

Séparation de composantes semi-aveugle pour des données polarisées du fond diffus cosmologique avec l'algorithme EM

Jonathan AUMONT

Laboratoire de Physique Subatomique et de Cosmologie,
Grenoble



Résumé

La mesure de la polarisation du fond diffus cosmologique (CMB) s'avère être un des défis majeurs de la cosmologie observationnelle contemporaine. En effet, son étude permettrait de lever les dégénérescences subsistant dans la détermination des paramètres cosmologiques à partir de la seule température du CMB. Sa mesure reste néanmoins délicate. En effet, il s'agit d'un signal très faible, qui se trouve être contaminé par la polarisation de rayonnements d'origine astrophysique. La connaissance de ces rayonnements parasites polarisés est insuffisante pour envisager une soustraction directe de leurs contributions. C'est pourquoi nous avons développé une méthode semi-aveugle de séparation de composantes appliquée aux données de type CMB polarisées. Nous présenterons ici brièvement le formalisme de cette méthode, ainsi que les grandes lignes de son algorithme. Nous présenterons également les simulations de type *Planck* que nous avons réalisées afin de tester et valider notre méthode. Pour une revue détaillée de la méthode et des résultats, se référer à [1]

56.1 Introduction

Le fond diffus cosmologique, ou *Cosmic Microwave Background* (CMB) est un rayonnement de corps noir à 2.725 K résultant du découplage entre matière et rayonnement, survenu environ 300.000 ans après le Big-Bang. Ce rayonnement, dont la température est en première approximation isotrope, comporte des anisotropies de l'ordre de 10^{-5} par rapport à la température moyenne. L'étude de ces anisotropies a permis au cours de ces dernières années à la cosmologie observationnelle de contraindre, de manière inégale, les différents modèles cosmologiques. En effet, la statistique spatiale des anisotropies dépend fortement des *paramètres cosmologiques* définissant notre Univers.

Le CMB est également un rayonnement polarisé. À l'instar de la température, la polarisation comporte aussi des anisotropies, plus délicates à détecter en raison de leur faible amplitude (entre 2 et 5 ordres de grandeur plus faibles que les anisotropies en température). L'étude de ces dernières permet de lever les dégénérescences restantes dans la détermination des pa-

ramètres cosmologiques. En outre, ce sera dans les anisotropies en polarisation que nous pourrions trouver des informations relatives à une phase de l'Univers primitif antérieure à la formation du CMB : *l'inflation*. L'un des enjeux de l'étude de la polarisation du CMB est la recherche de l'empreinte sur son spectre de puissance angulaire, des *ondes gravitationnelles primordiales* générées au cours de la période d'inflation.

La polarisation du CMB est linéaire, aussi son étude se fera dans le formalisme des paramètres de Stokes I , Q et U . Nous définirons, pour ne pas dépendre du repère choisi pour l'observation, les champs (modes) T correspondant à l'intensité I , ainsi que E et B correspondant à la polarisation linéaire Q et U [12]

La recherche de cette signature, et l'appréhension des anisotropies en polarisation de manière générale, en raison de la faiblesse du signal, doivent passer par une meilleure compréhension et une détermination explicite des contaminants que rencontre cette détection. En effet, aux fréquences d'étude du CMB, plusieurs contaminants d'avant-plan ou *foregrounds*, viennent se superposer aux mesures, dominant même à certaines fréquences le signal cosmologique. L'émission de rayonnement synchrotron due au mouvement hélicoïdal des électrons dans le champ magnétique Galactique domine à basses fréquences, alors que l'émission des grains de poussière Galactique domine aux hautes fréquences. De plus ces émissions semblent fortement polarisées comme le montrent des mesures récentes de la polarisation du rayonnement synchrotron à 1.41 GHz [11] ou les mesures de l'expérience ballon *Archeops* à 353 GHz qui montrent que l'émission diffuse de la poussière Galactique aurait un degré de polarisation de l'ordre de 5 à 10%, aussi bien près du centre Galactique [2], qu'à hautes latitudes Galactiques [8]. Ces émissions restent cependant mal connues, particulièrement en polarisation, ce qui ne permet pas une soustraction directe de ces contaminants. Il sera donc crucial pour les expériences à venir de développer des méthodes performantes pour séparer avec précision le CMB des différentes composantes Galactiques, particulièrement en ce qui concerne la polarisation.

