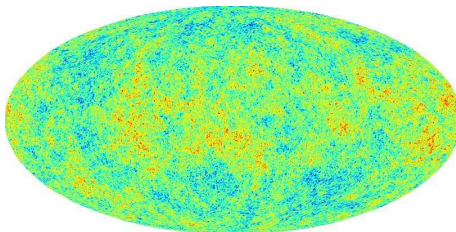


# Analyse des données du fond diffus cosmologique : simulation et séparation de composantes

Marc BÉTOULE

Direction : Jacques DELABROUILLE (APC) et Marc-Antoine MIVILLE-DESCHÊNES (IAS)

Soutenance de thèse, 25 septembre 2009



# Plan

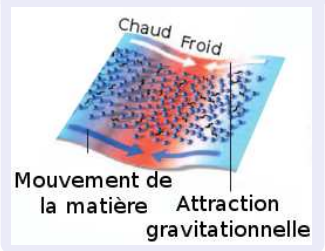
- 1 **Problématique**
  - Contexte
  - Le problème des avant-plans
- 2 **Modélisation de l'émission millimétrique**
  - Contexte
  - Le PLANCK Sky Model (PSM)
- 3 **Analyse de données**
  - Estimation spectrale en needlets
  - Extraction des sources ponctuelles
  - Avant-plans et détection des modes B
- 4 **Conclusion**
  - Conclusion
  - Perspectives



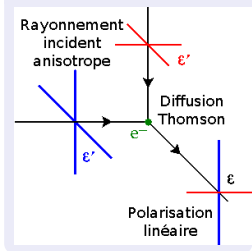
# Le fond diffus cosmologique



## Rayonnement anisotrope



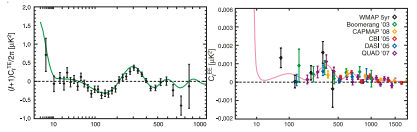
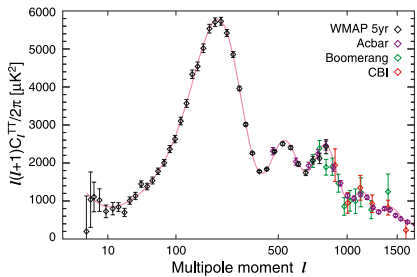
## et polarisé



# Mesures actuelles

## Fluctuations primordiales

- Fluctuations adiabatiques
- Invariance d'échelle
- Statistique gaussienne
- Oscillations acoustiques



## Modèle $\Lambda$ -CDM

Paramètre	WMAP+SN+BAO
$\Omega_b$	$0.0456 \pm 0.0015$
$\Omega_m$	$0.228 \pm 0.013$
$\tau$	$0.084 \pm 0.016$
$h$	$0.705 \pm 0.013$
$A_s$	$(2.445 \pm 0.096) 10^{-9}$
$n_s$	$0.960 \pm 0.013$



# Enjeux futurs

## Chercher des écarts au modèle standard

- Non gaussianité
- Fluctuations isocourbure

## Contraintes sur l'inflation

- Mesure plus précise de l'indice spectrale des modes scalaires  $n_s$
- Détection des modes B, mesure de  $r = T/S$

## Détection et exploitation d'autres effets

- Lentille gravitationnelle, réionisation, effet SZ, effet Vishniac
- ISW, contraintes sur l'énergie noire



# Enjeux futurs

## Chercher des écarts au modèle standard

- Non gaussianité
- Fluctuations isocourbure

## Contraintes sur l'inflation

- Mesure plus précise de l'indice spectrale des modes scalaires  $n_s$
- Détection des modes B, mesure de  $r = T/S$

## Détection et exploitation d'autres effets

- Lentille gravitationnelle, réionisation, effet SZ, effet Vishniac
- ISW, contraintes sur l'énergie noire



# Contexte expérimental

Lancement de PLANCK le 14 mai 2009



- 9 canaux entre 30 et 857 GHz (7 polarisés)
- Résolution 5 arcminutes  $\ell \sim 2000$
- Sensibilité

Haute résolution



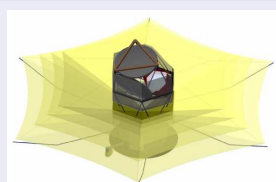
SPT, ACT...

Modes B



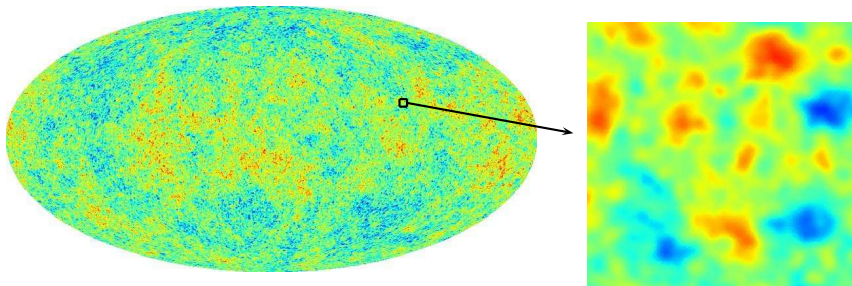
BICEP, QUAD,  
QUIET, EBEX,  
SPIDER...

