Measuring primordial B modes in presence of astrophysical foregrounds

Jacques Delabrouille
APC - Paris

The component separation problem

- Recognised as a key issue for a B-mode polarisation space mission
- If one targets to measure reliably (at good S/N) r=T/S levels undetectable with Planck and from the ground, one has to reject galactic foregrounds by a factor of 50-100 on large scales.
- Point sources cannot be neglected. The brightest ones have to be blanked. Contamination from the others should be corrected for.
- The key issue (for galactic foregrounds in particular) is whether we can predict (in a way or another), in a set of CMB channels, polarised FG emission with confidence at the 1-2 % level.

Galactic foregrounds

- Highly complex, with many uncertainties.

- Main foregrounds : Synchrotron and dust
  - Uncertainties concerning some components (anomalous dust)
  - All components can be polarised (at least locally in special regions) at the few percent level
  - Remain open minded : surprises are not excluded

- Two main approaches for subtracting the galactic emission
  - Either we can model it physically to within 1% error and subtract it (we are far from that)
  - Or we can use statistical methods, which use the independence of CMB from foregrounds to extract the CMB

- Our understanding will improve drastically with the analysis of the Planck data

PSM predictions – B modes

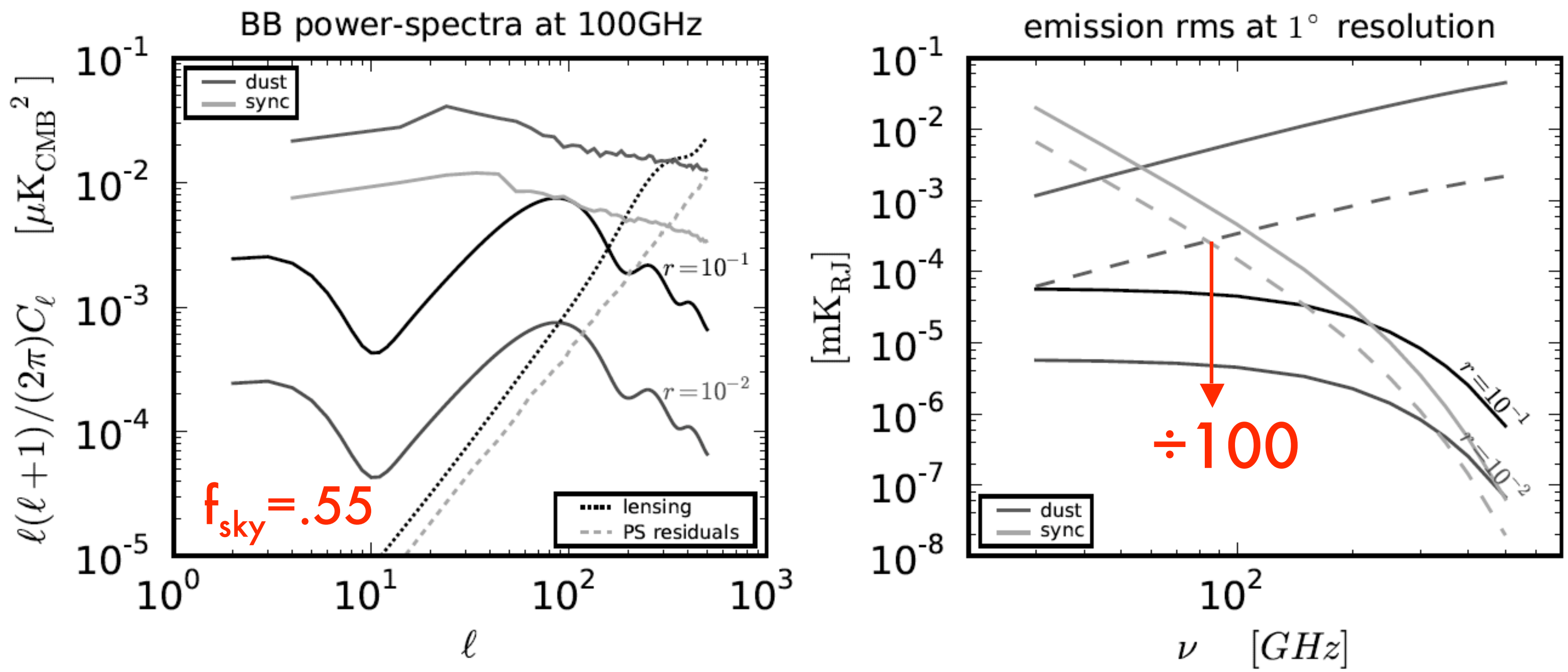
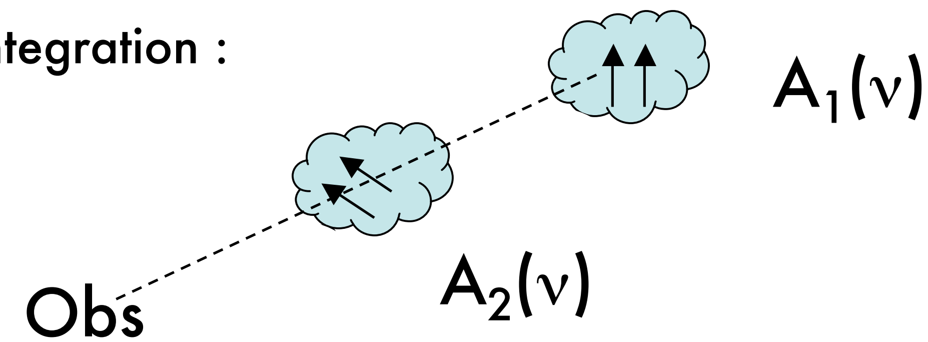


