Semi-blind component separation for measurement of CMB B-mode polarisation

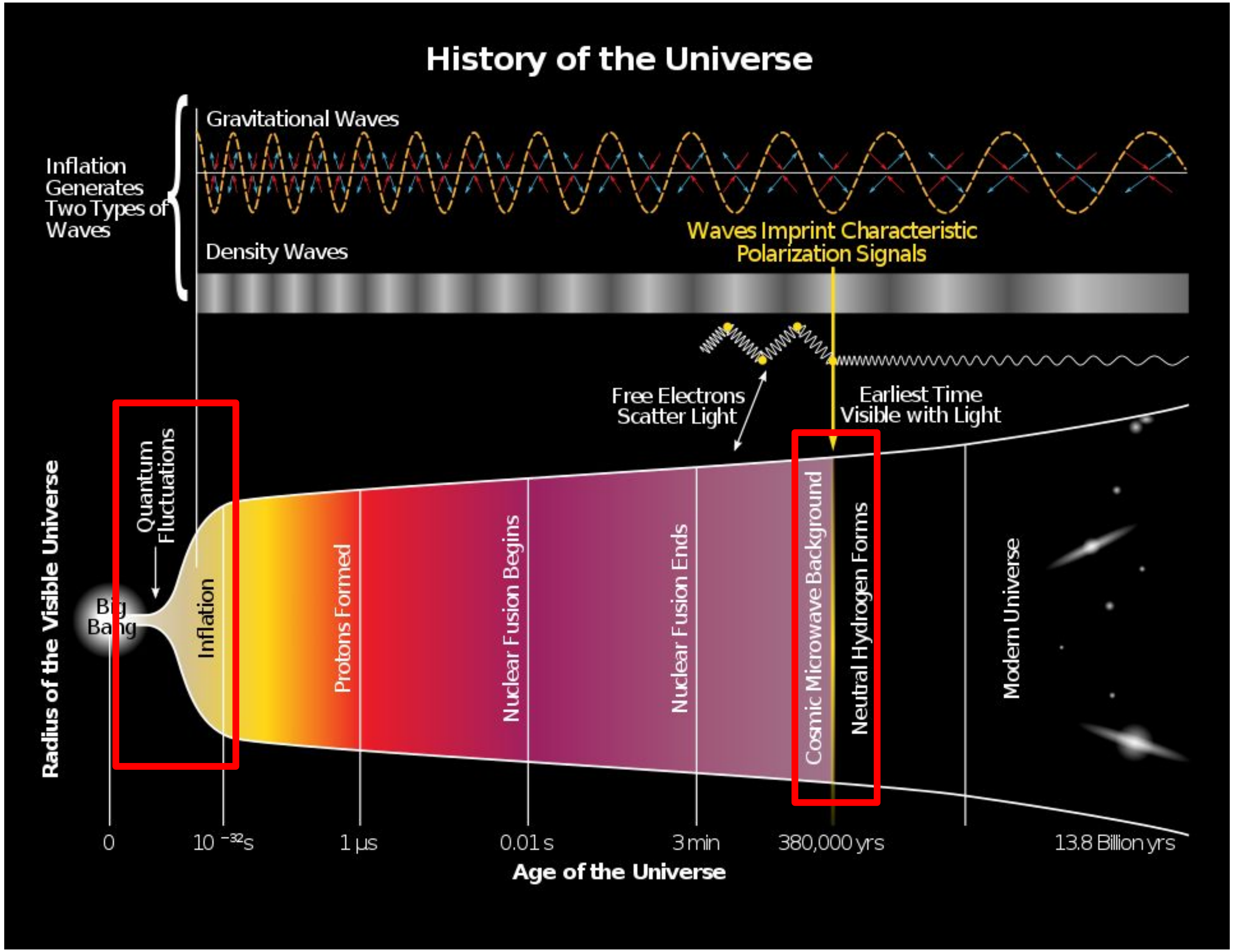
By: Tran Hoang Viet
Supervisor: Guillaume Patanchon - APC, Université Paris Cité

21st Rencontres du Vietnam
15th August 2025

Contents

1. Scientific context

Inflation & primordial perturbation



Inflation: A phase of exponential expansion of the universe, lasted for only $10^{-36} - 10^{-33}$ seconds.