4<sup>ème</sup> génération



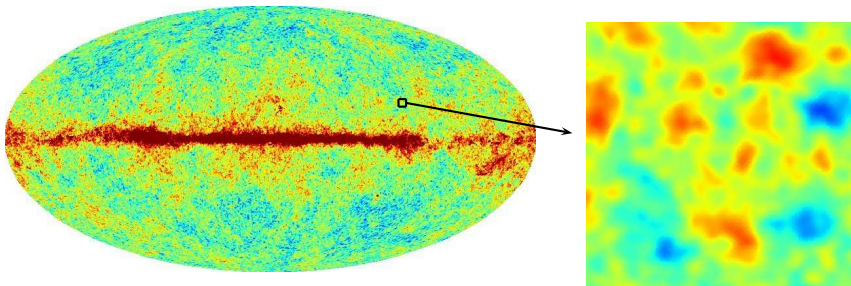
CMBPol et BPol.

# Contamination du signal





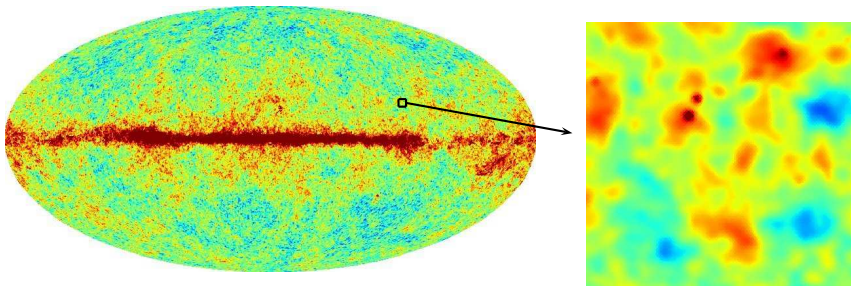
# Contamination du signal



- Émission galactique



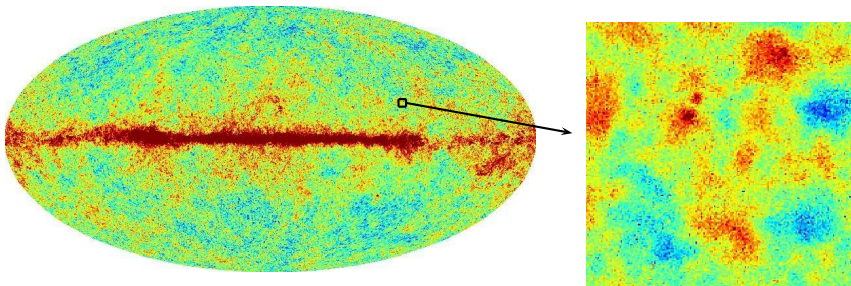
# Contamination du signal



- Émission galactique
- Sources extragalactiques



# Contamination du signal

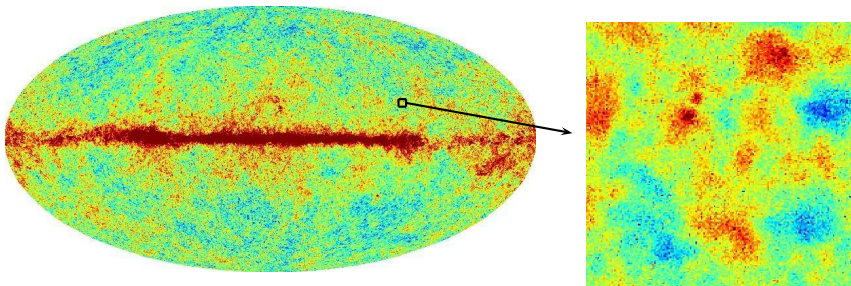


- Émission galactique
- Sources extragalactiques

- Bruit instrumental



# Contamination du signal

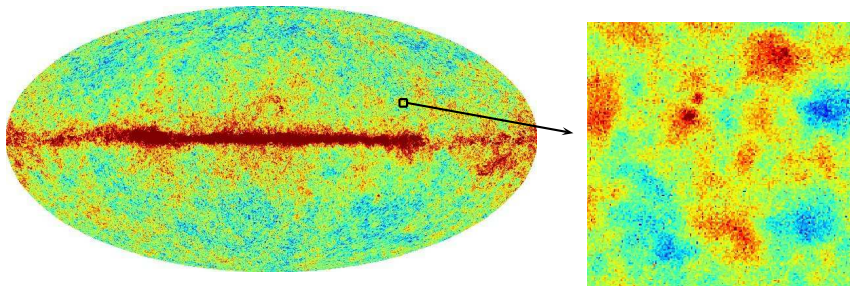


- Émission galactique
- Sources extragalactiques

- Bruit instrumental
- Anisotropies secondaires (Lensing, SZ ...)



## Contamination du signal



- Émission galactique
- Sources extragalactiques
- Bruit instrumental
- Anisotropies secondaires (Lensing, SZ ...)

La réalisation des objectifs scientifiques dépend de notre capacité à isoler ces différentes contributions.





# De nombreux processus à modéliser

## Galactique

synchrotron



rayonnement de freinage

poussière thermique



poussière anormale

Régions compactes (HII, cœurs denses...) émission spécifique.

## Extragalactique

- Processus propre à chaque source (Radio - Infrarouge)
- Fond de source infrarouge (CIB)

## “Cosmologique”

- Réionisation, effet SZ, effet Vishniac
- Effet de lentille gravitationnelle, ISW
- Fluctuations non gaussiennes



# Motivations pour le développement d'un outil cohérent

## Auparavant : modèles parcellaires pour chaque processus

### Simulation complète

- Processus fastidieux
- Unique
- Problèmes de traçabilité du travail

## Besoin au sein de PLANCK

- Simulations complètes et réalistes, faciles à réaliser
- Variabilité des simulations (Monte-Carlo, test de modèles)
- Traçabilité







# État

## Version stable du PSM : 1.6.6

- Mise en place de simulations complètes et raisonnablement complexes du ciel en température et en polarisation
- Effets instrumentaux (bruit et réponse instrumentale) minimaux
- Possibilité de produire :
  - Une prédiction (calibration, test de modèle ...)
  - Une simulation

## Limitations

- Polarisation de la poussière
- ...

