

Magistère des Sciences de la Matière

École Normale Supérieure de Lyon
Université Claude Bernard Lyon1

STAGE 2000-2001

LANCIEN Pierre
1^{ère} année
Option physique

Mesure du spectre de puissance des fluctuations du fond diffus cosmologique dans le cadre de l'expérience ARCHEOPS

Archeops est une expérience conçue pour mesurer les fluctuations du fond diffus cosmologique, rayonnement émis par l'Univers 300 000 ans environ après le Big-Bang. L'étude de ce rayonnement fossile donne accès aux paramètres cosmologiques de notre Univers : courbure, densité de matière et d'énergie... Un développement en harmoniques sphériques permet d'obtenir à partir de sa carte un spectre de puissance. De nombreux effets parasites faussent cependant ce spectre : les opérations de filtrage, le fait que seule une portion du ciel est balayée, et surtout le niveau de bruit, inévitablement important pour des mesures nécessitant une telle précision. J'ai contribué pendant ce stage à la mise au point d'un simulateur de type Monte-Carlo corrigeant ces effets, et plus particulièrement à l'estimation du spectre de bruit.

Mot-clés : cosmologie, rayonnement fossile, fluctuations,
spectre de puissance, traitement du signal

Physique Corpusculaire et Cosmologie, CNRS-IN2P3
Collège de France
<http://cdfinfo.in2p3.fr>

Sous la direction d'Alexandre Amblard et Jean-Christophe Hamilton

Stage effectué du 05/06/2001 au 02/08/2001

Table des matières

Les parties I. II. et III. présentent le cadre théorique et les éléments nécessaires à la compréhension de mon travail. La partie IV. expose l'ensemble du projet réalisé durant mon stage par l'équipe dont je faisais partie, et les parties V. et VI. traitent des aspects qui m'ont plus spécifiquement concerné.

I. Présentation : le CMB	
• La découverte du fond diffus cosmologique	5
• CMB et théorie du Big-Bang	5
• La mesure du CMB	5
• Origine des fluctuations	7
II. L'expérience ARCHEOPS	
• Présentation de l'instrument	8
• Les vols effectués ou prévus	9
III. Modes de représentation des données	
• La 'timeline'	10
• Le spectre de puissance temporel	11
• La carte	12
• Le spectre de puissance angulaire	13
IV. Estimation des CI grâce au simulateur MASTER	
• Motivations	14
• Formule de correction	15
• Estimation des différents termes	15
• Efficacité des corrections	17
V. Fit d'un spectre de bruit en $1/f$	
• Nécessité d'un nouveau programme de fit	18
• Test des programmes d'origine	18
• Analyse du code	19
• Corrections apportées	19
• Test de la nouvelle version	20
• Introduction de pics dans le spectre	20
VI. Estimation du bruit dans les données ARCHEOPS	
• Présentation du problème	21
• Modifications du simulateur	21
• Lien entre les deux spectres de puissance	22
• Proposition pour l'estimateur de bruit	23
• Application aux données	23
Conclusion	24
Annexes	25
Références et bibliographie	30

I. Présentation: Le CMB

La découverte du fond diffus cosmologique

En 1965, Penzias et Wilson, mettant au point une antenne radio de communication, notent l'omniprésence d'un bruit de fond, dans toutes les directions d'observation, comme si l'antenne était trop chaude de trois degrés. Leur talent d'expérimentateurs leur permet pourtant de conclure que l'origine de ce bruit est en fait un rayonnement provenant du ciel lui-même. Ils apprennent alors que son existence avait été prédite par Gamow, Alpher et Herman [1,2], en étudiant sur le plan théorique comment l'Univers avait évolué depuis la phase terriblement chaude et condensée du Big-Bang. C'est ainsi que la découverte du "Fond Diffus Cosmologique" ou CMB (Cosmic Microwave Background), devient un des arguments les plus probants en faveur du modèle du Big-Bang, et vaut à Penzias et Wilson le prix Nobel [3].

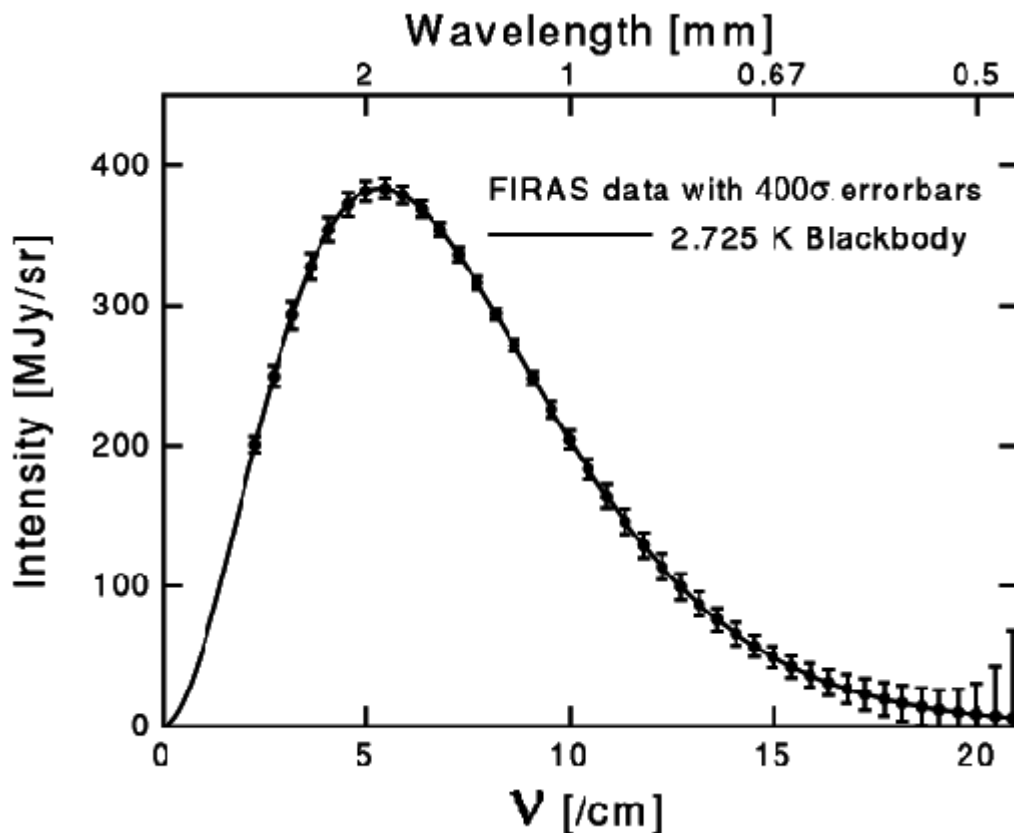
CMB et théorie du Big-Bang

Dans le cadre de la théorie du Big-Bang, l'Univers s'est trouvé il y a environ 10 à 15 milliards d'années dans une phase très chaude et très dense, à tel point que la densité moyenne d'énergie ne permettait pas que les électrons se combinent aux protons. La matière était alors totalement ionisée, ce qui réduisait considérablement le libre parcours moyen des photons, et induisait un fort couplage matière-rayonnement, rendant l'Univers "opaque". Cependant, l'expansion et le refroidissement prédits par le Big-Bang ont fait passer l'Univers par une période où la diminution de la température et de la densité d'énergie a permis aux électrons et aux protons de se combiner. Les photons interagirent ainsi beaucoup moins avec la matière, et l'Univers devint transparent. Ce moment, appelé recombinaison ou découplage matière-rayonnement a lieu environ 500000 ans après le Big-Bang, et les photons libérés sont à l'origine du fond diffus cosmologique qui nous parvient aujourd'hui (et qui est pour cette raison également nommé rayonnement fossile). L'intérêt de ce rayonnement est que les photons qui le constituent arrivent jusqu'à nous sans avoir été altérés, nous donnant ainsi une image très pure de l'Univers primordial.

La mesure du CMB

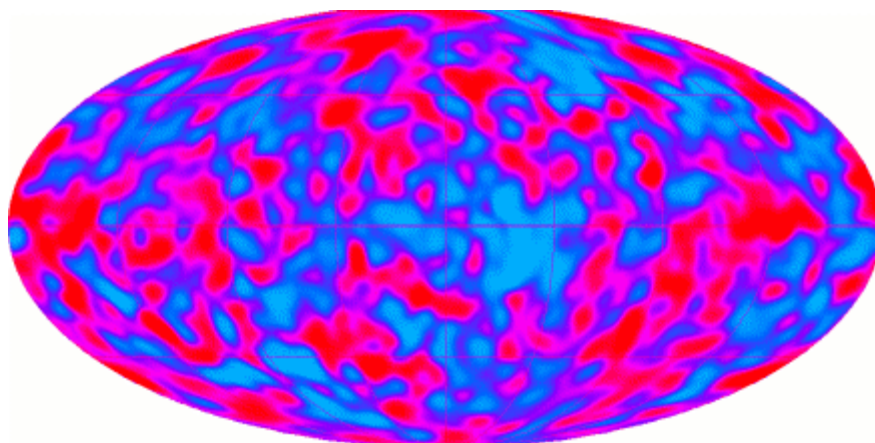
L'Univers étant en équilibre thermique au moment de la recombinaison, la théorie prévoit que le spectre du CMB est celui d'un corps noir. L'importance de la vérification de cette prédiction pour valider la théorie du Big-Bang a conduit les astrophysiciens à des mesures de plus en plus perfectionnées dans les trente années suivant la découverte par Penzias et Wilson. Il s'agissait également de vérifier le principe cosmologique, stipulant que l'Univers est homogène et isotrope à grande échelle, ce qui se traduit par une température uniforme du CMB sur la sphère céleste. Pour confirmer ces hypothèses, une précision bien plus grande que celle de la mesure de Penzias et Wilson était requise. Au début des années 1990, grâce au satellite COBE, le

Le spectre du CMB a été très bien mesuré dans plusieurs longueurs d'ondes. L'accord avec un spectre de corps noir à 2,73 Kelvins était parfait :



Points du spectre du CMB mesurés par COBE, superposés à un modèle de corps noir. Noter la taille minuscule des barres d'erreur à 400 σ !

Mieux encore, COBE combla les attentes des cosmologistes en produisant notamment la première carte du CMB :



Carte des fluctuations du CMB, par le satellite COBE

Sur cette image, les fluctuations relatives sont de l'ordre de 1/100000. On peut donc dire que le principe cosmologique n'est pas mis en défaut : la température de l'Univers à ses débuts était extrêmement homogène. De plus, la présence de fluctuations est normale et nécessaire pour expliquer la formation de structures de matière telles que les galaxies. Les modèles inflationnaires prévoient même les

caractéristiques de ces fluctuations, en fonction des paramètres cosmologiques, qui sont des constantes caractérisant notre Univers dans son ensemble et dans ses grandes lignes: densité de matière (et donc courbure de l'espace-temps), densité d'énergie, etc... L'obtention d'une carte la plus détaillée possible du CMB est donc primordiale pour répondre à nos grandes interrogations au sujet de l'Univers.

Origine des fluctuations

Les fluctuations du CMB témoignent de l'existence de faibles hétérogénéités dans la matière ionisée au moment de la recombinaison. Ces hétérogénéités ont entraîné des effondrements gravitationnels, créant plus tard les grandes structures de l'Univers. Cependant, pour que la gravitation, dont l'effet n'est pas instantané mais subordonné à la vitesse finie de toute information (qui est la vitesse de la lumière), s'exerce d'un point A sur un point B, il est nécessaire que B soit dans l'horizon de A (c'est à dire qu'un rayon lumineux aurait eu le temps d'aller de A à B depuis le Big-Bang). Ainsi, les petites fluctuations de densité sont les premières à rentrer dans l'horizon et à pouvoir s'effondrer gravitationnellement. Cependant avant le découplage lumière-matière, les photons exercent une forte pression de radiation, et la compétition des deux forces génère des oscillations acoustiques, créant des zones de compression et d'autres de dépression. La disparition de la pression de radiation, due au découplage, interrompt les oscillations, et les photons émis à ce moment gardent la trace de ces différentes zones.