Fig. 1. Respective emission levels of the various components as predicted by the PSM. Left: predicted power spectra of the various components at 100 GHz, compared to CMB and lensing level for standard cosmology and various values of r ($\tau = 0.07$, and other cosmological parameters follow Dunkley et al. (2008a)). The power spectra of diffuse galactic foregrounds are computed using the cleanest 55% of the polarised sky. The power spectrum from residual point sources is computed assuming that all sources brighter than 500 mJy (in temperature) in one of the Planck channels have been cut out. Right: typical frequency-dependence of the contributions to B-type polarisation of CMB, synchrotron and dust, at 1 degree resolution. The dashed lines correspond to the mean level of fluctuation as computed outside the mask used for the power spectra shown in the right panel.

CMB extraction with blind methods

- Blind separation tools permit to extract the CMB with practically no assumption about the foregrounds
- The key issue is not the level of galactic emissions, but their coherence (from channel to channel) and emission law(s). It is much easier to remove a strong foreground which scales rigidly with frequency, than a faint foreground mostly uncorrelated from channel to channel.
- An example is faraday rotation, which breaks the coherence between 1 GHz polarisation and 100 GHz polarisation.
- Decoherence by LOS integration :



Number of foreground components ?

- If we have a number *D* of "detectors" (i.e. channels)

$$x_d = c + f_d + n_d$$

CMB — foregrounds — noise

- Consider the matrix $R^{fg} = \langle f_d . f_{d'} \rangle$.
  - If foreground emission is completely incoherent between detectors, R^{fg} is of rank D (and is diagonal)
  - If foreground emission is completely coherent (same template at all frequencies), R^{fg} is of rank 1
  - If we are lucky enough that R^{fg} is of rank $F<D$ then we can (in principle) recover the CMB with no foreground residual

CMB with no foregrounds ?

- R^{fg} is a symmetric, positive matrix
- Let M be a matrix such that $R^{fg} = M \Delta M^T$

Consider

$$y_d = M^{-1} x_d = M^{-1} c + M^{-1} f_d + M^{-1} n_d$$

We have

$$\langle y_d \, y_d^T \rangle = M^{-1} C_l M^{-T} + \Delta + M^{-1} N_l M^{-T}$$

If Δ is not full rank then the component separation problem is solved (for the CMB) !