Pour s'affranchir des limitations dues à la pauvreté des connaissances sur les émissions d'avant-plan, un grand nombre de travaux ont été réalisés pour élaborer des algorithmes permettant de discriminer entre le

CMB et les contaminants. Ces méthodes permettent l'analyse de données multi-fréquences en s'appuyant sur la diversité des spectres électromagnétiques des différentes composantes du ciel micro-onde. Elles permettent également d'extraire les propriétés d'émission des avant-plans. La plupart de ces méthodes sont dédiées à la seule température. Nous présentons ici la méthode que nous avons développée, qui est une extension à la polarisation d'une méthode de comparaison des spectres angulaires [3]. Cette méthode est dite semi-aveugle, c'est à dire que nous n'aurons pas a priori ni sur les spectres de puissance angulaire des signaux, ni sur le bruit. Nous supposons seulement que le spectre d'émission du CMB est connu.

Nous ne présenterons ici qu'une description succincte de notre méthode, des simulations réalisées afin de pouvoir la tester, ainsi que nos résultats préliminaires. Pour plus de détails se référer à [1].

56.2 Modèle du ciel micro-onde

Pour pouvoir mesurer avec précision les anisotropies du CMB, les expériences dédiées à son observation doivent atteindre une grande précision, plus faible que le niveau attendu de la contamination due aux avant-plans astrophysiques. Pour opérer une telle séparation, la diversité des spectres électromagnétiques et spatiaux des différentes composantes est généralement utilisée. Les observations d'un instrument multi-fréquences peuvent être modélisées comme des combinaisons linéaires de multiples composantes physiques. C'est ce que l'on appelle une modélisation Multi Détecteurs Multi Composantes (MD-MC).

Faisons l'hypothèse d'une expérience avec n_ν détecteurs observant à des fréquences ν_i et n_c composantes dans les données. Pour chaque mode $X = \{T, E, B\}$ dans l'espace des harmoniques sphériques pour chaque couple $\{\ell, m\}$ en supposant une couverture complète du ciel et en négligeant l'effet des lobes instrumentaux, nous pouvons écrire les données $y_{\ell m}^{\nu, X}$ pour chaque détecteur :

$$y_{\ell m}^{\nu, X} = \sum_{c=1}^{n_c} A_c^{\nu, X} s_{\ell m}^{c, X} + n_{\ell m}^{\nu, X} \quad (56.1)$$

où $s_{\ell m}^{c, X}$ est la carte du ciel pour chacune des composante c , $n_{\ell m}^{\nu, X}$ est la carte de bruit pour chaque détecteur à la fréquence ν_i et $A_c^{\nu, X}$ est la *matrice de mélange* qui donne le comportement électromagnétique de la composante c à la fréquence ν_i . Le but des méthodes de séparation de composantes est d'extraire à partir des données $y_{\ell m}^{\nu, X}$, les différentes composantes $s_{\ell m}^{c, X'}$, la matrice de mélange $A_c^{\nu, X'}$ et le bruit $n_{\ell m}^{\nu, X}$.

Pour réduire le nombre de paramètres à extraire nous allons définir des matrices de densité qui contiennent les

spectres de puissance angulaire "binnés" des coefficients dans l'espace des harmoniques sphériques définis dans l'équation (56.1) :

$$R_\chi(b) \equiv \frac{1}{n_b} \sum_{\ell \in \mathcal{D}_b} \langle \chi_{\ell m}^{i, X} \cdot \chi_{\ell m}^{j, X' \dagger} \rangle, \quad \chi = \{y, s, n\} \quad (56.2)$$

où \mathcal{D}_b représente l'ensemble des ℓ appartenant au "bin" b . Nous pouvons ainsi réécrire l'équation (56.1) pour ces matrices :

$$R_y(b) = A R_s(b) A^T + R_n(b) \quad (56.3)$$

Le fait important est que ces matrices de densité contiennent les spectres d'auto-corrélation angulaire relatifs à la température et à la polarisation, ainsi que les spectres croisés température-polarisation et polarisation E -polarisation B . Séparer les différentes composantes de manière semi-aveugle reviendra donc à estimer les matrices $R_s(b)$, $R_n(b)$ et une partie de A à partir des données multi-canaux $R_y(b)$. Nous appellerons l'ensemble des paramètres à estimer $\theta(b) = \{R_s(b), R_n(b), A\}$.