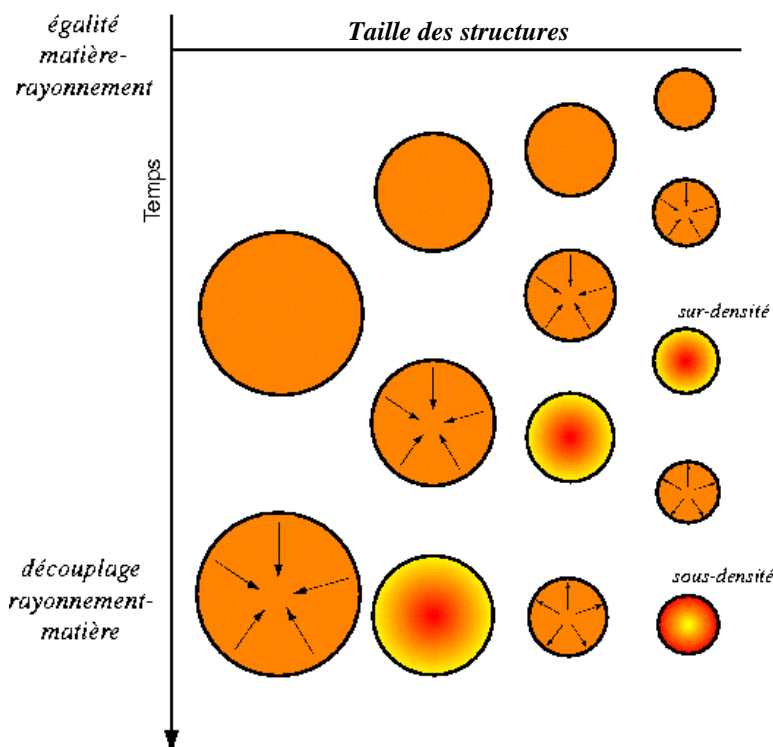


Schéma expliquant qualitativement l'origine de tailles privilégiées pour les fluctuations du CMB : après l'égalité matière-rayonnement, les oscillations acoustiques agissent sur des structures de taille de plus en plus grande. Les structures de différentes tailles n'oscillent pas en phase. Nous recevons le rayonnement émis au moment du découplage.

On s'attend donc à voir des différences dans les fluctuations du CMB, suivant leur taille caractéristique. Pour avoir accès à toute l'information véhiculée par le CMB, il faut donc obtenir une carte avec une bonne résolution spatiale, mais surtout une très bonne résolution en température, et analyser la distribution des structures suivant leur échelle grâce à une analyse par harmoniques sphériques.

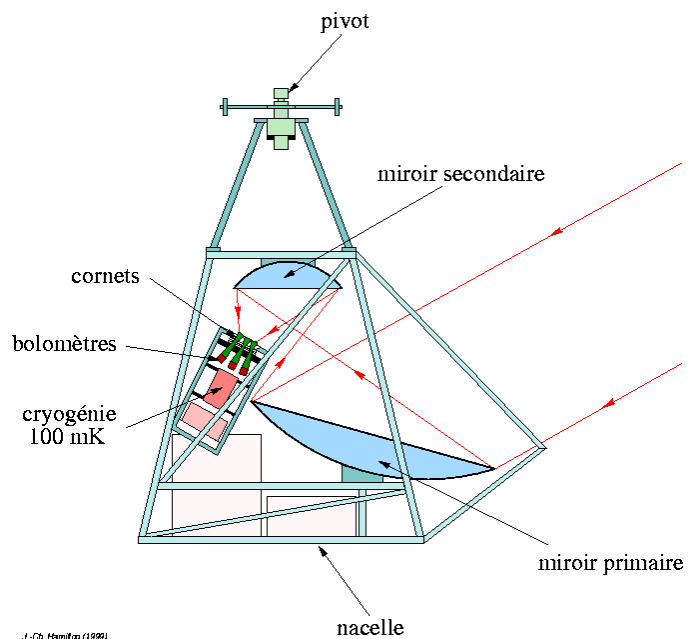
II. L'expérience ARCHEOPS

Présentation de l'instrument

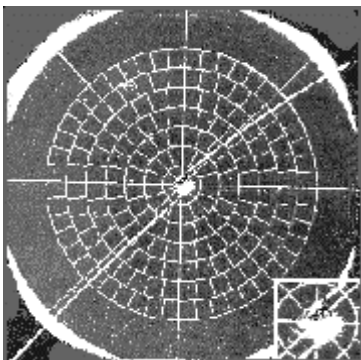
La difficulté d'observer les fluctuations du CMB rend nécessaire la réduction de tous les signaux parasites lors de leur mesure. Ceux-ci peuvent provenir des détecteurs eux-mêmes tout comme du ciel (nuages, atmosphère, etc.). Ainsi, l'observation par satellite est recommandée pour ce type de mesures, afin de s'affranchir de la plupart des effets parasites optiques. Ce sera le cas de Planck (ESA), normalement lancé en 2007, ou de MAP (NASA), lancé pendant mon stage. Une autre solution, beaucoup moins onéreuse, consiste à lancer un télescope fixé à un ballon stratosphérique, à une altitude raisonnable (plus de 30 km) pour que les effets de l'atmosphère permettent une mesure de qualité. BOOMERanG (voir résultats p.13) est une expérience de ce type.

C'est cette solution qui a été retenue pour ARCHEOPS: un télescope de 1,50m de diamètre, dans une nacelle accrochée à un ballon, pointe vers le ciel avec une inclinaison d'une quarantaine de degrés par rapport à l'horizon. Il balaie le ciel en suivant des cercles, grâce à une rotation de la nacelle avec une période de trente secondes environ. Combiné à la rotation de la Terre, qui fait se déplacer le centre de ces cercles, cela permet un balayage d'une grande portion du ciel (un quart) en plusieurs heures (voir fig. p.12)

La lumière collectée par le télescope est envoyée à un ensemble de bolomètres, sensibles à des longueurs d'ondes différentes, dont le principe est de mesurer l'élévation de la température d'un petit cristal lorsqu'il reçoit du rayonnement. Vu l'extrême précision requise par cette mesure, il est obligatoire de réduire au maximum le bruit instrumental. Le système de mesure est donc refroidi grâce à un cryostat de très grande qualité, qui devrait également équiper Planck. Ceci permet d'atteindre une température de l'ordre de 100 mK.



Coupe de la nacelle d'Archeops



Vue d'un bolomètre.

La forme en toile d'araignée permet un temps de réponse plus rapide, et réduit la probabilité de détection d'un rayon cosmique (dont la taille caractéristique est inférieure à une maille)

Avantages d'ARCHEOPS :

L'expérience ARCHEOPS est la première expérience de mesure du CMB dans laquelle la France est fortement impliquée. En effet, de nombreuses missions (une vingtaine), ont déjà eu lieu pour cartographier le CMB, menées principalement par les américains. L'ESA (agence spatiale européenne), désirant passer au premier plan, prépare pour 2007 la mission Planck, de grande ampleur, et qui devrait clore les mesures des fluctuations du CMB en apportant les arguments décisifs. Les équipes françaises avaient donc tout intérêt à programmer une mission si peu coûteuse qu'ARCHEOPS, sous l'égide du CNES. D'autre part, les atouts d'ARCHEOPS sont de taille: la couverture du ciel est très bonne (un quart), comparée à des missions comme Boomerang ou Maxima (quelques pour cents). De plus, la résolution de l'instrument est bien meilleure que celle des autres missions bénéficiant d'une couverture du ciel de même ampleur. Du côté des résultats espérés, ARCHEOPS devrait donc préciser les mesures pour les fluctuations à grande échelle du CMB, c'est à dire aux alentours de 5° . Enfin, les mesures de la polarisation du rayonnement sont rendues possibles par la présence de filtres polarisants devant certains bolomètres, permettant d'accéder à une caractéristique supplémentaire des objets d'avant plan (principalement la poussière galactique), que l'on commence tout juste à mesurer.

ARCHEOPS est donc une mission à petit budget mais complète, et dotée d'une instrumentation de très grande qualité. Elle permet à la France d'entrer de plain pied dans la course à la mesure des fluctuations du CMB amorcée depuis une dizaine d'années.

Les vols effectués ou prévus :

Deux vols ont eu lieu pour le moment:

- Juillet 1999: Trapani (Sicile)
Ce vol de test a permis de montrer le bon fonctionnement de l'instrument et a mis en évidence certains défauts à régler. Il a permis notamment une meilleure connaissance des effets systématiques. Le ballon a volé trois heures à une altitude de 40 km.
- Janvier 2001: Kiruna (Suède)
Premier vol scientifique, qui a permis d'enregistrer sept heures de données. Le vol était prévu pour durer 24h, en profitant de la nuit polaire. Malheureusement, le vent étant trop fort, le ballon a dérivé trop vite, survolant le territoire russe. Il fallut donc le faire atterrir. Le temps de vol et la qualité des mesures ont été cependant bien améliorés par rapport à Trapani.

Un vol est également prévu en décembre 2001 à Kiruna, suivi d'un deuxième en janvier 2002 si les conditions météo le permettent. Ces mesures seront essentielles pour l'obtention de bons résultats, car il est nécessaire d'augmenter la durée de vol à 24h: le ballon passerait ainsi de nombreuses fois au même endroit, réduisant considérablement l'importance des systématiques.

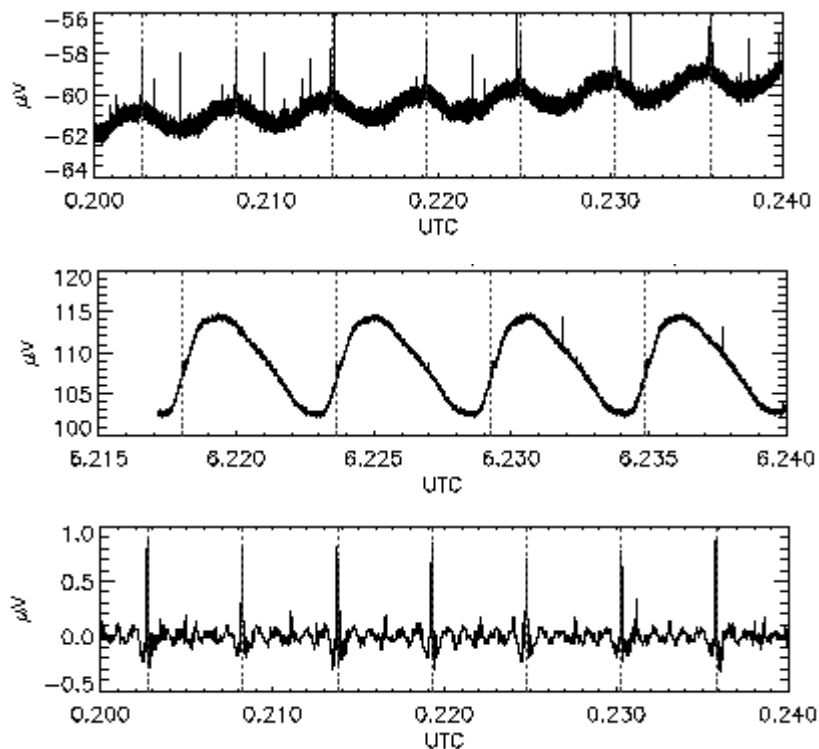
III. Modes de représentation des données pour l'analyse

Mon travail a consisté à analyser les données des vols effectués. Voici les différentes formes de représentations que j'ai utilisées pour leur visualisation et leur traitement :

La 'timeline' :

C'est l'élément de départ de toute l'analyse des données : il s'agit tout simplement de l'intensité du rayonnement, mesurée par un bolomètre, en fonction du temps. La timeline de base est en μV , mais elle peut être convertie en μK (micro Kelvins), grâce une opération appelée calibration, qui s'avère délicate du fait de divergences selon les étalons choisis (Jupiter, dipôle, galaxie...).