If we knew R^{fg}, we could find M, and solve the problem

Measuring *r* and C_l^{BB} with SMICA

- Reminder: SMICA implements a Maximum Likelihood estimate of parameters defining the power spectra and frequency scalings of a set of astrophysical components (i.e. C_l and R^{fg} – and N_l).

- These components can be specified (e.g. with physical parameters, such as cosmological parameters) or unspecified (components being described empirically by $A(\nu)$, C_l

- SMICA matches model covariances to empirical covariances:

R_{ij}(parameters) is a fit to $\langle x_i . x_j \rangle$

Any set of parameters describing the auto and cross power-spectra of the astrophysical components

Measured, empirical auto and cross power-spectra of the observations

Numerical experiments

- We use PSM simulations: polarised CMB+foregrounds
- We investigate a set of observations with different instruments
- We assume $C_l^{BB} = r.\mathcal{S}_\ell$, the shape $\mathcal{S}_\ell$ being known, but not r
- Very flexible foreground model : any set of F<D correlated components, with free auto and cross spectra, with free frequency scalings.

Unknown foreground contribution can be modelled as the mixed contribution of D correlated sources:

$$\mathsf{R}_\ell^{\mathrm{fg}} = \mathsf{A}\Sigma_\ell\mathsf{A}^\dagger \quad (17)$$

where A is a $F \times D$ mixing matrix and Σ_ℓ is the $D \times D$ spectral covariance matrix of the sources. The model of the spectral covariance matrix of the observations is then:

$$\mathsf{R}_\ell = r\mathcal{S}_\ell \boldsymbol{A}_{\mathrm{cmb}} \boldsymbol{A}_{\mathrm{cmb}}^\dagger + \underline{\mathsf{A}\Sigma_\ell\mathsf{A}^\dagger} + \mathsf{R}_\ell^{\mathrm{noise}}$$

Betoule et al. 2009

Toy-model instruments

Experiment	frequency (GHz)	beam FWHM (')	NET ($\mu K \sqrt{s}$)	T_{obs} (yr)	sky coverage (f_{sky})
PLANCK	30, 44, 70 100, 143, 217, 353	33, 24, 14 10, 7.1, 5, 5	96, 97, 97 41, 31, 51, 154	1.2	1
EPIC-LC	30, 40, 60 90, 135, 200, 300	155, 116, 77 52, 34, 23, 16	28, 9.6, 5.3 2.3, 2.2, 2.3, 3.8	2	1
EPIC-CS	30, 45, 70, 100 150, 220, 340, 500	15.5, 10.3, 6.6, 4.6 3.1, 2.1, 1.4, 0.9	19, 8, 4.2, 3.2 3.1, 5.2, 25, 210	4	1
EPIC-2m	30, 45, 70, 100 150, 220, 340, 500(,800)	26, 17, 11, 8 5, 3.5, 2.3, 1.5(, 0.9)	18, 7.6, 3.9, 3.0 2.8, 4.4, 20, 180(, 28k)	4	1
Ground-Based	97, 150, 225	7.5, 5.5, 5.5	12, 18, 48	0.8	0.01
Deep field	30, 45, 70, 100 150, 220, 340, 500	15.5, 10.3, 6.6, 4.6 3.1, 2.1, 1.4, 0.9	19, 8, 4.2, 3.2 3.1, 5.2, 25, 210	4	0.01

Betoule et al. 2009

Investigate *r* with B modes specifically

- Assume all other parameters obtained from CMB T and E, and other astro/cosmo probes
- Assume *r* is low enough that it does not impact measurably T and E (all the info is in B)

---

we compute the Fisher information matrix $\mathsf{I}_{i,j}(\boldsymbol{\theta})$ deriving from the maximised likelihood (14) for the parameter set $\boldsymbol{\theta} = (r, \mathsf{A}, \Sigma_1, \cdots, \Sigma_{\mathsf{Q}})$:

$$\mathsf{I}_{i,j}(\boldsymbol{\theta}) = \frac{1}{2} \sum_q w_q \,\mathrm{trace} \left(\frac{\partial \mathsf{R}_q(\boldsymbol{\theta})}{\partial \theta_i} \mathsf{R}_q^{-1} \frac{\partial \mathsf{R}_q(\boldsymbol{\theta})}{\partial \theta_j} \mathsf{R}_q^{-1} \right) \quad (19)$$

The lowest achievable variance of the r estimate is obtained as the entry of the inverse of the FIM corresponding to the parameter r:

$$\sigma_r^2 = \mathsf{I}_{r,r}^{-1} \quad (20)$$

Results

(If you do not trust the PSM, these results should be interpreted qualitatively more than quantitatively)

Where does the info come from ?