56.3 Un algorithme MD-MC de séparation de composantes pour la polarisation

Pour pouvoir estimer les différents paramètres ci-dessus à partir des données, nous allons essayer de trouver la meilleure adéquation entre le modèle $\hat{R}_y(b)$ calculé à partir des estimées $\hat{\theta}(b)$ et les données $R_y(b)$. La log-vraisemblance sera une mesure raisonnable de cette adéquation. Nous ferons également l'hypothèse que les différentes composantes physiques ainsi que le bruit sont des réalisations de champs stationnaires gaussiens (approximation de Wittle). La maximisation de la vraisemblance (i.e. minimisation de la log-vraisemblance) est réalisée grâce à un algorithme d'*Expectation-Maximization* (EM) [4]. Cet algorithme procède itérativement depuis un état initial $\theta_0(b)$ en suivant une série de mises à jour des paramètres $\theta_i(b)$, appelées *étapes EM*. Par construction, chaque étape EM améliore l'adéquation entre les données et le modèle en maximisant la vraisemblance [4]. Pour une revue détaillée de cet algorithme, voir [3].

Le spectre du CMB étant déterminé de manière théorique, nous avons supposé connu, au cours des travaux présentés ici le spectre électromagnétique du CMB, c'est à dire que nous fixons la colonne de la matrice A relative au CMB. En revanche, nous ne pouvons pas trouver de motivations à une telle supposition en ce qui concerne le comportement des composantes d'origine astrophysique. Nous avons en outre apporté un

soin particulier à l'initialisation des paramètres (calcul de $\theta_0(b)$), car la convergence de l'algorithme y est très sensible.

56.4 Simulations du ciel observé avec *Planck*

Afin de tester les performances de notre méthode nous avons fait des simulations de cartes du ciel complètes, en température et en polarisation, pour le satellite *Planck*. Nous simulons une carte pour chacune des composantes aux fréquences de l'instrument que nous souhaitons simuler et nous additionnons les cartes de toutes les composantes à chaque fréquence. Nous avons pris en compte trois composantes physiques – le CMB et les émissions de rayonnement synchrotron et de la poussière Galactique – ainsi que du bruit instrumental :

- Les cartes I , Q et U du CMB sont générées aléatoirement à partir des spectres de puissance angulaire polarisés calculés à l'aide du logiciel *CAMB* [7] pour un ensemble de paramètres cosmologiques. Nous avons utilisé les paramètres cosmologiques calculés à partir des mesures du satellite *WMAP* [10] : $H_0 = 71 \text{ km} \cdot \text{s}^{-1} \cdot \text{Mpc}^{-1}$, $\Omega_b = 0.044$, $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\tau = 0.17$ et $r = 0.7$. Nous avons également tenu compte de l'effet des lentilles gravitationnelles sur les spectres polarisés du CMB.
- Pour le rayonnement synchrotron diffus, nous avons utilisé des cartes patrons en température et polarisation données par [5]. Le spectre électromagnétique de ce rayonnement est supposé isotrope, suivant une loi de puissance d'indice spectral $\alpha = -2.7$.
- L'émission de la poussière Galactique est générée aléatoirement à partir d'un modèle en loi de puissance [9]. Nous avons normalisé ce modèle pour qu'il soit compatible à grande échelle angulaire avec le spectre de corrélation croisée TE d'*Archeops* à 353 GHz [8]. L'émissivité de la poussière suit une loi de puissance d'indice spectral $\alpha = 2$, dans nos simulations.
- Le bruit instrumental est un bruit blanc généré aléatoirement et normalisé pour chaque fréquence à la sensibilité de l'instrument que nous souhaitons simuler.

Pour chacune des composantes ci-dessus, nous avons supposé les spectres électromagnétiques similaires en température et en polarisation.