La 'timeline' brute présente de nombreux pics (les 'glitches') dus aux rayons cosmiques, qui se détachent nettement du reste, ainsi que des dérives basses fréquences, par exemple l'effet du réchauffement des instruments pendant le vol. Il est possible de nettoyer une partie importante de ces bruits parasites et des systématiques en supprimant les pics et en filtrant les basses fréquences grâce à une FFT. Il reste tout de même , après ce traitement, un niveau de bruit très important, que l'on va pouvoir quantifier avec le spectre de puissance temporel.



*Extraits de timeline : en début de vol (en haut), en fin de vol (milieu).
Noter l'importance des dérives basse fréquence, bien éliminées sur la timeline
de début de vol après filtrage (bas). Les grands pics résiduels sont dus à la galaxie.*

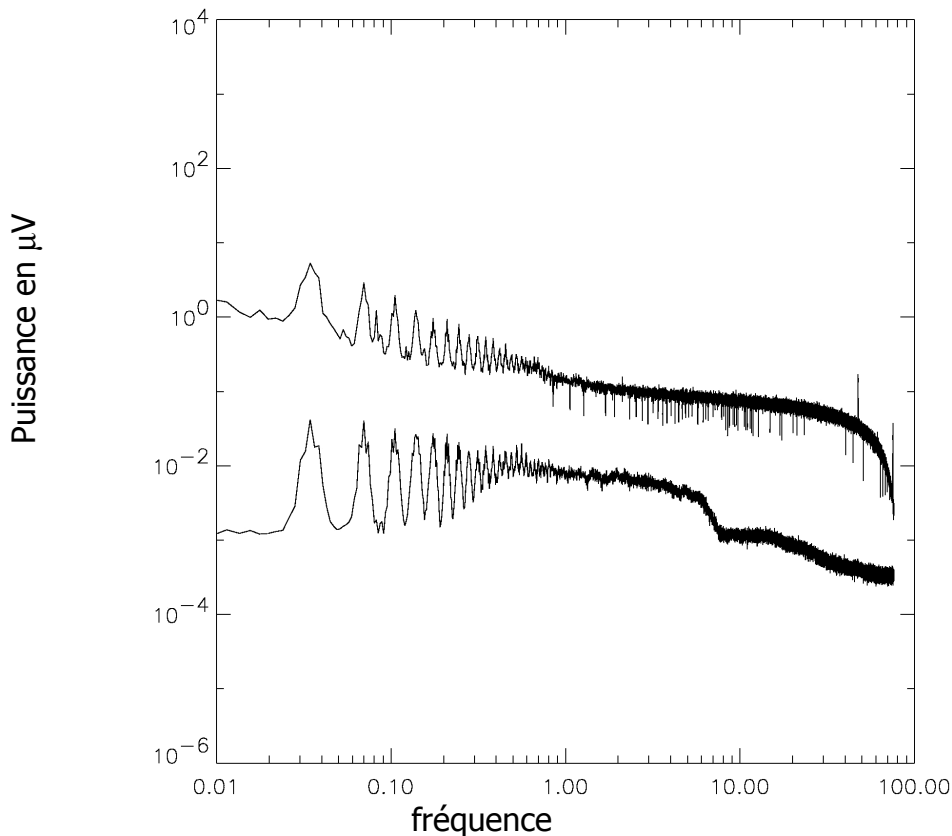
Le spectre de puissance temporel :

C'est le module au carré de la transformée de Fourier des données. Sa forme approximative est du $1/f$, c'est à dire de la forme :

$$P(f) = \sigma_0^2 \left[1 + \left(\frac{f}{f_{knee}} \right)^\alpha \right]$$

Où σ_0 est le niveau de bruit blanc, dominant à hautes fréquences, et α donne le comportement à basses fréquences. f_{knee} est la fréquence caractéristique, transition entre les domaines de dominance de α et σ_0 .

Pour l'analyse des données des vols de Trapani et Kiruna, on constate que le bruit est très important, malgré toutes les précautions prises et la qualité de l'instrumentation utilisée : j'ai simulé un spectre pour des données de CMB seulement pour le comparer au spectre mesuré.



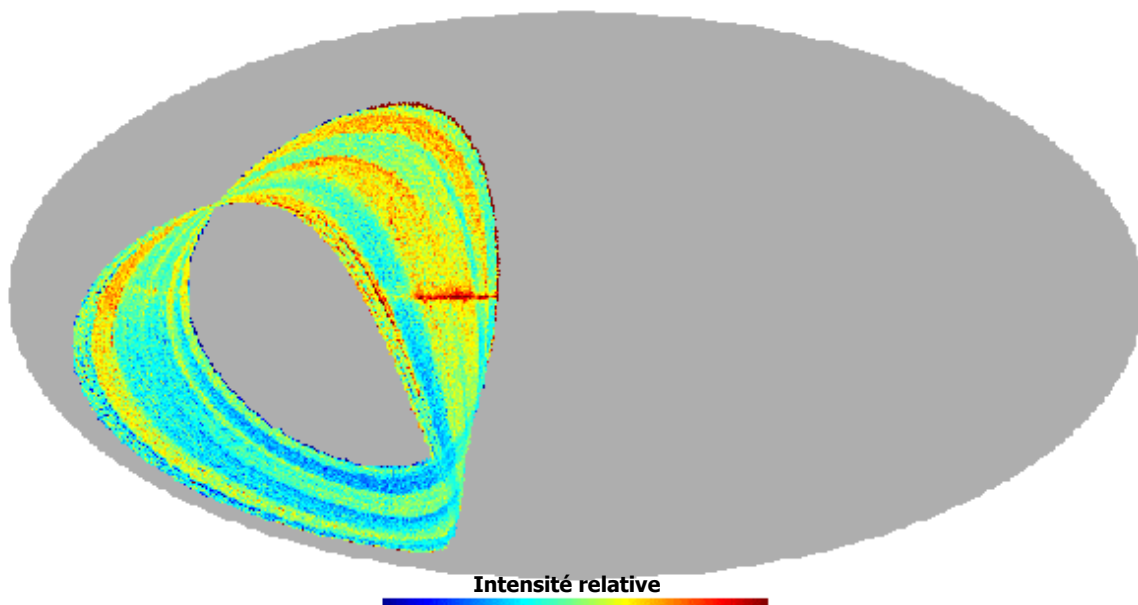
Spectres de puissance en coordonnées log-log: CMB (en bas) et données (en haut)

On voit à quel point la mesure des fluctuations du CMB requiert une précision importante. L'expérience Archeops a pour le moment un rapport signal sur bruit de 1/100 sur le spectre de puissance. Il est donc primordial, pour pouvoir soustraire ce bruit lors de l'analyse des données, de le connaître, c'est à dire de le modéliser, de façon très précise.

La carte :

A partir de la 'timeline' et des données de pointage du télescope, on peut construire une carte du CMB. Cette opération est cependant bien moins simple qu'il n'y paraît : il faut choisir une pixelisation appropriée et prendre en compte pour chaque pixel tous les points de la 'timeline' qui y tombent. En effet, le télescope pointe plusieurs fois vers les mêmes points, ce qui permet de réduire l'importance du bruit.

La manière la plus simple de procéder est d'affecter à un pixel la moyenne des valeurs des mesures qui y tombent : le bruit tend à se moyennner à 0, contrairement au signal. L'inconvénient est qu'on ne tient pas compte des corrélations dans le bruit, qui permettraient de minimiser mieux encore son effet. Dans le cas le plus général, on formalise ce passage d'une 'timeline' de n éléments à une carte de m pixels grâce à une multiplication par une matrice $n \times m$. Il existe ensuite pour cette matrice une formule analytique pour obtenir la carte optimale (qui maximise la fonction de vraisemblance). Le problème reste alors d'effectuer le calcul pour des 'timelines' de plus de 3 millions d'éléments (il y a notamment besoin d'inverser de telles matrices...). Le temps CPU requis pour traiter les données ainsi est rédhibitoire ; on est forcés d'utiliser des approximations, par exemple par des méthodes itératives (Doré et al. [4]).



Carte des données Archeops, non nettoyées.

L'ellipse représente une projection de l'ensemble de la voûte céleste, en coordonnées galactiques.

Noter la présence de la galaxie sur l'axe horizontal.

Sur cette carte, où les données ne couvrent pas toute la voûte céleste, il est possible de visualiser l'importance de la stratégie de scan, qui est la manière dont le télescope balaye le ciel. Pour Archeops, ce sont des cercles qui sont tracés, et dont le centre se déplace avec la rotation de la Terre. Pour l'expérience BOOMERanG, la portion du ciel couverte était de quelques pour cents seulement, mais le télescope passait de très nombreuses fois au même point, et cette stratégie de scan s'est révélée très fructueuse du fait de la minimisation du bruit qu'elle a permis. Archeops vise des échelles plus larges sur la voûte céleste, mais c'est aussi sa faiblesse, à cause du niveau de bruit élevé que cela entraîne.

Le spectre de puissance angulaire :

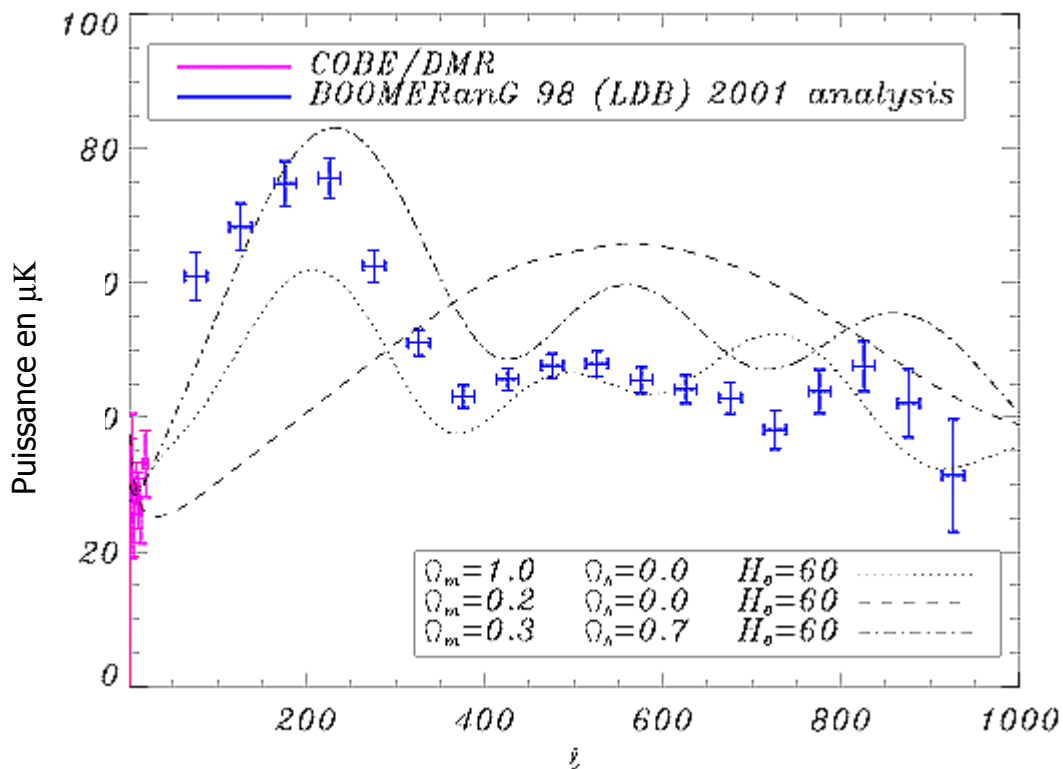
But de toute la chaîne de traitement des données d'Archeops, le spectre de puissance angulaire est ce qui permet de relier les mesures à la cosmologie, et de trancher en faveur de tel ou tel modèle d'Univers, si les mesures sont de qualité. Il donne les importances relatives des structures du CMB en fonction de leur échelle.