## Applications

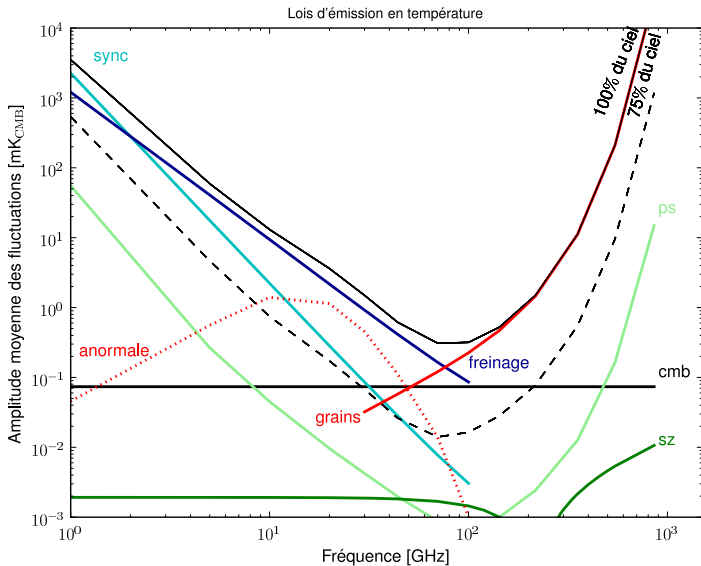
- Pour PLANCK : "Data Challenges" des groupes de travail, test end-to-end, calibration, analyse rapide
- Études extérieures



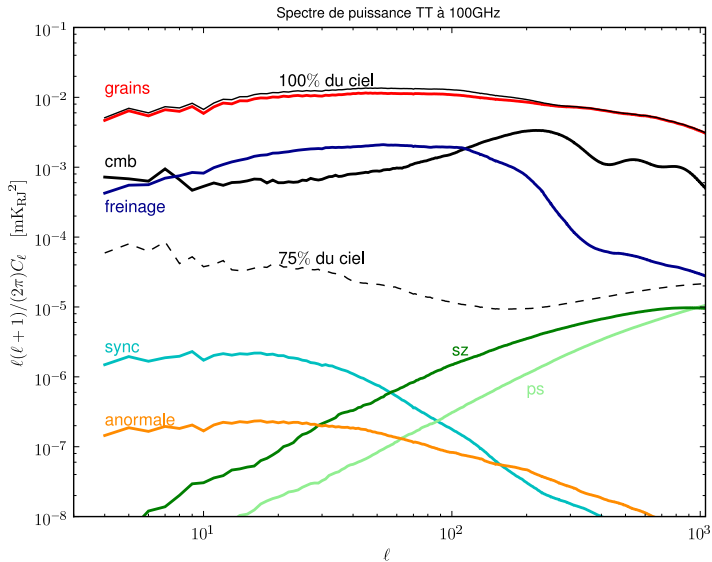




# Contributions aux échelles $>1^\circ$



# Contributions aux différentes échelles













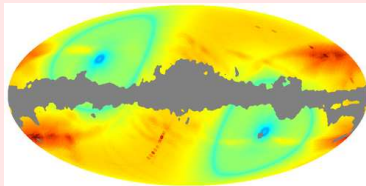
# Estimation spectrale

## Ciel complet sans bruit

Pseudo-spectre :  $\hat{C}_\ell = \frac{1}{2\ell+1} \sum_m |a_{\ell m}|^2$

## Problème

- Données incomplètes
- Bruit inhomogène



## Alternative

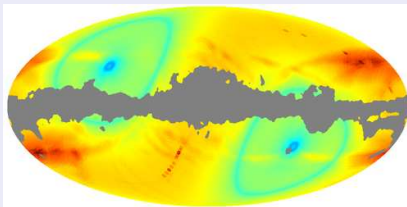
$$\hat{C}^{(j)} := \sum_k w_k^{(j)} \left\{ \left( \gamma_k^{(j)} \right)^2 - \left( n_k^{(j)} \right)^2 \right\}$$

$$w_k^{(j)}(\bar{C}^{(j)}) := \left( \bar{C}^{(j)} + \left( n_k^{(j)} \right)^2 \right)^{-2} \left[ \sum_{k'} \left( \bar{C}^{(j)} + \left( n_{k'}^{(j)} \right)^2 \right)^{-2} \right]^{-1}$$

# Caractérisation de la méthode

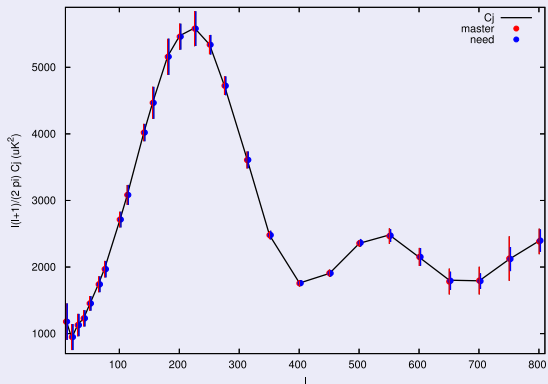
## Simulation Monte-Carlo

- Erreur quadratique moyenne de l'estimateur obtenue sur des simulations Monte-Carlo
- Pour un cas représentatif de l'analyse des données WMAP