Nous avons simulé des cartes complètes du ciel aux fréquences où les instruments LFI et HFI du satellite *Planck* sont sensibles à la polarisation (30, 44, 70, 100, 143, 217 et 353 GHz), additionnant les cartes de CMB, synchrotron, poussière et bruit, simulées à chacune de

ces fréquences. Le niveau de bruit correspond à une mission *Planck* durant 14 mois. Nous avons simulé ces cartes du ciel en utilisant la pixelisation *Healpix* [6] pour un $N_{\text{side}} = 128$ (pixels ayant une aire de $\sim 27 \text{ arcmin}^2$), permettant une reconstruction théorique des spectres jusqu'aux multipoles $\ell \sim 400$.

Pour chaque simulation réalisée, nous calculons la matrice de densité des données $R_y(b)$ à partir du jeu de cartes I , Q et U à la fréquence donnée. Le "bin" b , se référera ici à des pas de 10 en ℓ .

56.5 Résultats et conclusions

Pour chacune des 300 simulations de type *Planck* que nous avons réalisées de la manière décrite en 56.4, nous appliquons l'algorithme EM aux données polarisées décrites en 56.3. Nous avons essayé de reconstruire les spectres de puissance angulaire en température et polarisation des différentes composantes, $R_g(b)$, le spectre de bruit de chaque détecteur, $R_n(b)$ ainsi que les éléments de la matrice de mélange A correspondant aux spectres électromagnétiques de la poussière et du synchrotron.

Les résultats que nous avons obtenus pour les spectres de puissance angulaire du CMB sont portés sur la figure 56.1. Les spectres reconstruits sont représentés par des croix, les modèles utilisés pour construire les simulations par des traits pleins. Les barres d'erreur sont la dispersion calculée sur les résultats obtenus à partir des 300 simulations réalisées.

Pour le CMB (voir figure 56.1), nous pouvons voir que les spectres C_ℓ^{TT} et C_ℓ^{EE} sont reconstruits avec une grande précision jusqu'à $\ell \simeq 400$, ce qui représente la limitation théorique à la résolution à laquelle nous travaillons. Le spectre C_ℓ^{BB} est reconstruit avec efficacité jusqu'à $\ell \simeq 60$. À plus petite échelle angulaire, la reconstruction est biaisée. Ceci est dû au fait que le signal des modes BB est très faible (5 ordres de grandeur plus faibles que la température). De ce fait, l'algorithme de séparation ne "voit" pas ce signal et laisse un bruit blanc résiduel dans ce cas. Nous pouvons néanmoins voir que même si nous étions en mesure de soustraire efficacement ce bruit résiduel, les barres d'erreur seraient compatibles avec zéro, indiquant ainsi que la limitation provient du rapport signal sur bruit du signal recherché. L'intérêt de notre méthode, par rapport à une extrapolation directe des méthodes en température à la polarisation, est qu'elle permet également de reconstruire les spectres croisés TE , TB et EB . Nous pouvons voir que le spectre C_ℓ^{TE} est reconstruit avec précision jusqu'à $\ell \simeq 400$ et que les spectres TB et EB sont compatibles avec le modèle, nul pour ces deux spectres.

Les spectres de la poussière Galactique et de l'émission synchrotron sont quant à eux reconstruits jusqu'à