Dimensions of FG component (found by SMICA)

case	r	noise-only σ_r/r	noise-only $\sigma_r^{\ell\leq 20}/r$	noise-only $\sigma_r^{\ell>20}/r$	known foregrounds σ_r/r	known foregrounds $\sigma_r^{\ell\leq 20}/r$	known foregrounds $\sigma_r^{\ell>20}/r$	SMICA σ_r/r	SMICA $\sigma_r^{\ell\leq 20}/r$	SMICA $\sigma_r^{\ell>20}/r$	r^{est}	$l_{\text{min}} - l_{\text{max}}$	f_{sky}	D^3
PLANCK	0.3	0.075	0.17	0.084	0.1	0.2	0.12	0.15	0.22	0.2	**0.26**	2 - 130	0.95	3
	0.1	0.17	0.25	0.22	0.23	0.34	0.32	0.29	0.34	0.55	0.086			
EPIC-LC	0.01	0.019	0.084	0.019	0.05	0.18	0.053	0.079	0.18	0.1	**0.0098**	2 - 130	0.86	4
	0.001	0.059	0.15	0.064	0.27	0.4	0.38	0.37	0.43	0.82	0.00088			
EPIC-2m	0.01	0.016	0.083	0.016	0.027	0.12	0.027	0.032	0.11	0.036	**0.0096**	2 - 300	0.87	4
	0.001	0.051	0.14	0.055	0.14	0.25	0.16	0.16	0.24	0.24	**0.001**			
EPIC-CS	0.01	0.017	0.084	0.017	0.029	0.12	0.03	0.036	0.11	0.041	**0.0096**	2 - 300	0.87	4
	0.001	0.058	0.15	0.063	0.15	0.27	0.19	0.18	0.26	0.29	**0.00098**			
Ground-based	0.1	0.083	–	–	0.15	–	–	0.24	–	–	**0.11**	50 - 300	0.01	2
	0.01	0.18	–	–	0.8	–	–	1.6	–	–	0.018			
Grnd-based+Planck	0.01	0.18	–	–	0.51	–	–	0.69	–	–	0.0065	50 - 300	0.01	2
Deep field mission	0.001	0.082	–	–	0.1	–	–	0.13	–	–	**0.00092**	50 - 300	0.01	4

How much foregrounds are a problem ?

How much SMICA is ineffective ?

Betoule et al. 2009

Why don't we get perfect separation ?

- R^{fg} is full rank, but the smallest components are very local (not "seen" on average statistics)
  - Mask peculiar regions
  - Use local statistics
  - Increase number of bands

- There is not enough statistics to measure R^{fg} perfectly
  - Measure as many independent modes as possible

- There is a trade-off between residual noise and residual foregrounds

- Note that in real life, multiplicative errors are a worry (e.g. beam uncertainties, + possibly calibration errors)

CMB extraction with blind methods

- In order to achieve residuals of order 1%, templates representing galactic emission must be (loosely speaking) 99% correlated from channel to channel.
  - This is unlikely to be the case over a large frequency range with one single template for each emission.
  - The alternative is to assume that multifrequency galactic emission can be represented with a set of templates (possibly correlated in space) so that the residual of this representation is < 1%. The number of templates needed sets the number of frequency channels required.
  - This is what is discussed here as the 'dimension of the galactic component'

- We do not have the information now
  - The PSM is far too simplistic to investigate this in detail
  - Details of foreground polarised emission remain to be understood
  - Safe option is : *many frequencies, many modes !*