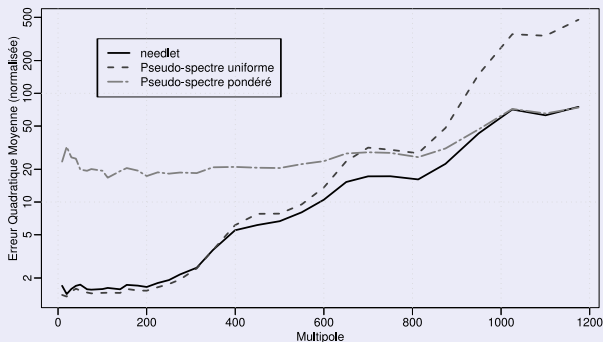
# Caractérisation de la méthode

## Premiers résultats



# Caractérisation de la méthode

## Caractérisation complète (Faÿ, Guilloux)



- Régime signal sur bruit intermédiaire
- Fusion d'expérience

Publication : Faÿ, Guilloux, Betoule, Cardoso, Delabrouille, Le Jeune, *Phys. Rev. D* (2008)



# Plan

- 1 Problématique
  - Contexte
  - Le problème des avant-plans
  
- 2 Modélisation de l'émission millimétrique
  - Contexte
  - Le PLANCK Sky Model (PSM)
  
- 3 **Analyse de données**
  - Estimation spectrale en needlets
  - **Extraction des sources ponctuelles**
  - Avant-plans et détection des modes B
  
- 4 Conclusion
  - Conclusion
  - Perspectives



# Problématique de la séparation des sources ponctuelles

## Théorie de la détection et de l'estimation

Source unique de forme connue dans un bruit gaussien stationnaire :  
Le filtre optimal est le *filtre adapté* (anglais Matched Filter)

## Mise en pratique délicate

- Arrière-plan non stationnaire
- Statistique non gaussienne
- Confusion ?

## Proposition

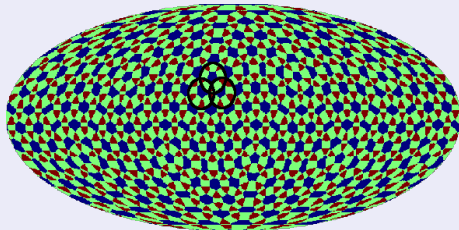
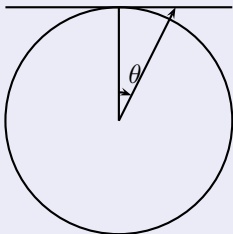
- Division du ciel en zones de dimensions restreintes
- Estimation empirique du spectre de l'arrière-plan





# Méthode localisée

## Découpage de la sphère et projection sur un plan tangent



## Pipeline

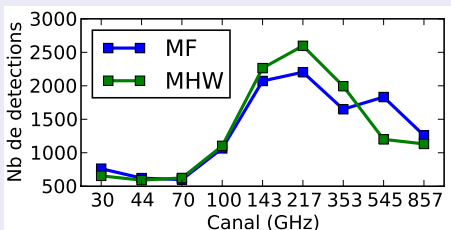
- 1 Projection sur la collection de plans tangents
- 2 Filtrage adapté et détection des sources fortes (seuil  $20\text{-}\sigma$ )
- 3 Masquage des sources fortes
- 4 Filtrage adapté et détection des sources faibles (seuil  $5\text{-}\sigma$ )
- 5 Association des catalogues



## Résultats pour la détection

### Data challenge séparation de composantes

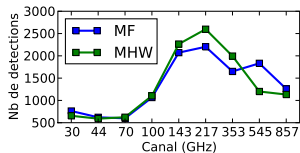
- Simulation d'un jeu de données PLANCK
- 2 méthodes :
  - Filtre adapté (mon travail)
  - Mexican Hat Wavelet (MHW) (groupe de Santander Gonzales-Nuevo et al., 2006)
- Nombres de détection pour un taux de fausse détection de 5% :



- Publié dans Leach et al. (2008) *A&A*

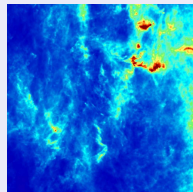


## Discussion



### Limitations

- Localisation identique pour toutes les échelles
- Compromis estimation spectrale du fond / variabilité



### Améliorer

- Localisation dans un espace d'ondelettes (MF needlets)
- Itération avec la séparation de composantes diffuses



# Estimation du flux et de la position des sources

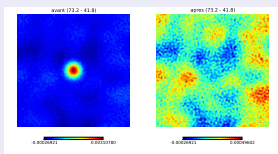
- 2 Objectifs :
- 1 Catalogue
  - 2 Soustraction pour CMB

## Ajustement d'une forme connue $\theta$

Modèle local du bruit : iid gaussien de moyenne  $b$  et de variance  $\sigma^2$ ,  
 $n(\xi) \sim \mathcal{N}(b, \sigma^2)$  :

$$\chi^2(\xi_0, A, b) = \sum_i^{N_i} \left( \frac{x(\xi_i) - b - \theta(\xi_i; \xi_0, A)}{\sigma} \right)^2. \quad (1)$$

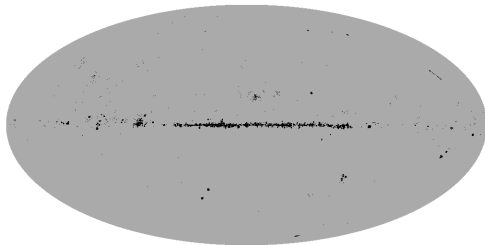
## Résultats sur le jeu de données PLANCK'Full Focal Plane'