Inflation generates the primordial perturbations, seeding the structures of the universe we see today

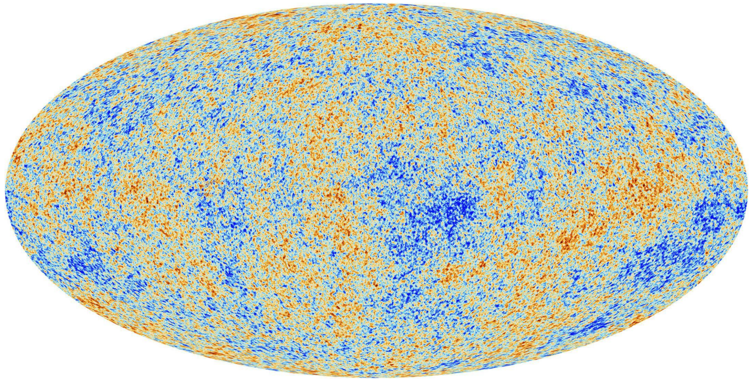
Inflation must have been driven by some process or mechanism not yet seen => **Search for new physics**

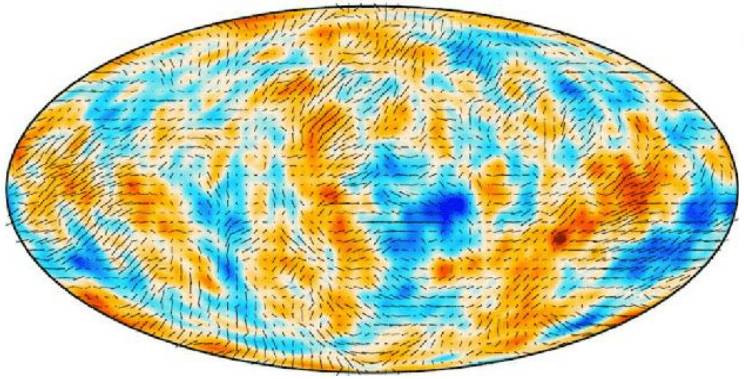
Perturbations evolve till some time after inflation, neutral hydrogen forms, emission of photons from the primordial plasma => **CMB emission**

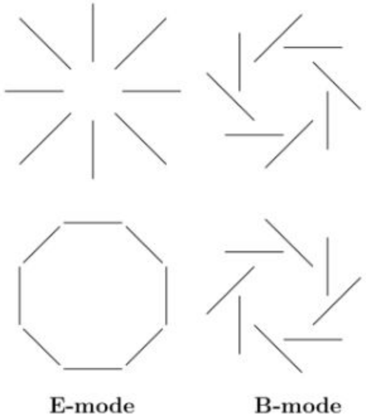
The CMB pattern that we can see today was **imprinted** with the **characteristics of primordial perturbation**

The Cosmic Microwave Background

<u>Cosmic Microwave Background</u>: first light of the universe – photons from primordial universe