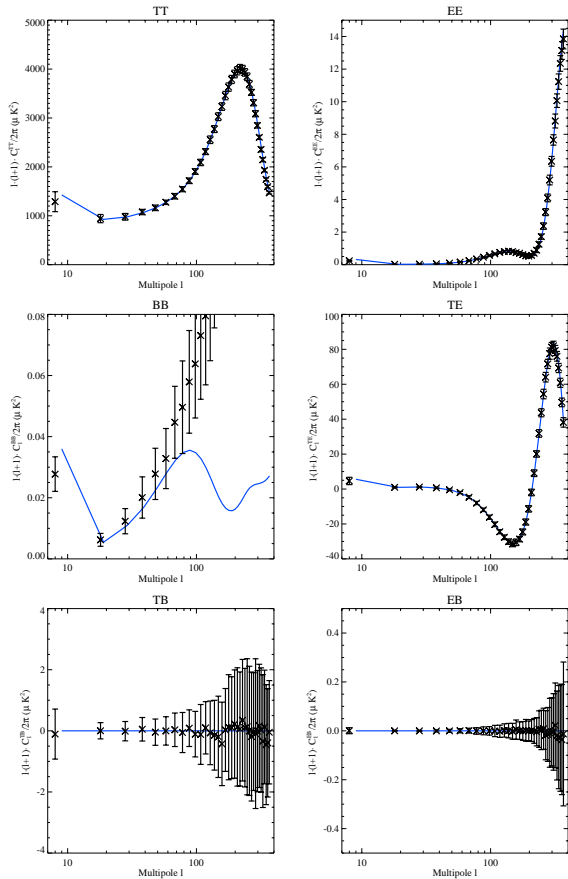


FIG. 56.1 – Spectres de CMB reconstruits à 100 GHz (croix). Le modèle utilisé pour générer les simulations est également porté (bleu). Les barres d'erreur sont la dispersion de la reconstruction sur 300 simulations.

$\ell \simeq 400$ pour tous les modes (TT , EE , BB , TE , TB et EB). Pour plus de détails et pour les figures, voir [1].

Les spectres de bruit de chaque détecteur simulé sont également reconstruits sur toute la gamme des échelles angulaires accessibles. La matrice de mélange A est reconstruite pour toutes les fréquences dans le cas de la poussière, en température et en polarisation. Pour le synchrotron, la matrice A est reconstruite à toutes les fréquences, en polarisation. En température, le spectre est biaisé à hautes fréquences. Ceci est dû au fait que en température, le synchrotron est sous-dominant par rapport à la poussière et par rapport au CMB, alors qu'en polarisation, ce n'est le cas que par rapport à la poussière. Pour plus de détails sur ces résultats, se référer à [1].

Nous avons développé une méthode de séparation de composantes applicable à des données de type CMB polarisées. Nous avons réalisé des simulations de type *Planck* afin de tester les performances de notre méthode. Outre le fait d'extraire des données les spectres

de puissance angulaire TT , EE et BB avec une très grande précision, pour le CMB et les avant-plans, cette méthode permet également une reconstruction des spectres croisés TE , TB et EB . De plus, nous avons montré que nous étions également capables de retrouver les spectres spatiaux du bruit en température et polarisation, pour chaque détecteur, ainsi que le spectre électromagnétique des avant-plans.

Cette méthode renvoie néanmoins un spectre biaisé dans le cas où le rapport signal sur bruit dans la carte du ciel est trop faible (de l'ordre de 10^{-2}). Il sera nécessaire d'améliorer cette méthode, l'algorithme EM convergeant très lentement au voisinage du maximum de vraisemblance, par l'ajout d'un algorithme de convergence rapide de type quasi-Newton afin de s'affranchir d'un tel effet. Il sera également nécessaire de tenir compte d'effets comme celui des lobes instrumentaux, ou encore celui de la couverture incomplète du ciel.

Références

- [1] Aumont J. & Macías-Pérez J.F., 2006, soumis à MNRAS, astro-ph/0603044
- [2] Benoît A., et al., 2004, A&A, 424,571
- [3] Delabrouille J., Cardoso J.-F. & Patanchon G., 2003, MNRAS, 346, 1089
- [4] Dempster A., Laird N. & Rubin D, 1977, J. of the Roy. Stat. Soc. B, 39, 1
- [5] Giardino G., Banday A. J., Górski K. M., Bennett K., Jonas J. L. & Tauber J., 2002, A&A, 387, 82
- [6] Górski K. M., Hivon E., Wandelt B. D., 1999, astro-ph/9812350
- [7] Lewis A., Challinor A. & Lasenby A., 2000, ApJ, 538, 473
- [8] Ponthieu N. et al., 2005, A&A, 444, 327
- [9] Prunet S., Sethi S. K., Bouchet F. R. & Miville-Deschêne M. -A., 1998, A&A, 339, 187
- [10] Spergel D. N. et al., 2003, ApJS, 148, 175
- [11] Wolleben M., Landecker T. L., Reich W. & Wielebinski R., 2005, A&A, submitted, astro-ph/0510456v1
- [12] Zaldarriaga M. & Seljak U., 1997, Phys. Rev. D, 55, 1830