# Traitement des sources non ajustées

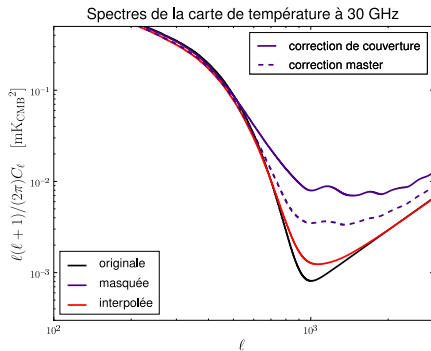
Certaines sources sont mal ajustées

- Arrière-plan complexe
- Sources étendues



# Traitement des sources non ajustées

- Les masques de sources ponctuelles ont des effets spectraux importants
- La majorité de l'effet provient de la discontinuité introduite
- Solution : restaurer au minimum la continuité par interpolation

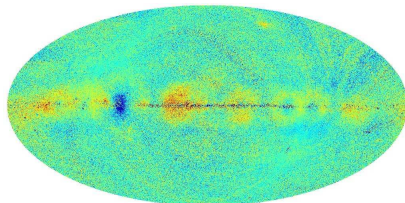
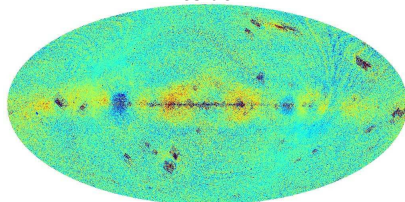


# Interaction avec la séparation de composantes diffuses

## Données test : "Full Focal Plane" dans PLANCK

- ILC needlets (moins sensible au phénomène)
- Comparaison de deux pipelines
  - Sans traitement
  - Soustraction  
Masquage / Interpolation

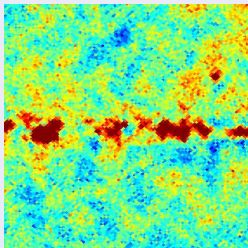
Résidu :



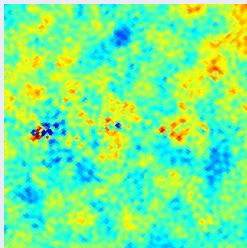
# Traitement des sources ponctuelles dans WMAP

## Zoom sur le plan galactique

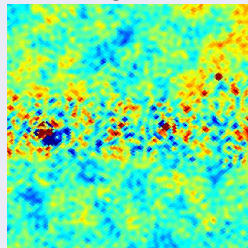
WMAP



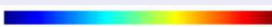
ADAMIS



Tegmark



-0.50



0.50

(45.0, 0.0) Galactic

Publié dans Delabrouille et al. (2008), *A&A*





# Plan

- 1 Problématique
  - Contexte
  - Le problème des avant-plans
- 2 Modélisation de l'émission millimétrique
  - Contexte
  - Le PLANCK Sky Model (PSM)
- 3 Analyse de données
  - Estimation spectrale en needlets
  - Extraction des sources ponctuelles
  - Avant-plans et détection des modes B
- 4 Conclusion
  - Conclusion
  - Perspectives



# Étude de mission CMBPol

L'inflation prédit l'existence d'un fond d'ondes gravitationnelles

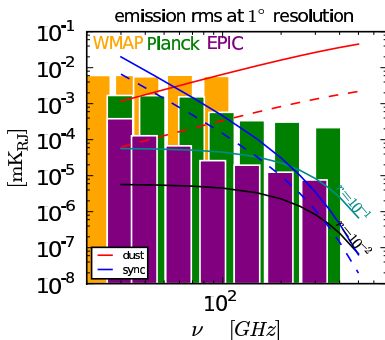
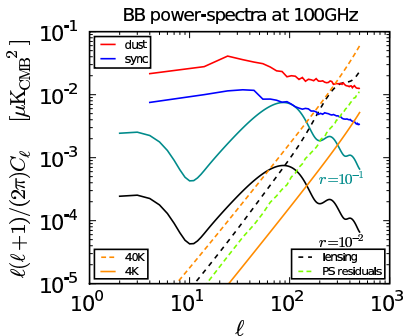
- Modes tenseurs  $\rightarrow$  motif de polarisation de type B
- Polarisation du CMB : un moyen unique de détection de ces modes
- Mesure de  $r = T/S \rightarrow$  contraintes sur la physique de l'inflation

## Contexte expérimental

- $r \lesssim 0.22$  (WMAP - Komatsu et al. 2008)
- Nombreuses expériences (dont PLANCK)
- Motive la conception d'une quatrième mission spatiale



# Problématique de la détection des modes B



- Émission diffuse galactique est le principal contaminant
- Séparation de composantes cruciale pour la détection des modes B



## Contexte

### Deux appels à projets

- BPol
- CMBPol : Plusieurs réponses, dont EPIC (Bock et al. 2008)

### La séparation de composantes gouverne des choix critiques

- Couverture spectrale (nombre de canaux)
- Résolution et sensibilité

### Plusieurs études prospectives

- Hypothèses sur la contamination résiduelle
  - Verde (2006), Gratton (2009), soustraction de templates
  - Amblard (2007), ILC
- Modèle physique des avant-plans
  - Dunkley (2008)
- Notre approche : Modèle flexible sans paramétrisation physique des avant-plans
  - Betoule (2009)

# Séparation aveugle des composantes avec SMICA

SMICA (Delabrouille et al. 2003, Patanchon 2003, Cardoso, Le Jeune, Delabrouille, Betoule, Patanchon 2008)

- Ajustement d'un modèle des matrices de covariance spectrales
- Analyse en composantes indépendantes
- Modèle flexible