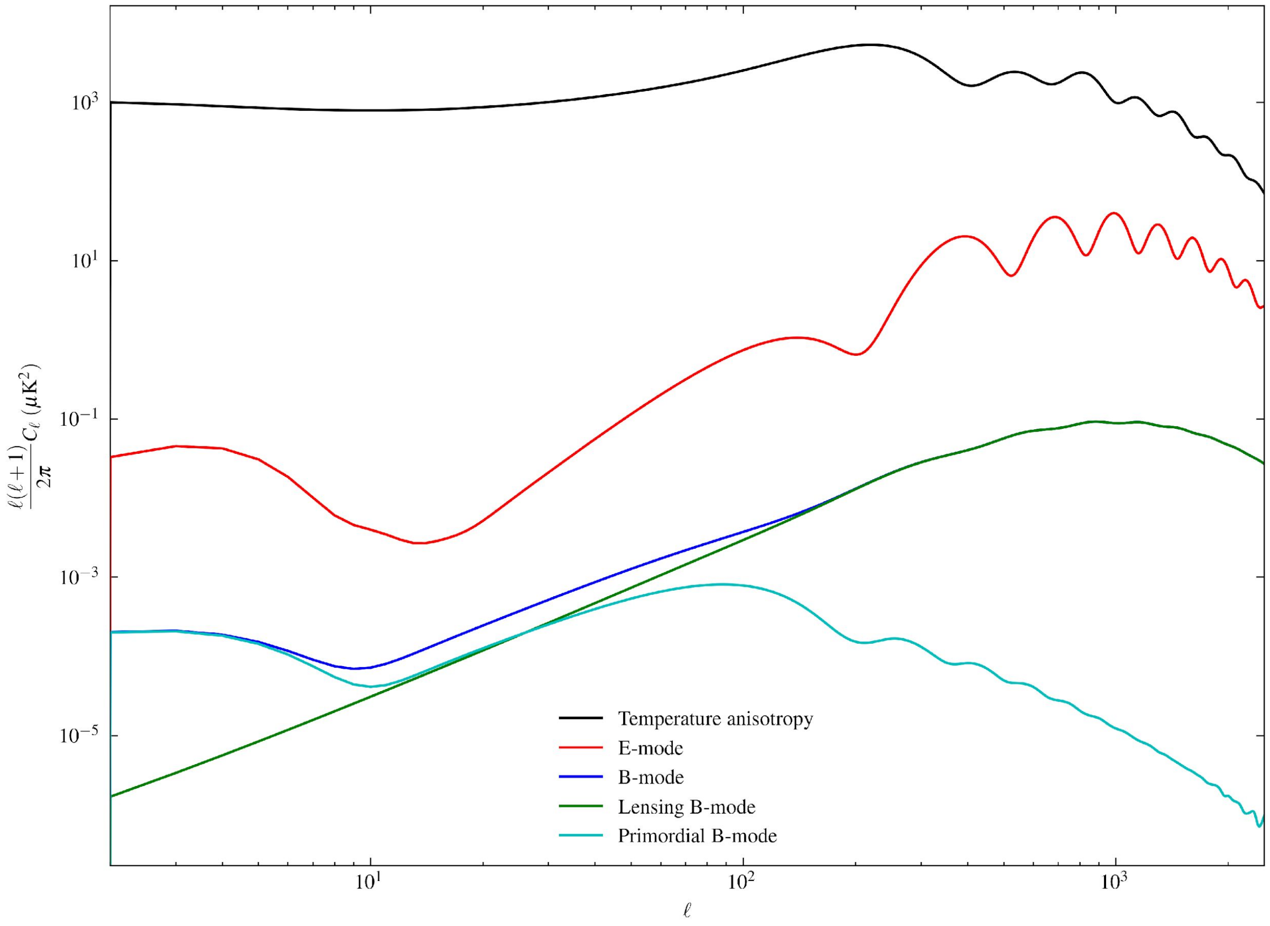




Monopole of 2.725 K

Temperature anisotropy of about 100 μK from perturbation in the primordial photon/baryon plasma

Polarization from anisotropic scattering with electrons and primordial gravitational wave, often decomposed into positive parity E-mode & negative parity B-mode



The **CMB angular power spectra**: shows the "power" of the perturbation with respect to the multipole which is related to the inverse of the angular separation in the sky