On l'obtient à partir de la carte, qu'on développe en harmoniques sphériques. Comme pour la transformée de Fourier, cette méthode s'appuie sur la décomposition d'une fonction dans une base de fonctions orthogonales. Ici, la fonction à décomposer est la carte du CMB obtenue à partir des données des bolomètres nettoyées et de la reconstructions du pointage (notre espace d'analyse est la sphère céleste entière). La base des harmoniques sphériques se compose de fonctions multipolaires Y_ℓ^m , ou ℓ est inversement proportionnel à la taille caractéristique des structures du multipôle, et m lié à la direction de l'origine du multipôle.

$$\frac{\Delta T(\theta, \varphi)}{T_{moy}} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_\ell^m(\theta, \varphi)$$

Dans l'analyse des fluctuations du CMB, en vertu du principe cosmologique, la direction du multipôle n'a pas d'importance (homogénéité et isotropie des fluctuations). On se contentera donc des coefficients C_ℓ .

$$C_\ell = \langle |a_{\ell m}|^2 \rangle = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$$



Spectre de puissance du CMB (coefficients C_ℓ en fonction de ℓ), avec les données de COBE et BOOMERanG, ainsi que les courbes théoriques de trois modèles différents.

IV. Estimation des Cl grâce au simulateur MASTER

Venons en à présent à la méthode utilisée pour l'analyse des données, et qui a fait l'objet d'un projet mené à bien au cours de mes deux mois de stage par l'équipe dont j'ai fait partie.

Motivations :

L'analyse en harmoniques sphériques doit, en théorie, être appliquée à la voûte céleste entière pour donner des résultats valables. Avec l'expérience Archeops, la couverture du ciel est incomplète, ce qui biaise fortement la mesure des Cl : on doit dans l'analyse remplacer les points de la carte non observés par la valeur nulle, pour permettre le calcul. Non seulement la puissance totale de chaque structure est sous-estimée d'un facteur proportionnel à la portion du ciel non couverte, mais le ciel est ainsi convolué par une fonction fenêtre, ce qui limite les modes auxquels on a accès.

Les opérations de filtrage appliquées aux données ne laissent pas non plus identiques les Cl. Enfin, le bruit, ici bien plus important que le CMB, fausse toute la mesure.

Il faut donc absolument apporter des corrections aux 'pseudo' Cl calculés. Pour cela, il faudrait en théorie calculer le biais introduit par chaque facteur dans les pseudo Cl grâce aux formules analytiques de l'analyse en harmoniques sphériques, de la fonction fenêtre de la couverture du ciel, et de l'estimation du bruit. Dans la pratique, il est complètement impossible d'effectuer ce calcul, ainsi que de séparer le bruit du signal recherché.

Sous la direction de Jean-Christophe Hamilton, nous avons donc opté au PCC pour une méthode de type Monte-Carlo nommée MASTER, exposée dans un article récent (Hivon et al., [5]) : au lieu de calculer théoriquement la correction à apporter, nous la mesurons grâce à des simulations. Nous avons constitué une petite équipe de quatre personnes (A. Amblard, M. Douspis, J.C. Hamilton, P. Lancien), afin de nous lancer dans la réalisation d'un simulateur reproduisant toutes les étapes du traitement de données. Nous partons d'un modèle de CMB défini par la donnée des Cl. Il y a une infinité de cartes de CMB possibles qui correspondraient à ce modèle. Nous en générons une en supposant les fluctuations du CMB gaussiennes (ce qui est couramment fait), et c'est ce que nous appellerons réalisation. De même avec un modèle de bruit : nous partons d'un spectre de puissance temporel donné, et pour repasser dans l'espace réel nous prenons des phases aléatoires, et nous générons une réalisation de bruit, en le supposant gaussien. Le générateur de bruit sera plus précisément présenté dans la partie suivante.

Nous introduisons donc en entrée de notre simulateur une réalisation de CMB et une réalisation de bruit simulés à partir de modèles, et nous leur appliquons le même traitement qu'aux données (filtrage et projection sur le ciel), avec la même stratégie de scan, pour obtenir des pseudo Cl. Il est ensuite possible de comparer les pseudo Cl simulés aux Cl des modèles. Un nombre suffisant de simulations de ce type permet une analyse statistique, afin d'en tirer la correction moyenne à appliquer, ainsi que les barres d'erreurs correspondantes.

Formule de correction :

Il est possible de formaliser l'opération de correction des pseudo Cl par des opérations vectorielles et matricielles, en distinguant l'influence de chaque terme.

On part du cas le plus général, où chaque pseudo Cl peut dépendre de tous les Cl non biaisés, avec des coefficients différents. Ceci se comprend bien si l'on considère que la fonction fenêtre induite par une couverture du ciel incomplète va mélanger les modes par convolution. En outre, les pseudo Cl dépendent fortement du bruit. La linéarité dans l'analyse en harmoniques sphériques permet d'écrire :

$$\langle \tilde{C}_\ell \rangle = \sum_{\ell'} a_{\ell\ell'} \langle C_{\ell'} \rangle + \langle \tilde{N}_\ell \rangle$$

$\langle \tilde{C}_\ell \rangle$ est la moyenne d'ensemble des pseudo Cl mesurés à partir des simulations.

$\langle C_{\ell'} \rangle$ se rapporte aux Cl moyens des modèles utilisés. Ce sont les Cl non biaisés.

$\langle \tilde{N}_\ell \rangle$ représente les Cl du bruit. Plus rigoureusement, ce sont des pseudo Cl car ils sont calculés sur le même plan que $\langle \tilde{C}_\ell \rangle$.

$a_{\ell\ell'}$ est un coefficient réel, que nous allons maintenant détailler : il doit contenir les effets du filtrage dans la FFT, de la fonction fenêtre appliquée pour une couverture du ciel incomplète, mais aussi du 'Beam', c'est à dire de la convolution des données due à la résolution finie du télescope (10' d'arc pour Archeops) et à la pixelisation de la carte. Nous avons vu que la fonction fenêtre était responsable d'un éventuel mélange des modes. C'est donc la correction qu'elle induit qui dépendra de ℓ et ℓ' . Les effets du filtrage et du Beam ne dépendent que de ℓ dans notre approximation. L'équation finale est alors :

$$\langle \tilde{C}_\ell \rangle = \sum_{\ell'} M_{\ell\ell'} F_{\ell'} B_{\ell'}^2 \langle C_{\ell'} \rangle + \langle \tilde{N}_\ell \rangle \quad (1)$$

$F_{\ell'}$ est l'effet du filtrage, $B_{\ell'}^2$ celui du beam, et $M_{\ell\ell'}$ celui de la fenêtre de couverture du ciel.

Estimation des différents termes :

L'équation (1) exprime aussi que l'estimation du bruit et celle des facteurs correctifs peuvent être traitées à part :

- Pour estimer les $\langle \tilde{N}_\ell \rangle$, on effectue des réalisations de bruit à partir d'un modèle concordant avec les données (trouver le modèle adéquat est en fait très délicat ; c'est cette opération qui a fait plus précisément l'objet de mon stage, et qui sera décrite dans les parties suivantes). Des cartes de bruit pur sans CMB sont dressées et les $\langle \tilde{N}_\ell \rangle$ sont alors calculés en faisant des moyennes sur de nombreuses réalisations. Cette simulation de type Monte-Carlo a l'énorme avantage de donner

aussi les barres d'erreurs (pour une distribution supposée gaussienne). Il est à noter que pour un bruit blanc, on attend des $\langle \tilde{N}_\ell \rangle$ constants. Le bruit dans l'expérience Archeops s'approche assez du bruit blanc, après filtrage du comportement 1/f des basses fréquences, et on peut prévoir un niveau à peu près constant pour les $\langle \tilde{N}_\ell \rangle$.

- Pour le terme du beam, $B_{\ell'}^2$, il est possible d'utiliser une formule analytique, dans l'hypothèse d'un beam gaussien (de largeur σ) et symétrique :

$$B_\ell = e^{-\ell(\ell+1)\sigma^2/2}$$

- Pour $M_{\ell\ell'}$, il existe également une forme analytique, plus complexe, dont la démonstration ne peut être abordée ici :

$$M_{\ell\ell'} = \frac{2\ell'+1}{4\pi} \sum_{\ell''} (2\ell''+1) W_{\ell''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2$$

Où $\begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}$ est le symbole de Wigner 3-j, et $W_{\ell''}$ est le spectre de puissance angulaire de la fonction fenêtre formée par la zone couverte de la sphère céleste. Pour vérifier que cette partie du simulateur était correcte, nous avons effectué des réalisations de CMB seul sans filtrage, afin d'obtenir l'équation simplifiée :

$$\langle \tilde{C}_\ell \rangle = \sum_{\ell'} M_{\ell\ell'} \langle C_{\ell'} \rangle$$

Nous avons ainsi vérifié sur des moyennes d'ensemble que l'accord était excellent. Mais l'application de cette formule est supposée aller au delà de cette vérification et nous donner les $\langle C_{\ell'} \rangle$ à partir des $\langle \tilde{C}_\ell \rangle$. Comme la couverture du ciel est incomplète, nous ne pouvons avoir accès à tous les modes, et cela entraîne la non inversibilité de la matrice M. Il faut accepter cette perte de modes et sous-échantillonner la matrice. Avec des largeurs de bin raisonnables (plus de 3 ou 4 modes par bin), M devient inversible et nous pouvons appliquer la correction grâce à sa matrice inverse. Nous n'avons de toute façon pas à regretter une perte d'information, car les bins nous permettent de réduire considérablement les barres d'erreurs. En pratique nous prenons des bins comprenant 10 à 50 modes. De plus, une modification que j'ai apportée au simulateur permet de prendre des bins variables pour optimiser notre mesure. On peut ainsi profiter de la bonne qualité des mesures à bas ℓ en prenant des bins de 10 et augmenter leur taille jusqu'à 50 ou plus en arrivant aux grands modes (qui correspondent à une petite échelle angulaire à laquelle le bruit domine).

- $F_{\ell'}$, terme introduit pour corriger le filtrage effectué (passe bande 1 Hz - 45 Hz pour Kiruna), dépend en théorie aussi de la stratégie de scan et du modèle de CMB. Nous avons supposé pour établir la formule générale de MASTER que le terme de filtrage ne dépendait que du mode ℓ . Cette approximation a été vérifiée en déterminant les

F_{ℓ} pour des modèles de CMB très différents : les écarts maximaux des courbes obtenues étaient inférieurs à quelques pour cents.