Adaptable à la mesure de  $r$

- Analyse sur les cartes de modes B
- Forme générale du  $\mathcal{S}_\ell$  spectre  $BB$  connue

$$\mathbf{R}_\ell^{\text{cmb}} = \mathcal{C}_\ell \mathbf{A}_{\text{cmb}} \mathbf{A}_{\text{cmb}}^\dagger \longrightarrow \mathbf{R}_\ell^{\text{cmb}} = r \mathcal{S}_\ell \mathbf{A}_{\text{cmb}} \mathbf{A}_{\text{cmb}}^\dagger$$



## Modèle flexible

### 3 composantes indépendantes

$$\mathbf{R}_\ell = \mathbf{R}_\ell^{\text{cmb}} + \mathbf{R}_\ell^{\text{avant-plan}} + \mathbf{R}_\ell^{\text{bruit}}$$

- $\mathbf{R}_\ell^{\text{cmb}} = \mathcal{C}_\ell \mathbf{A}_{\text{cmb}} \mathbf{A}_{\text{cmb}}^\dagger \longrightarrow \mathbf{R}_\ell^{\text{cmb}} = r \mathcal{S}_\ell \mathbf{A}_{\text{cmb}} \mathbf{A}_{\text{cmb}}^\dagger$
- $\mathbf{R}_\ell^{\text{bruit}}$  connu

### Modèle des avant-plans

- Mélange linéaire de  $N_{\text{obs}}$  "composantes" corrélées
- Matrice  $\mathbf{R}_\ell^{\text{avant-plan}}$  de rang  $N_{\text{obs}} < N_{\text{freq}}$

$$S_\nu(\xi) = S_0(\xi) \left( \frac{\nu}{\nu_0} \right)^{\beta(\xi)}$$

$$S_\nu(\xi) \approx S_0(\xi) \left( \frac{\nu}{\nu_0} \right)^\beta + S_0(\xi) \delta\beta(\xi) \left( \frac{\nu - \nu_0}{\nu_0} \right) \left( \frac{\nu}{\nu_0} \right)^\beta$$

Représentation approximative de toute émission mesurée à un nombre fini de fréquences.



# Résultats principaux

## Étude sur des simulations PSM

### PLANCK : Borne pessimiste $r \sim 0.1$

- Mission nominale
- Seul *BB* est exploité

### Sol (Clover)

- Peu de fréquences (90,150,225) → avant-plans limitent  $r \sim 0.09$
- Sol+PLANCK  $r \sim 0.035$

### EPIC : 3 versions, coût croissant →

Version	EPIC-LC	EPIC-2m	EPIC-CS
Diamètre	0.3	2 m	3 m
Canaux (GHz)	7 (30-300)	9 (30-800)	8 (30-500)
Durée	2 ans	4 ans	4 ans
sensibilité	Comparable $\sim 1.5\mu K\sqrt{s}$		
$r \sim 0.001$	3- $\sigma$	6- $\sigma$	5.5- $\sigma$



# Résultats principaux

## Étude sur des simulations PSM

### PLANCK : Borne pessimiste $r \sim 0.1$

- Mission nominale
- Seul *BB* est exploité

### Sol (Clover)

- Peu de fréquences (90,150,225) → avant-plans limitent  $r \sim 0.09$
- Sol+PLANCK  $r \sim 0.035$

### EPIC : 3 versions, coût croissant →

Version	EPIC-LC	EPIC-2m	EPIC-CS
Diamètre	0.3	2 m	3 m
Canaux (GHz)	7 (30-300)	9 (30-800)	8 (30-500)
Durée	2 ans	4 ans	4 ans
sensibilité	Comparable $\sim 1.5\mu K\sqrt{s}$		
$r \sim 0.001$	3- $\sigma$	6- $\sigma$	5.5- $\sigma$

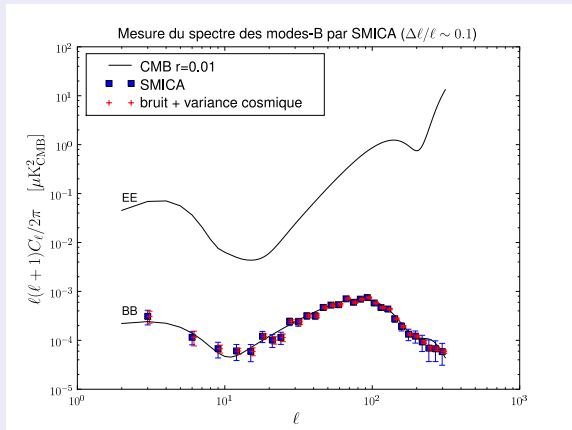




# Mesure du spectre

## L'objectif scientifique évolue pour EPIC-IM 4K

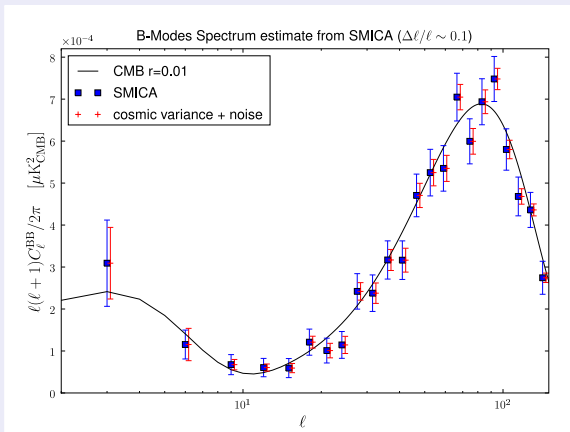
- Similaire à EPIC-2m
- Télescope entièrement refroidi à 4K



