5 Cartes de température, indice spectral et profondeur optique

En dernier lieu, j'ai essayé de concilier les cartes Planck et le signal du fond diffus en pondérant ce signal par un coefficient k . J'ai cherché ce coefficient de manière à minimiser par la suite l'écart entre les cartes d'indice spectral, de profondeur optique et de température officielles et celles obtenues par interpolation des données Planck 'brutes'. J'ai obtenu : $k \simeq 2.5 \times 10^{-4}$. Observons cela sur la figure 10 : les résultats sont plus concluants que précédemment. On retrouve pour chaque carte les mêmes intervalles de variation des grandeurs, et les allures sont à peu près cohérentes. Ce n'est malgré tout pas optimal : une étude plus approfondie serait à faire.

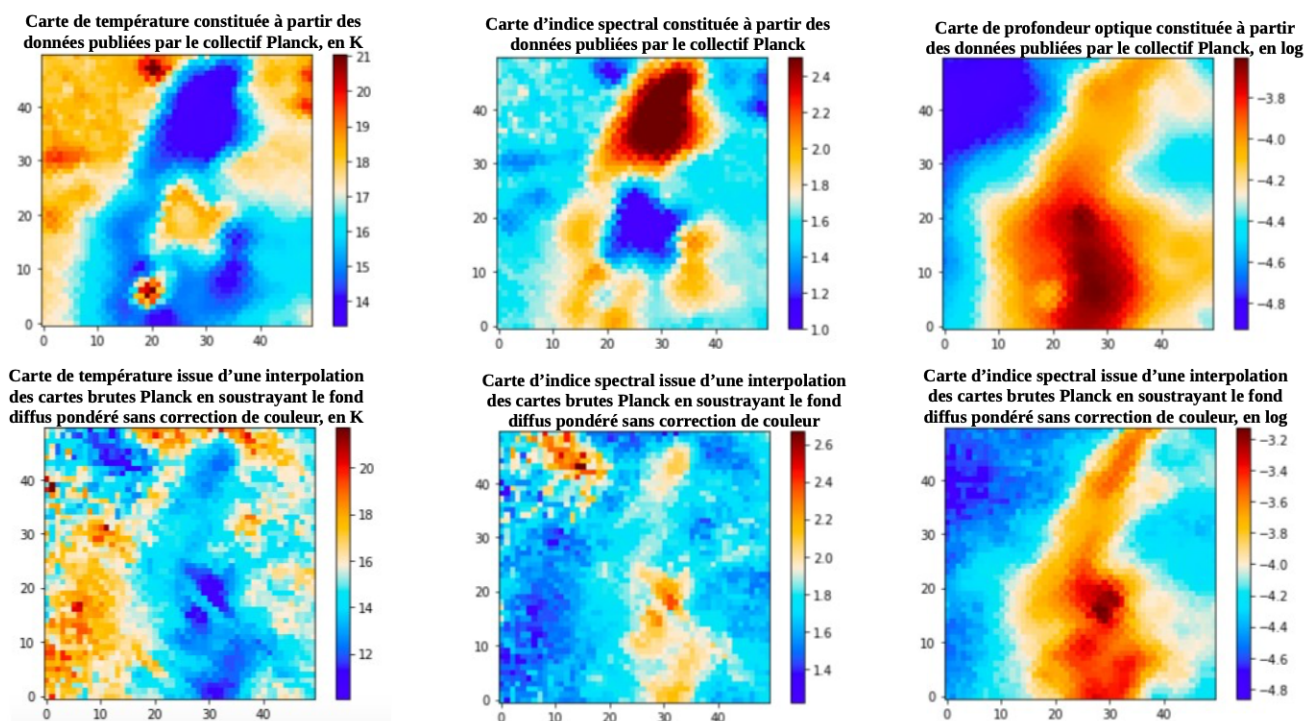


FIGURE 10 – En haut : cartes officielles publiées par le collectif Planck - En bas : cartes de température (en K), indice spectral et profondeur optique obtenues par interpolation des données brutes Planck moins le fond diffus pondéré