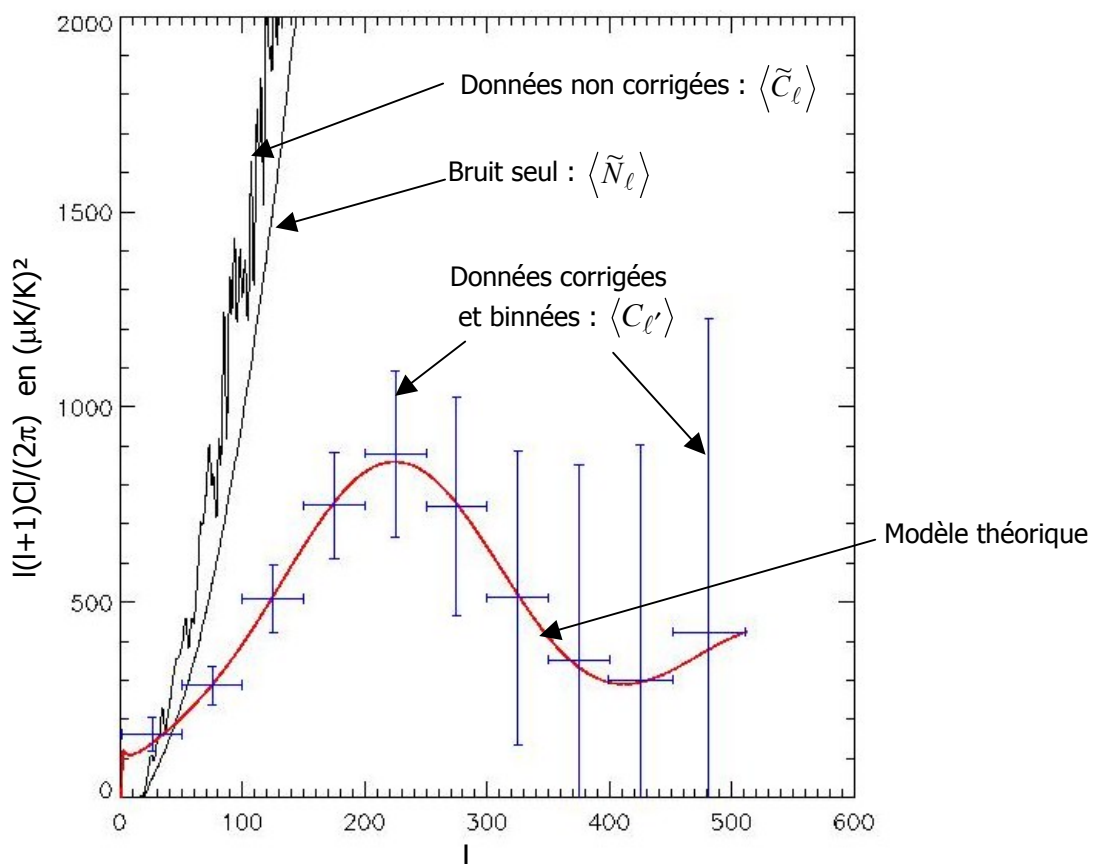
La détermination des F_{ℓ} intervient après celle de la matrice M , et on part toujours de réalisations de CMB seul. On a donc la formule :

$$\langle \tilde{C}_{\ell} \rangle = \sum_{\ell'} M_{\ell\ell'} F_{\ell'} \langle C_{\ell'} \rangle$$

Dans laquelle on connaît tous les termes sauf les F_{ℓ} . $M_{\ell\ell}$ n'est cependant pas inversible. Ce système d'équations est dégénéré : il y a une infinité de solutions. Une seule est cependant physiquement acceptable. Pour la trouver, on se sert d'un algorithme itératif, qui converge rapidement (une ou deux itérations) vers la solution.

Efficacité des corrections :

Après avoir fini le codage de toutes les parties composant le simulateur MASTER, nous avons testé son efficacité pour un vol comme celui de Kiruna. Nous avons simulé du CMB ainsi que du bruit en $1/f$ (avec des caractéristiques proches du bruit des données de Kiruna), et appliqué le filtrage passe-bande mentionné ci-dessus. Le graphe obtenu montre l'efficacité de notre correction, qui supprime tout biais.



Spectre de puissance simulé pour le vol de Kiruna (7 heures), avec un bruit parfaitement connu. (moyennes sur 200 réalisations)

V. Fit d'un spectre de bruit en 1/f

Cette première partie de mon stage a été le moyen de me familiariser avec mon environnement de travail et de prendre progressivement conscience de l'importance de l'estimation du bruit dans ARCHEOPS. J'ai dû apprendre le langage de programmation IDL, assez intuitif, et l'utiliser pour simuler du bruit en 1/f, puis le traiter.

Nécessité d'un nouveau programme de fit.

En essayant de faire tourner des programmes déjà existants au laboratoire pour simuler des 'timelines' de bruit en 1/f, obtenir leur spectre de puissance par transformée de Fourier, et en faire un fit, je me suis aperçu que l'écart entre les paramètres du 1/f en entrée et ceux retrouvés par le programme de fit était relativement important (de l'ordre de 10%). Était-ce dû au fait que la réalisation de bruit en 1/f s'écartait de la forme théorique ou, ce qui est plus grave, à une erreur systématique du programme de fit ? Il me fallait revoir les codes, rentrer dans les algorithmes, et faire de nombreuses simulations... Deux programmes rentraient en ligne de compte :

- L'un génère une 'timeline' de bruit gaussien (avec un sigma de 1), c'est à dire un signal dont l'intensité en chaque instant est tirée avec une statistique gaussienne. Il en prend le spectre de puissance et le multiplie au spectre de puissance théorique en 1/f de formule :

$$P(f) = \sigma^2 \left[1 + \left(\frac{f}{f_{knee}} \right)^\alpha \right]$$

En revenant dans l'espace réel, on obtient alors une 'timeline' de bruit gaussien en 1/f, qui peut être passée au programme de fit pour le tester.

- Le programme de fit repose sur la méthode des moindres carrés : il détermine les coefficients σ , α et f_{knee} pour lesquels la somme des carrés des différences entre le spectre interpolé et le spectre réel est minimale. Cette partie est une routine déjà implémentée dans IDL. Le programme de fit est cependant nécessaire pour préparer les données à subir cette routine : il faut choisir des bins astucieux pour réduire le nombre des points rentrant dans le calcul des moindres carrés, et il est possible de raffiner la méthode en affectant des poids aux bins, qu'on peut relier directement au sigma de la distribution supposée gaussienne des points à l'intérieur de chaque bin.

Test du programme d'origine :

Pour tester la précision et la justesse du programme de fit d'origine, j'ai développé un programme de test, `testfit.pro`, qui m'a beaucoup aidé lorsque, plus tard, j'ai eu à comparer plusieurs versions de fit. L'utilisateur rentre les trois coefficients de son bruit en 1/f théorique, le nombre de simulations à effectuer, et le(s) programme(s) de fit à utiliser. Le programme se charge alors de faire les réalisations de bruit gaussien en 1/f et leur applique le(s) fit(s). Il représente ensuite la dispersion de l'ensemble des résultats de fit pour les trois paramètres σ , α et f_{knee} , de forme gaussienne, et en donne les caractéristiques : le sigma donne la précision et la valeur centrale permet de juger de la justesse. Voici les résultats obtenus pour des valeurs d'entrée assez proches des caractéristiques du bruit des données :

	α	σ	f_{knee}
Valeurs réelles	-2.00	0.100	1.00
Valeur du fit	-1.86	0.098	1.12
Sigma	0.20	0.0019	0.18
Erreur systématique	7.0%	2.0%	12%
Incertitude	10.0%	1.9%	18%

Il y a donc une erreur systématique (qui reste toutefois inférieure à 1σ), et l'incertitude est beaucoup trop élevée pour permettre d'utiliser un tel fit avec les données Archeops, où il est nécessaire de connaître le bruit à mieux que 3 ou 4%. Il faut donc réduire grandement l'incertitude, ce qui impose aussi de diminuer le biais systématique.

Analyse du code :

Après analyse du code, j'ai d'abord remarqué un choix judicieux : comme la loi en $1/f$ se manifeste surtout à basse fréquences, et que la densité d'échantillons par tranche de fréquence est constante dans le spectre de puissance à traiter, les fréquences ont été binnées de manière logarithmique (en fait la loi en $1/f$ à une forme naturellement simplifiée en coordonnée log-log).

Cependant, j'ai également constaté certains choix arbitraires dans l'algorithme: la largeur des bins, les seuils à partir desquels on élimine les données, et le logarithme utilisé pour binner (log de 10)... De plus, les poids donnés à chaque bin pour la méthodes des moindres carrés étaient bien calculés en fonction de la dispersion des échantillons dans chaque bin, et ceci grâce au sigma d'une interpolation gaussienne, mais ce calcul ne tenait pas compte du nombre d'échantillon par bin : une gaussienne pouvait être interpolée à partir de 3 échantillons par bin seulement. Les bins contenant 1 ou 2 points n'étaient pas pris en compte. Il s'avère que tout cela joue à basse fréquence, car les échantillons par bin y sont peu nombreux (du fait du logarithme). Cela explique le fait que α et f_{knee} soient faussés, car c'est justement à basse fréquence que ces coefficients interviennent : de nombreux bins étaient rejetés, ne contenant qu'un ou deux points, et les bins de 3 ou 4 points avaient des sigmas non significatifs (il suffit d'une statistique plaçant ces points très proches, mais loin de la moyenne des autres, pour leur donner à tord un poids trop important).

Corrections apportées :

Après avoir essayé de modifier des détails du code, je me suis rendu compte que pour faire un programme plus cohérent, il était préférable de tout recoder. J'ai donc fait un nouveau programme, reprenant dans les grandes lignes le précédant, mais avec des modifications de taille :

- Tous les nombres arbitraires ont été passés en paramètres, ce qui permet à l'utilisateur de les choisir et de les changer plus facilement. Il est ainsi possible, par exemple, de décider du nombre de bins, ce qui permet de plus ou moins lisser le spectre avant d'y appliquer les moindres carrés.
- La valeur affectée à chaque bin n'est pas la moyenne des seuls points du bin, car elle inclut 3 ou 4 points de plus de chaque coté (ce nombre est paramétrable). Pour les bins de 100 éléments, cela ne change rien, mais cela permet aux bins de 1 à quelques points d'être plus représentatifs.

- Le poids de chaque bin dans les moindres carrés est déterminé par le sigma d'une gaussienne à partir des points d'une fenêtre glissante paramétrable (typiquement de 100 échantillons).
- Enfin, la routine a été accélérée d'un facteur 3 par rapport au programme d'origine. Ces modifications ont permis un meilleur lissage à basse fréquence, et une souplesse plus grande pour l'utilisateur grâce au choix des paramètres.

Test de la nouvelle version :

Vu que celle-ci demande à l'utilisateur de choisir des paramètres, j'ai modifié testfit.pro afin qu'il teste aussi plusieurs valeurs des paramètres, pour déterminer les meilleurs réglages. Il se trouve qu'ils sont optimaux pour 50 bins et une fenêtre glissante de 100 échantillons pour le calcul du poids dans les moindres carrés. La qualité du fit a été grandement améliorée par rapport au programme original, surtout pour α , paramètre important pour la modélisation basses fréquences, qui est connu avec une précision acceptable :

	α	σ	f_{knee}
Valeurs réelles	-2.00	0.10000	1.000
Valeur du fit	-1.98	0.10005	1.032
Sigma	0.04	0.00030	0.043
Erreur systématique	1.0%	0.05%	3.2%
Incertitude	2.0%	0.30%	4.3%

Introduction de pics supplémentaires...

Mon programme de fit semblait donc être au point lorsqu'en entrée on lui donnait du bruit en $1/f$ généré. Il était cependant destiné à être appliqué aux vraies données. Une étape intermédiaire a été de le tester sur un spectre en $1/f$ simulé sur lequel j'ai rajouté de nombreux pics allant jusqu'à 10 fois la valeur normale, et fabriqués à partir des vraies données (ces pics y sont nombreux, notamment en raison de l'électronique). Les résultats ont été bien moins bons, redonnant des erreurs de plus de 10%. J'ai du alors coder un nouveau programme permettant de nettoyer ces pics. La version la plus aboutie fait un premier fit très grossier avec les pics puis supprime tous les points qui en sont éloignés de plus de 5 sigma. Un deuxième fit est fait sur les points restant, meilleur que le premier, et tous les points éloignés de 4 sigma ou plus sont supprimés, et ainsi de suite, en affinant progressivement le fit. Le nombre d'itérations est demandé à l'utilisateur, mais il reste aux alentours de 3 à 8 suivant la qualité des données.