# Mesure du spectre

## L'objectif scientifique évolue pour EPIC-IM 4K

- Similaire à EPIC-2m
- Télescope entièrement refroidi à 4K



## Résultats fiables ?

### Les simulations sont-elles représentatives de la réalité ?

- Pas de modèle physique des données
- Rang  $N_{\text{comp}}$  de la matrice  $\mathbf{R}^{\text{avant-plan}}$  réelle comparable à la simulation

### Exploiter le PSM pour faire varier les simulations

- Synchrotron :  $S_{\nu}^X(\xi) = S_{\nu_0}^X(\xi) \left( \frac{\nu}{\nu_0} \right)^{\beta_s(\xi) + \alpha C(\xi) \log(\nu/\nu_1)}$  pour

$$\alpha = -0.3$$

- Poussière : Spectre de puissance  $\times 2$

→ La qualité de l'ajustement ne varie pas, la dimension  $N_{\text{comp}}$  nécessaire non plus.

Betoule et al. (2009) *A&A*, Dunkley et al. (2009), Bock et al. (2009)



# Bilan

## Mise à disposition du Planck Sky Model

- Simulation aisée
- Possibilité de réaliser des simulations Monte-Carlo
- Standard pour PLANCK

## Estimateur Spectral en Needlets

Gain pour les régimes intermédiaires/fusion d'expériences

## Traitement des sources ponctuelles

- Amélioration notable de la reconstruction du CMB
- Implémenté pour PLANCK
- Appliqué à WMAP

## Mesure des modes tenseurs

- Tenant compte de l'incertitude sur les avant-plans
- Application à la conception de mission

# Perspectives

## Dans la continuité

- Modélisation des données `PLANCK` dans le PSM
- Estimateur Spectral en Needlets croisé (Application possible à WMAP)
- Estimateur Spectral en Needlets polarisé
- Filtre adapté dans l'espace needlets (SZ)
- Application directe de la détection des modes B à `PLANCK`

## À plus long terme

Intégration séparation de composantes diffuses et ponctuelles au sein d'un pipeline needlets

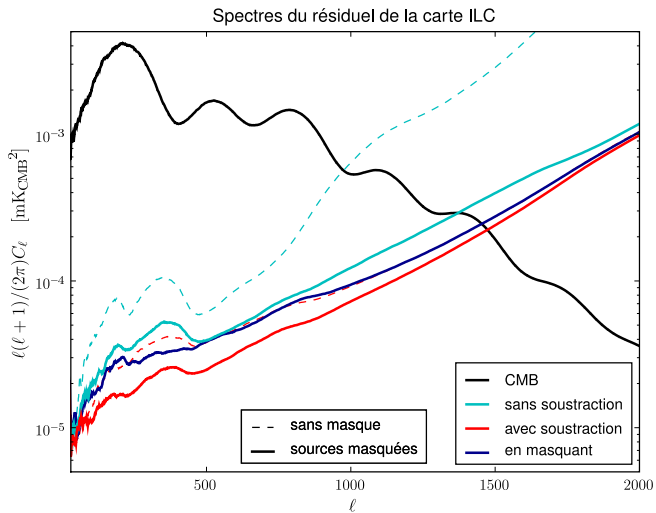


# Bibliographie

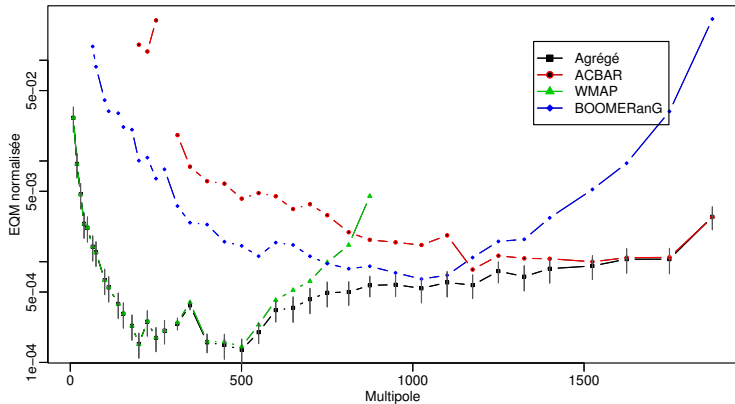
-  M. Betoule, E. Pierpaoli, J. Delabrouille, M. Le Jeune, J.F. Cardoso.  
*A&A*. Volume 503, Issue 3, 2009, pp.691-706
-  Delabrouille, J. ; Cardoso, J.-F. ; Le Jeune, M. ; Betoule, M. ; Fay, G. ; Guilloux, F.  
*A&A*, Volume 493, Issue 3, 2009, pp.835-857
-  Faÿ, G. ; Guilloux, F. ; Betoule, M. ; Cardoso, J.-F. ; Delabrouille, J. ; Le Jeune, M.  
*Physical Review D*, vol. 78, Issue 8
-  Cardoso, Jean-François ; Le Jeune, Maude ; Delabrouille, Jacques ; Betoule, Marc ; Patanchon, Guillaume  
*IEEE Journal of Selected Topics in Signal Processing*, vol. 2, issue 5, pp. 735-746
-  Leach, S. ; Cardoso J.-F. et al.  
*A&A*, Volume 491, Issue 2, 2008, pp.597-615