Ce programme marche relativement bien, cependant la valeur de σ (le bruit blanc) mesurée à la première itération est toujours meilleure que les suivantes. La raison est que les points sont plus nombreux dans les bins à haute fréquence, et les pics y sont négligeables. A chaque itération, on enlève toujours un peu de puissance, et la valeur du bruit blanc en pâtit plus que α et f_{knee} . J'ai donc appliqué les itérations pour déterminer ces deux dernières valeurs, alors que σ est désormais déterminé à la première itération.

Après de nombreux essais, j'ai pu estimer l'incertitude (assez variable suivant les pics introduits) entre 2 et 3% pour σ , 3 et 4 % pour α et 8 et 10% pour f_{knee} (c'est compréhensible car les pics se concentraient surtout dans la région de f_{knee}).

L'étape suivante a été d'appliquer ce programme aux vraies données avec MASTER...

VI. Estimation du bruit dans les données ARCHEOPS

Présentation du problème.

Le simulateur MASTER a besoin d'un modèle du bruit pour pouvoir le retirer en calculant les $\langle \tilde{N}_\ell \rangle$. Ce modèle doit être déduit des données. En général, dans les expériences de ce type, le bruit est en $1/f$. Le fait de faire cette modélisation réduit le nombre de paramètres à trois : σ , α et f_{knee} . C'est pour les estimer que mon programme de fit a été utile. La qualité du résultat obtenu dépend de la façon dont les données collent au modèle : elle est médiocre s'il y a trop de pics et pire si l'on constate des dérives importantes. Nous espérons cependant nous en tirer avec un fit en $1/f$.

J.C. Hamilton a testé MASTER sur les données du vol de Kiruna 2001 avec mon programme de fit pour estimer le bruit. Malheureusement le spectre de puissance angulaire issu du traitement n'a pas montré une allure de spectre de CMB mais plutôt celle d'un bruit blanc. Ceci montrait que le bruit avait été mal soustrait. Après vérification, nous avons constaté que le spectre du modèle issu du fit se superposait mal au spectre des données (l'erreur dépassait les 10% par endroits). Le programme de fit ayant été testé avec des résultats plus probants, il n'a pas été mis en défaut. Nous en avons conclu que le bruit n'était pas vraiment du $1/f$. En effet, en coordonnées log-log, au lieu du plateau supposé pour le bruit blanc à haute fréquence pour un modèle en $1/f$, le spectre du bruit montrait une décroissance constante !

A défaut de trouver un autre modèle théorique de bruit, ce qui semblait impossible, J.C. Hamilton a préféré utiliser différentes techniques pour mieux coller aux données. Nous nous sommes servis du seul avantage d'avoir un rapport signal sur bruit de $1/100$: le spectre des données est proche celui du bruit. Nous avons convenu que pour valider les essais d'A. Amblard et J.C. Hamilton, je devrai refaire les en parallèle les mêmes manipulations sur des simulations.

Modifications du simulateur :

Pour effectuer ce travail, j'ai pu apprécier la souplesse de la conception du simulateur MASTER, qui était au départ prévu pour un modèle de bruit en $1/f$: il est composé de nombreux modules indépendants (calcul de la matrice M, simulation de CMB, applications des corrections, etc.) reliés entre eux par un script que chacun peut composer à son goût. J'ai donc écrit mon propre script, plus spécifique à la tâche qui m'était affectée que le script existant. J'ai dû aussi modifier certains modules de MASTER, codés uniquement pour du bruit en $1/f$, et en ajouter de nouveaux. Voici les principales modifications :

- Le module principal de MASTER, `master_simulator`, codé en C++, permet d'obtenir des simulations de CMB et de bruit. L'utilisateur indique le nombre de simulations voulues, donne un modèle de CMB (par la valeur des C_ℓ) et/ou un modèle de bruit en $1/f$ (par ses trois coefficients), et indique la stratégie de scan utilisée. En sortie, le programme génère autant de fichiers que de simulations voulues, contenant les \tilde{C}_ℓ finals. Ceux-ci sont calculés en appliquant l'analyse en harmoniques sphériques à une carte de CMB simulée, convoluée par la fonction fenêtre de la couverture du ciel, et à laquelle a été ajouté le bruit. L'utilisateur peut aussi obtenir des simulations de CMB

seul ou de bruit seul. Les modifications que j'ai apportées ont concerné la partie simulant le bruit : au lieu de rentrer les trois coefficients du bruit en $1/f$, l'utilisateur peut donner le spectre de bruit complet (grâce à un fichier contenant les points du spectre). Cela permet et de généraliser à tous les spectres possibles. A partir de là, le programme fabrique les 'timeline' de bruit. Pour cela, il faut passer de l'espace de Fourier à l'espace réel. La méthode utilisée génère un bruit gaussien. Pour des raisons exposées plus loin, j'ai inclus la possibilité de passer dans l'espace réel en tirant uniquement des phases aléatoires (alors que le module est aussi affecté dans le cas d'un bruit gaussien).

- Ajout d'un simulateur de 'timeline' de CMB, à partir d'une carte de CMB et de la stratégie de scan simulées. `master_simulator` avait déjà cette fonction, mais en interne : la 'timeline' n'était pas donnée en sortie.

Lien entre les deux spectres de puissance.

Parallèlement à ce travail de codage, il m'a semblé important de comprendre mieux comment le spectre de puissance temporel et le spectre de puissance angulaire étaient reliés. Les opérations pour passer de l'un à l'autre sont la fabrication d'une 'timeline', sa projection sur le ciel grâce à la connaissance de la stratégie de scan du ballon, et le calcul des harmoniques sphériques, corrigées a posteriori des effets de la fonction fenêtre et du filtrage. Il faut remarquer tout d'abord qu'à un spectre temporel ne correspond pas un spectre angulaire : la fabrication de la 'timeline' force à introduire des phases aléatoires, dont va dépendre la carte obtenue ensuite. C'est ce qui exige de faire de nombreuses simulations et de prendre la moyenne pour accéder aux $\langle \tilde{C}_\ell \rangle$. Il est impossible de procéder autrement, car on ne connaît pas les phases de notre bruit (ceci reviendrait à pouvoir distinguer, sur la 'timeline' des vraies données, ce qui est du bruit et ce qui est du CMB ! c'est la raison pour laquelle on utilise le spectre de puissance, d'où l'on peut tirer des informations sur le bruit, moyennant la perte des phases).

Une question se pose donc : est-il possible, compte tenu du fait que l'on moyenne sur de nombreuses simulations, de trouver de la structure sur la carte à partir d'un spectre temporel, c'est à dire d'avoir un spectre de puissance angulaire non constant ? Il semble raisonnable de penser que les phases aléatoires, combinées à une stratégie de scan circulaire à grande échelle (voir partie III.) vont effacer toute information. Il est cependant possible d'envisager la formation de structure dans des cas précis : si on envisage un pic unique à une fréquence f dans le spectre angulaire, cela se traduira par une 'timeline' sinusoidale, et lors de la projection sur la carte on peut s'attendre à retrouver, si f est grand devant la fréquence de rotation du ballon (environ 0.03Hz), des structures angulaires dont la taille est de l'ordre de v/f où v est la vitesse angulaire de parcours du ciel par le télescope. Ceci a effectivement été observé par A.Amblard dans une simulation. Mais dès que le spectre de bruit ne présente pas de pic dominant, ce qui est le cas pour Kiruna, tous les pics créent des structures à v/f avec f qui balaie tout le spectre, et on s'attend à ce que les $\langle \tilde{N}_\ell \rangle$ soient constants.

Pour vérifier ce raisonnement, j'ai simulé une 'timeline' de CMB pur, puis j'ai pris son spectre de puissance temporel, que j'ai mis en entrée de `master_simulator` afin de calculer les $\langle C_\ell \rangle$. Il y a donc une reprojction du spectre temporel du CMB sur la carte.

Effectivement, la forme caractéristique du spectre a complètement disparu, et le spectre obtenu s'apparente à celui d'un bruit blanc très faible. (voir annexe 2)

Proposition pour l'estimateur du bruit.

Vu que l'information sur la structure du CMB n'est pas contenue dans le spectre temporel des données de Kiruna, et comme le bruit y est 100 fois plus fort que le signal, j'ai proposé de prendre le spectre des données, tel quel, comme spectre de bruit. C'est d'ailleurs dans ce but que j'avais modifié MASTER. J'ai donc testé cette hypothèse en prenant le spectre de Kiruna comme modèle de bruit, à partir duquel j'ai simulé une nouvelle 'timeline' de bruit pur. J'y ai ajouté une 'timeline' de CMB simulé, pour obtenir une 'timeline' CMB+bruit, avec laquelle j'ai calculé les \tilde{C}_ℓ . Ensuite, je me suis de nouveau servi du spectre de Kiruna comme modèle de bruit pour calculer les $\langle \tilde{N}_\ell \rangle$. C'est pour cette dernière opération que les modifications de master_simulator ont été utiles, car j'ai pu lui faire générer ses 'timelines' à partir du spectre, et en utilisant uniquement des phases aléatoires au lieu de générer un bruit gaussien. Par ce choix, toutes les 'timelines' ainsi fabriquées ont bien un spectre de puissance temporel identique au spectre de Kiruna, non multiplié par un spectre de bruit blanc car cela n'a plus lieu d'être ici. En contrepartie, le calcul des barres d'erreurs sur un modèle gaussien n'est plus possible, ce qui est sans importance vu que la mesure que je cherchais à faire était celle du biais introduit par mon estimateur de bruit. En appliquant alors MASTER dans ces conditions, j'ai pu montrer qu'on retrouvait bien le spectre de CMB en sortie, avec un biais mineur (voir annexe 3).

Application aux données :

J.C. Hamilton a appliqué aux données l'estimateur de bruit qu'il avait réalisé avec A.Amblard : un fit amélioré, avec une interpolation polynômiale, sans supposer la forme en $1/f$. Dans un second temps, il a aussi utilisé mon estimateur de bruit (en réalité une légère variante, un lissage du spectre au lieu du spectre tel quel, que j'avais également testée au préalable). Dans les deux cas, le bruit à soustraire était surestimé : le spectre était même négatif... J'ai testé par des simulations les deux estimateurs de bruit, qui semblaient tous les deux corrects, le fit étant légèrement biaisé (annexe 4).

J.C. Hamilton a alors arbitrairement multiplié le spectre du bruit (les $\langle \tilde{N}_\ell \rangle$) par un facteur correctif afin de redresser le spectre final. Pour un facteur 0.88, il a obtenu un spectre tout à fait satisfaisant. Il restait alors à trouver pourquoi le bruit était surestimé de 12%. Après de nombreuses réflexions, et de nombreux essais infructueux, nous nous sommes aperçu que nous avions estimé le bruit à partir du spectre des données entières, alors que le calcul des \tilde{C}_ℓ se faisait après avoir éliminé de la carte les mauvais échantillons, et surtout la zone du ciel où le télescope croisait la Voie Lactée : notre modèle surestimait le bruit, prenant en compte les échantillons les plus bruités, n'intervenant pas dans le calcul des \tilde{C}_ℓ ! Après rectification, nous avons obtenu des résultats corrects, sans facteur arbitraire, assortis de barres d'erreurs, et à la suite desquels nous avons publié une note destinée aux membres de la collaboration Archeops.

N.B : cette note doit être acceptée avant que les résultats soient publiés hors de la collaboration, et c'est la raison pour laquelle ils ne seront pas présentés ici.

Conclusions et perspectives

Perspectives scientifiques

Le simulateur MASTER que nous avons codé d'après l'algorithme de E. Hivon a donc démontré ses capacités à analyser correctement les données d'Archeops, moyennant une estimation du bruit plus raffinée que du simple $1/f$. Nous pouvons penser que cette méthode sera utilisée avec profit dans les expériences à venir. Nous avons pu, grâce à MASTER, simuler les vols prévus fin 2001 et début 2002 à Kiruna, en constatant que les barres d'erreurs étaient considérablement réduites pour un vol de 24 heures (voir annexe 5). Ce résultat nous rend confiant dans l'avenir d'Archeops, qui n'a pas pu pour le moment utiliser pleinement ses capacités, et permettra d'argumenter avec plus de poids en faveur de l'augmentation du temps de vol.

Conclusion personnelle

Ce stage a été pour moi très enrichissant, à plusieurs niveaux. Cela aura été mon premier contact avec le travail de recherche, un bref aperçu de deux mois, qui donne l'envie d'y revenir vite, et pour plus longtemps... Je considère comme une chance exceptionnelle d'avoir vu une petite équipe se former autour d'un projet comme MASTER le jour même de mon arrivée, d'avoir pu contribuer à l'évolution et aux tests de ce simulateur à des niveaux variés pendant deux mois, et de quitter le laboratoire le jour même de la publication de la note concrétisant les résultats obtenus !

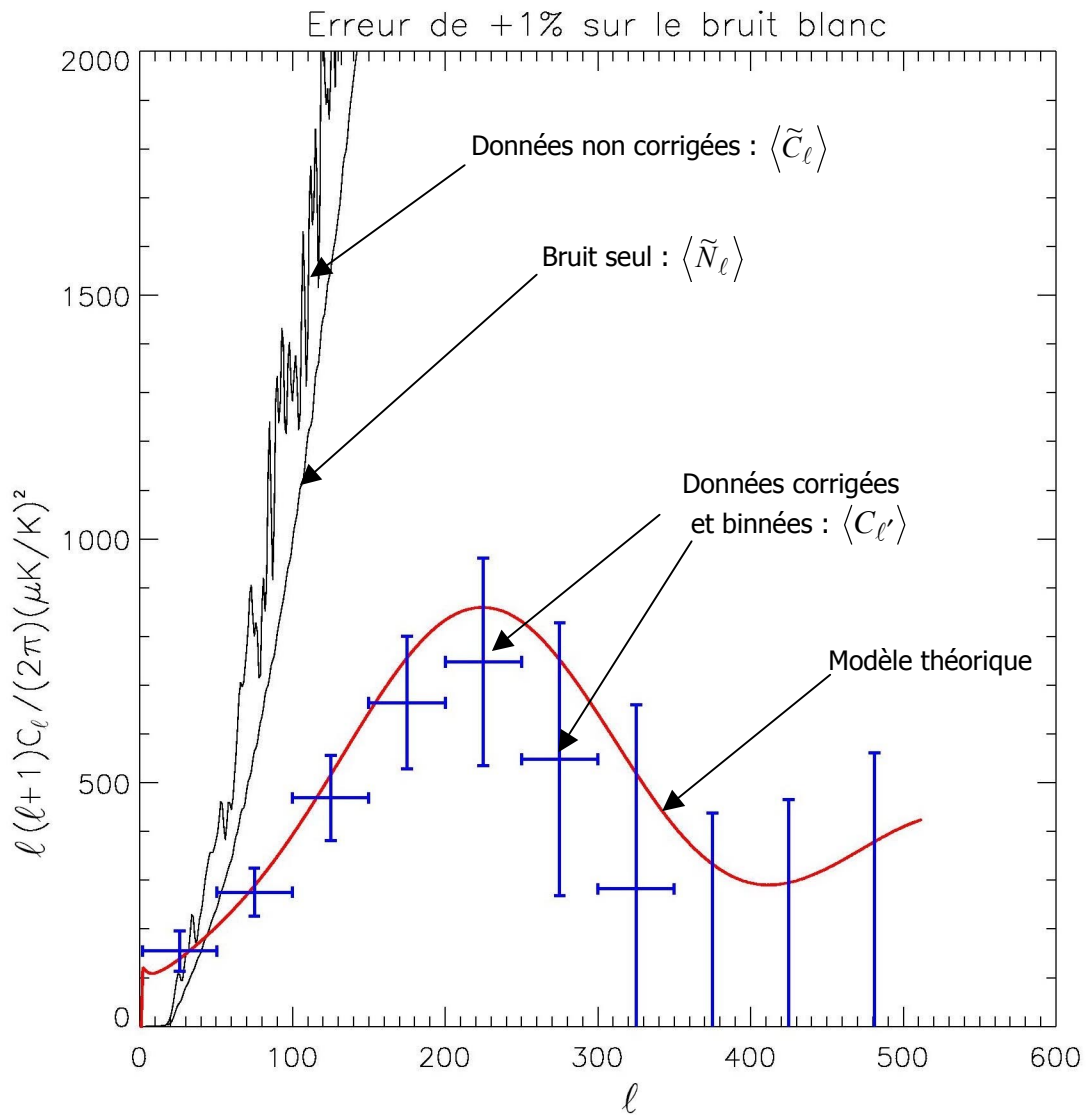
Certes, l'avancement du projet n'a pas été aussi linéaire que pourrait le laisser supposer cette synthèse : il y a eu des moments de découragement, où le but à atteindre semblait inaccessible, du temps perdu à cause de problèmes informatiques, de bugs... Certes, l'expérience Archeops n'a pas pleinement réussi sa mission, et les résultats obtenus, ainsi que les barres d'erreurs sont modestes par rapport aux expériences précédentes. Mais c'est ce qui a justement donné de l'intérêt à mon stage : le succès de notre projet a été d'arriver à tirer des résultats de données aussi bruitées. Cet important niveau de bruit nous a poussé à raffiner sa modélisation, et cela a été bénéfique : des données d'une qualité parfaite réduiraient le chercheur à exécuter des tâches automatiques, sans se poser de questions, sans faire avancer la recherche.

Enfin, j'ai pu apprécier grâce à l'encadrement de qualité offert par le PCC la vie d'un laboratoire et de ses chercheurs au quotidien. J'ai même été invité à participer à la réunion Archeops de Juin, à Grenoble, qui m'a permis de mieux situer mon sujet de stage par rapport aux travaux et aux préoccupations de la collaboration.

Je remercie donc vivement J.C. Hamilton et A. Amblard pour m'avoir accepté comme stagiaire et pour m'avoir donné sans compter autant de leur temps et de leurs connaissances.

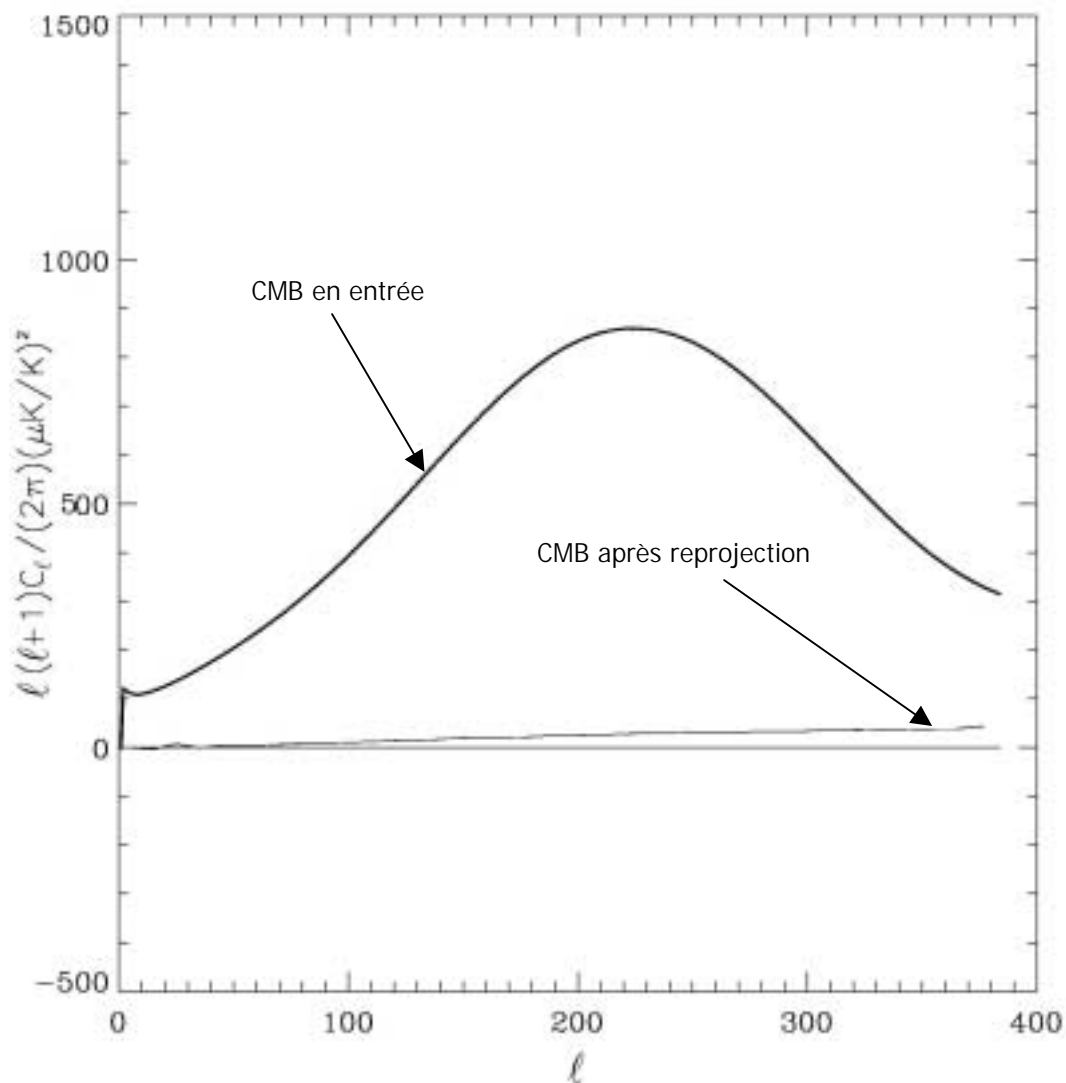
J'attends avec impatience le prochain stage...