# Prétraitement des sources ponctuelles et ILC



# Fusion d'expériences





## SMICA

$$\mathbf{x}(\xi) = \sum_{c=1}^C \mathbf{x}^c(\xi), \mathbf{a}_{\ell m} = \sum_{c=1}^C \mathbf{a}_{\ell m}^c \quad (2)$$

$$\mathbf{R}_\ell(\theta) = \sum_{c=1}^C \mathbf{R}_\ell^c(\theta^c). \quad (3)$$

$$-2 \ln \mathcal{L} = \sum_{\ell} (2\ell + 1) f_{\text{ciel}} K(\hat{\mathbf{R}}_\ell, \mathbf{R}_\ell(\theta)) + \text{const.} \quad (4)$$

où  $\hat{\mathbf{R}}_\ell$  est l'estimateur de  $\mathbf{R}_\ell$  :

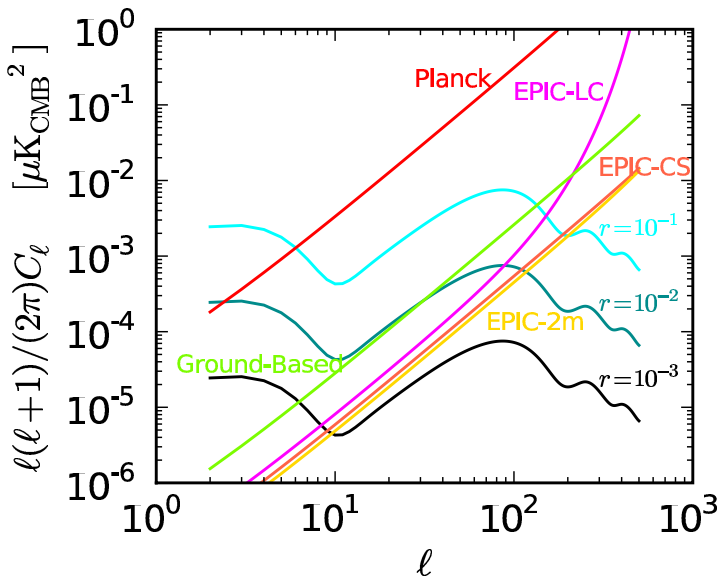
$$\hat{\mathbf{R}}_\ell = \frac{1}{2\ell + 1} \frac{1}{f_{\text{ciel}}} \sum_{m=-\ell}^{\ell} \mathbf{a}_{\ell, m} \mathbf{a}_{\ell, m}^\dagger \quad (5)$$

et où  $K(\cdot, \cdot)$  est une mesure de la divergence entre deux matrices définies positives :

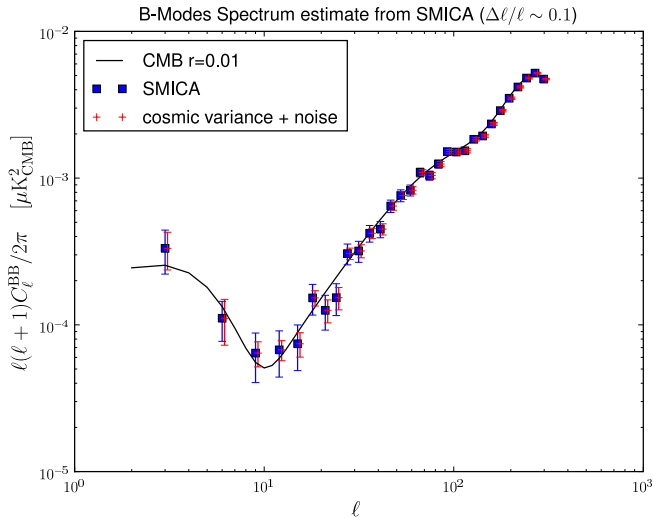
$$K(\hat{\mathbf{R}}, \mathbf{R}) = \frac{1}{2} \left[ \text{trace}(\mathbf{R}^{-1} \hat{\mathbf{R}}) - \log \det(\mathbf{R}^{-1} \hat{\mathbf{R}}) - F \right]. \quad (6)$$



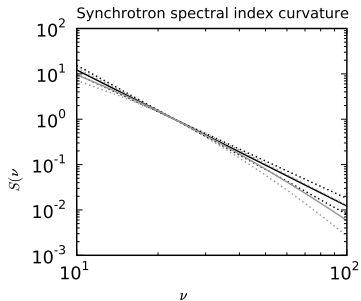
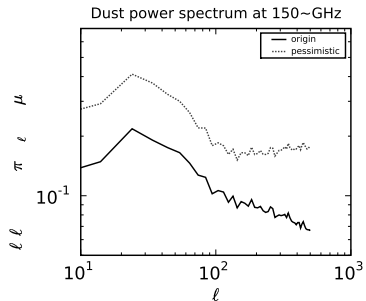
# Sensibilité



# Lensing



# Variations



## Sources ponctuelles - lensing

$r$	$r^{\text{no-ps}}$	$r^{\text{ideal}}$	$r^{\text{no-model}}$	$\sigma_r^{\text{no-ps}}$	$\sigma_r^{\text{ideal}}$	$\sigma_r^{\text{no-model}}$
0.001	$1.07 \cdot 10^{-3}$	$1.04 \cdot 10^{-3}$	$2.00 \cdot 10^{-3}$	$1.84 \cdot 10^{-4}$	$1.91 \cdot 10^{-4}$	$2.49 \cdot 10^{-4}$

Expérience	sans lensing			lensing		
	$\sigma_r/r$	$\sigma_r^{\ell \leq 20}/r$	$\sigma_r^{\ell > 20}/r$	$\sigma_r/r$	$\sigma_r^{\ell \leq 20}/r$	$\sigma_r^{\ell > 20}/r$
EPIC-CS	0.18	0.26	0.29	0.21	0.26	0.38
Ciel profond	0.13	—	—	1.1	—	—