Annexe 1



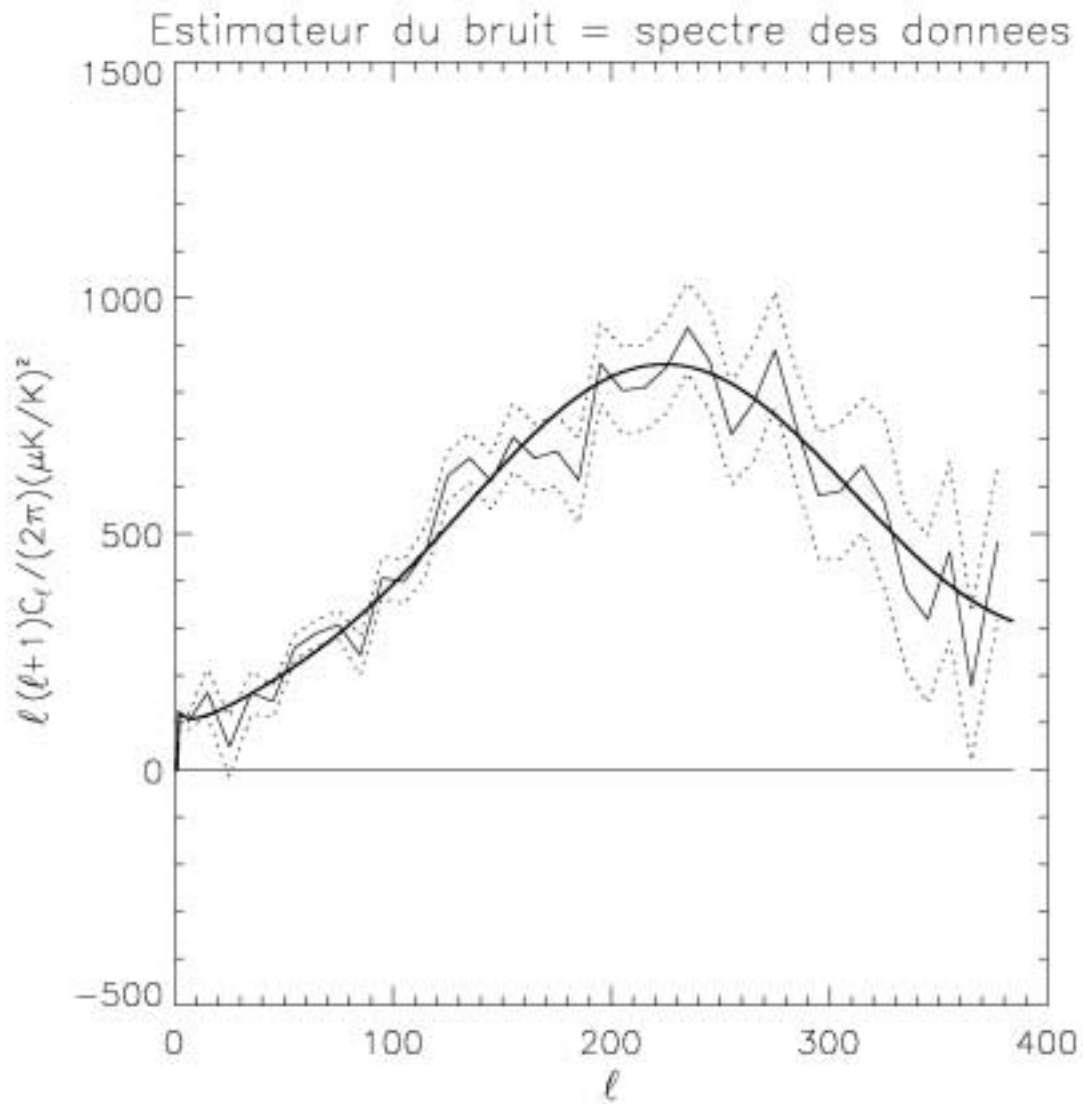
Simulation de l'effet d'une mauvaise estimation du bruit pour les données de Trapani. Ici, le modèle de bruit introduit est en $1/f$. L'erreur simulée est de 1% sur le paramètre σ . Cela permet de se rendre compte de la qualité de l'estimateur de bruit requis.

Annexe 2



*Effet de la reprojection d'un spectre temporel simulé de CMB pur .
Les phases aléatoires introduites pour obtenir une 'timeline',
puis une carte, ont totalement dilué la structure contenue initialement.*

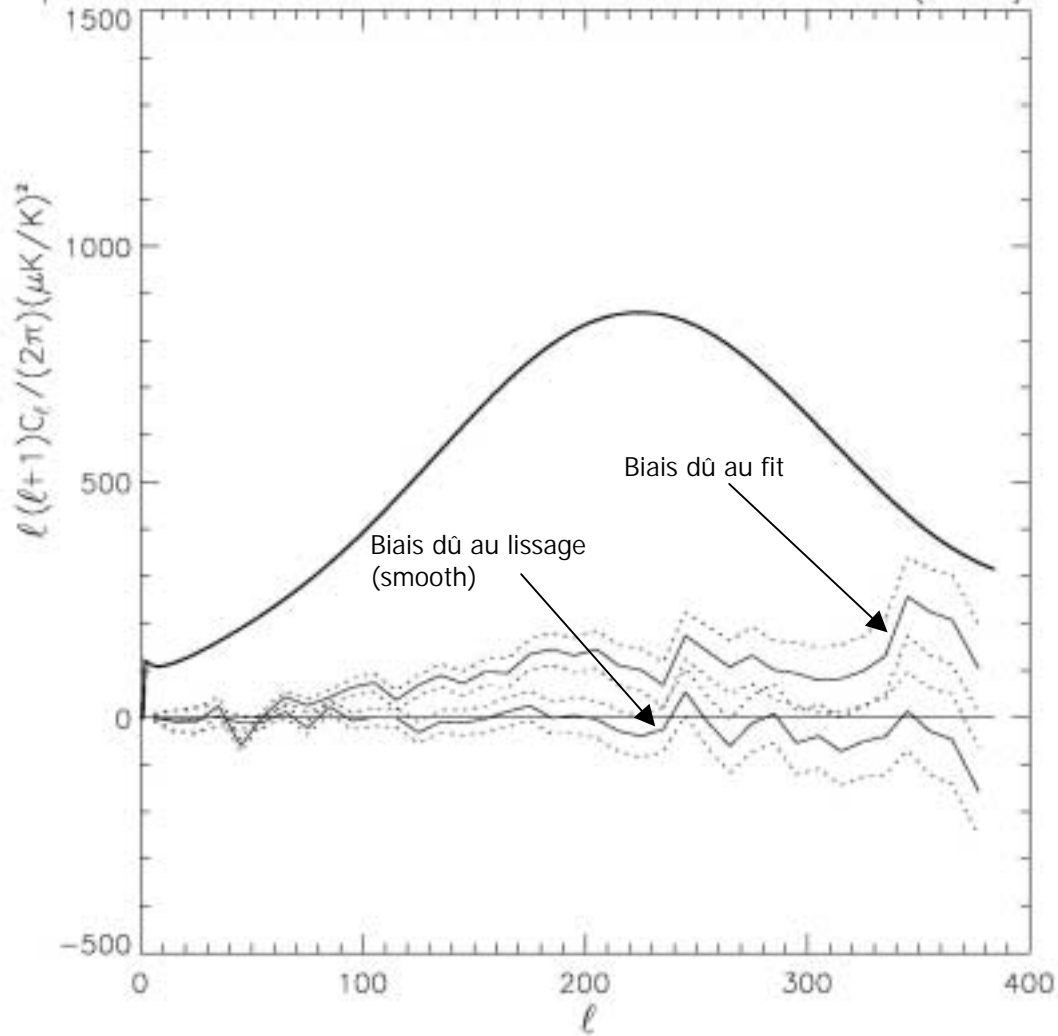
Annexe 3



Erreur due au modèle de bruit, en prenant comme estimateur du spectre de bruit le spectre des données lui-même. Les lignes pointillées représentent l'incertitude à 1σ sur cette erreur.

Annexe 4

Comparaison de deux estimateurs du bruit: un smooth(1000) et un fit

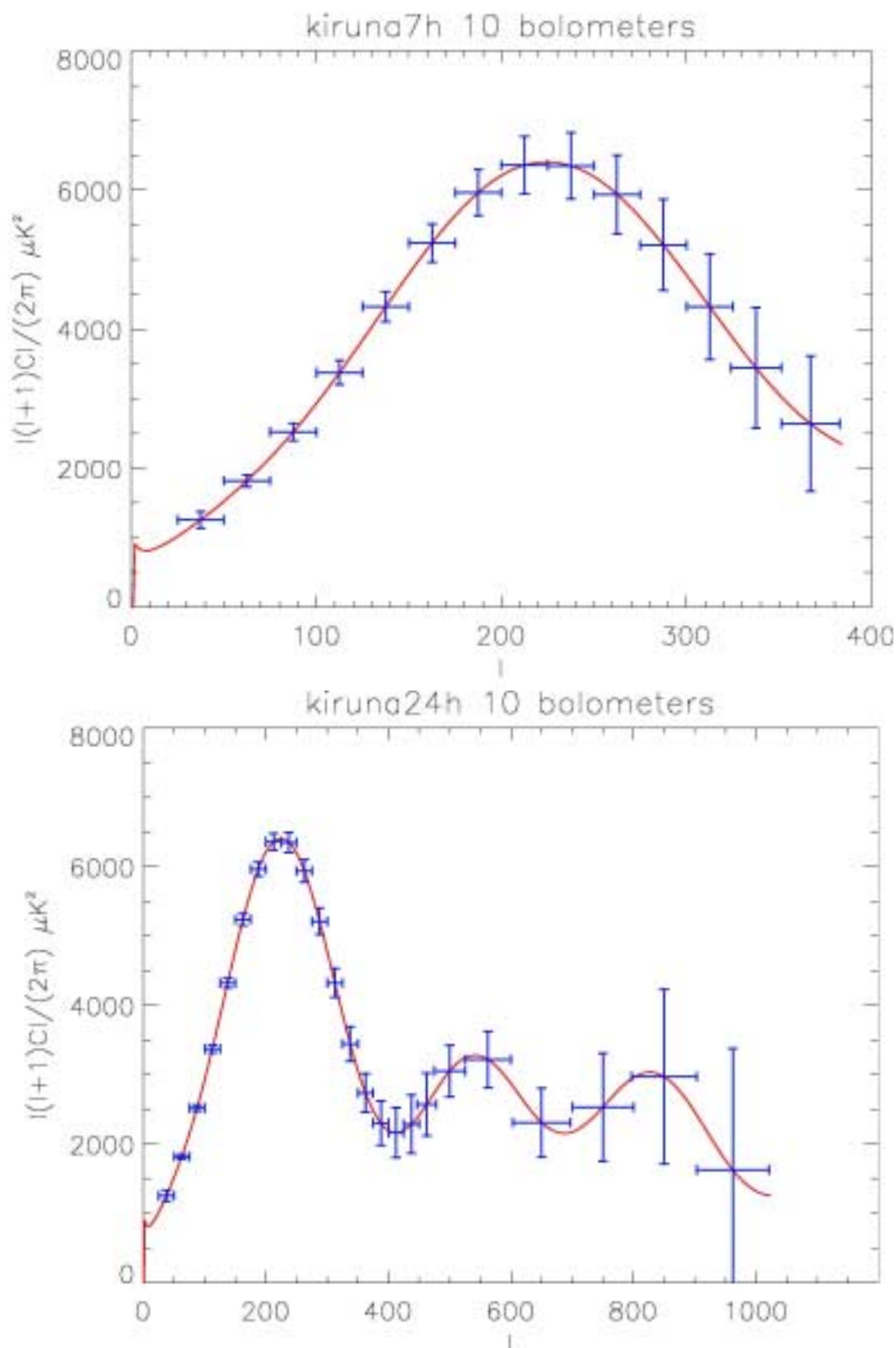


Simulation du biais introduit par deux estimateurs du spectre de bruit différents : un fit du spectre des données et un lissage du même spectre.

Les pointillés indiquent l'incertitude à 1σ sur ces biais.

Le modèle de CMB a été ajouté pour permettre de comparer les niveaux.

Annexe 5



Simulation des barres d'erreur, avec le niveau de bruit mesuré sur les données de Kiruna, et en appliquant MASTER à un vol de 7 heures (comme en Janvier 2001) et un vol de 24 heures (comme il est prévu pour l'hiver 2001/2002). Cette simulation suppose toutefois que l'on utilise des données de 10 bolomètres (nous en avons pour le moment combiné 3). On voit ici clairement l'importance d'effectuer de longs vols.

Références :

- [1] G. Gamow, Nature **162**, 680 (1948)
- [2] R. Alpher et R. Herman, Nature **162**, 774 (1948)
- [3] A.A. Penzias et R.W. Wilson, ApJ Lett. **142**, 419 (1965)
- [4] O. Doré, R.Teyssier, F.R. Bouchet , D. Vibert, S. Prunet , (2001),
soumis à A&A, astro-ph/0101112
- [5] E. Hivon et al., (2001) astro-ph/0105302

Bibliographie :

- Anisotropies de température et polarisation du rayonnement fossile : méthodes de détection et traitement de données, Thèse de B. Revenu (PCC, 05/2001)
- Naissance et Histoire du Cosmos , La Recherche, Hors Série n°1, (04/1998)
- Master of the CMB Anisotropy Power Spectrum : a fast method for statistical Analysis of large and complex CMB data sets, E.Hivon et al, (2001), astro-ph/0105302
- Ressources diverses sur le site : www.archeops.org.