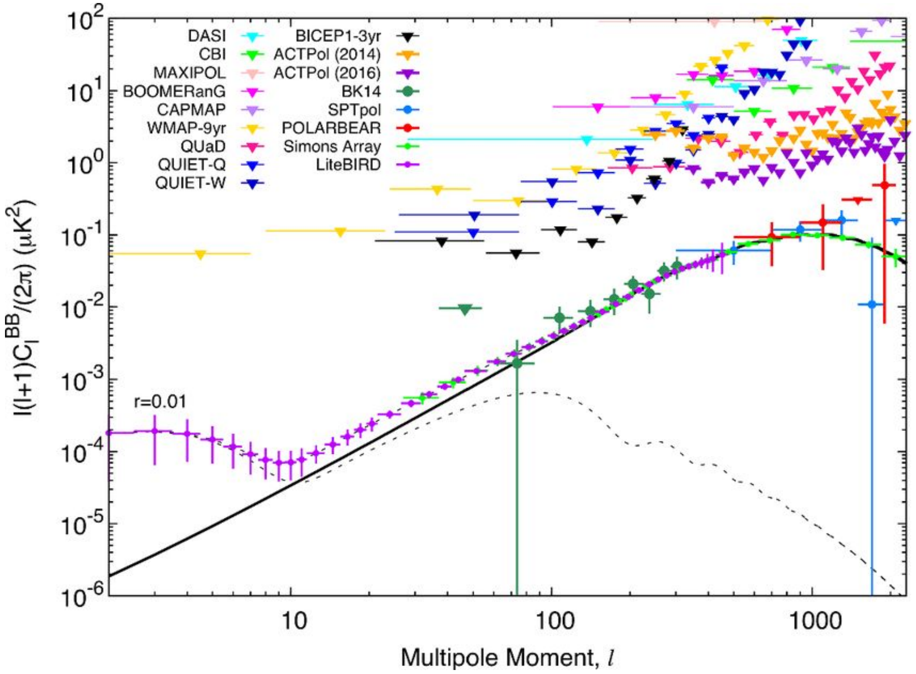
Cosmology and the LiteBIRD mission

Perturbation can be decomposed into: **Scalar** (Newtonian potential), Vector (Destroyed) & **Tensor** (Gravitational Wave)

Scalar perturbation can only generate E-mode and temperature anisotropy
The tensor-to-scalar ratio – directly related to the energy scale of inflation:

$$r = \frac{\Delta_h^2}{\Delta_k^2} = 16\epsilon$$

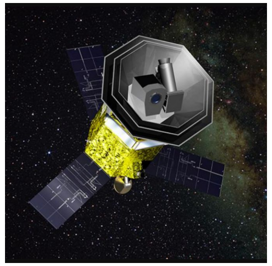
r is fully diluted in the anisotropy & E-mode due to the large amplitude of scalar perturbation
=> B-mode is our direct look into primordial gravitational waves



Hazumi, M. et al. (2019). “LiteBIRD: A Satellite for the Studies of B-Mode Polarization and Inflation from Cosmic Background Radiation Detection.”

Currently, we have only the **upper limit** $r < 0.032$ at 95% confidence

(Tristram et al. 2022, combining BK18 and Planck PR4)



LiteBIRD is a satellite mission under development by an international collaboration led by JAXA

⇒ Future with LiteBIRD: Determining r at a precision of:

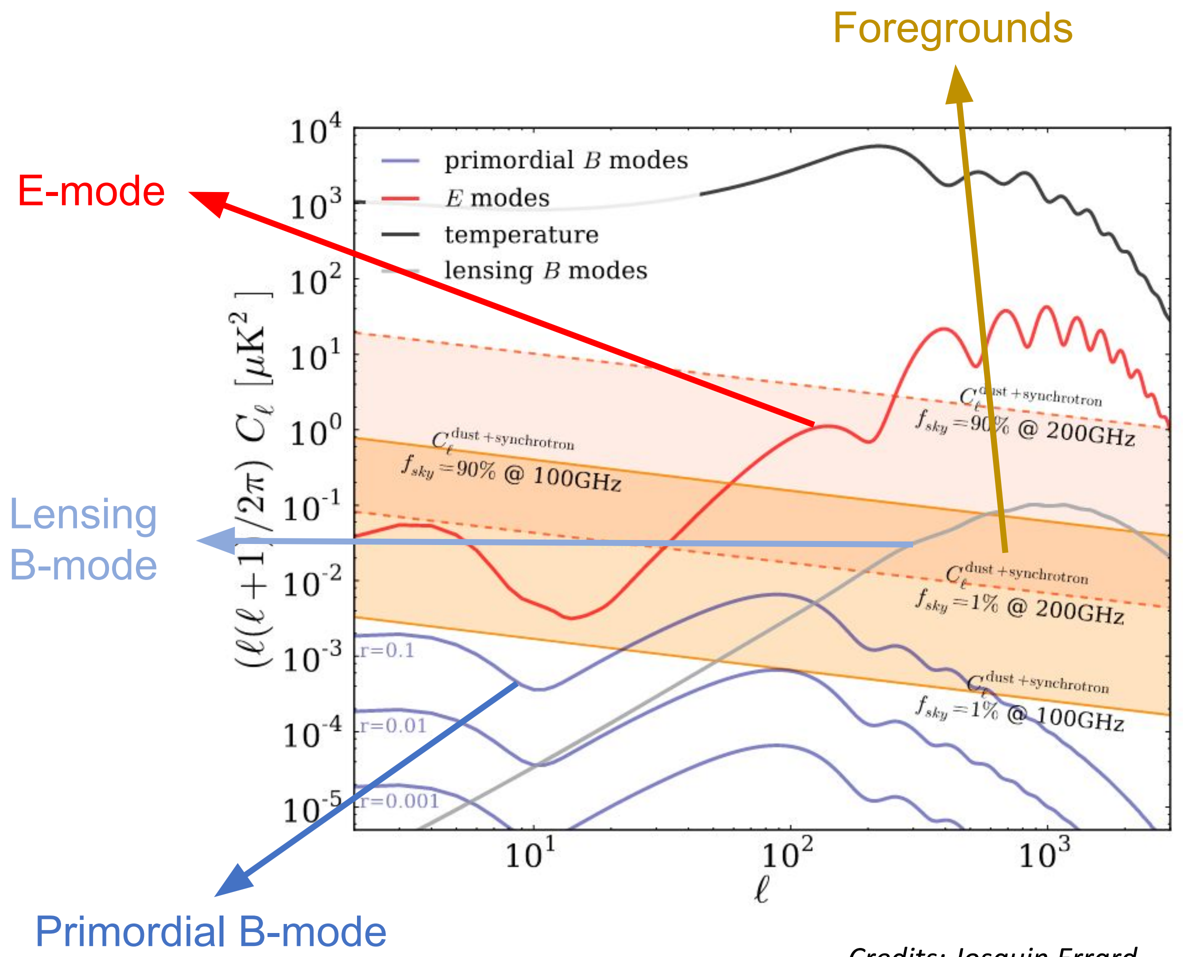
$$\Delta r < 0.001$$

Requires unprecedented removal of systematics and other contamination sources

Foregrounds emission and component separation

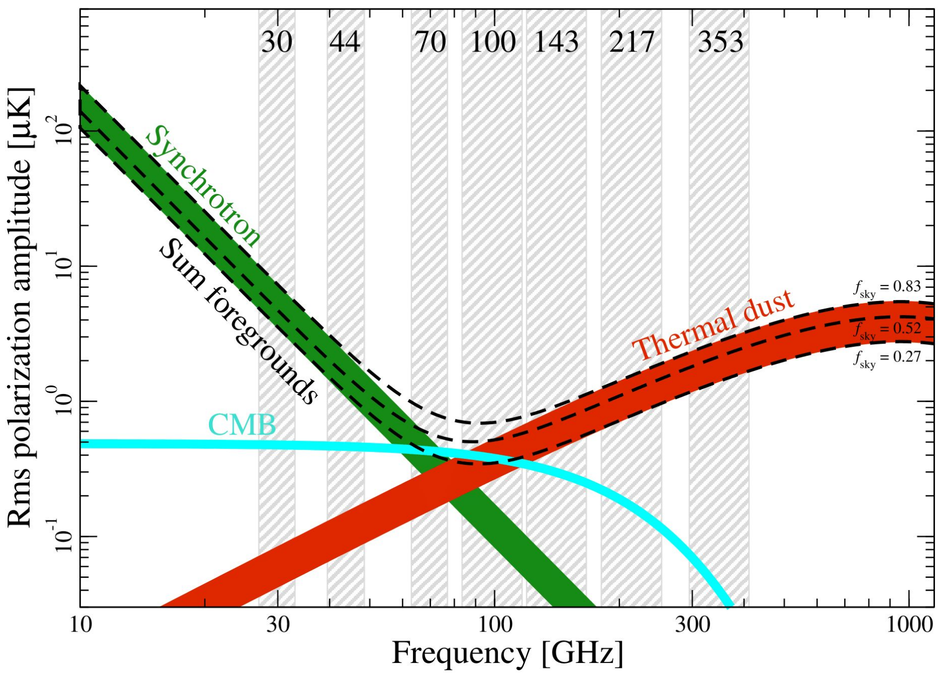
<u>Context</u>: Other sources of contamination are foreground emissions: dust & synchrotron polarization
These are the most important effects when it comes to polarization measurement

⇨ **<u>Component separation method are needed !</u>**



Credits: Josquin Errard

These methods utilize the fact that we have measurement of the components in multiple different frequencies

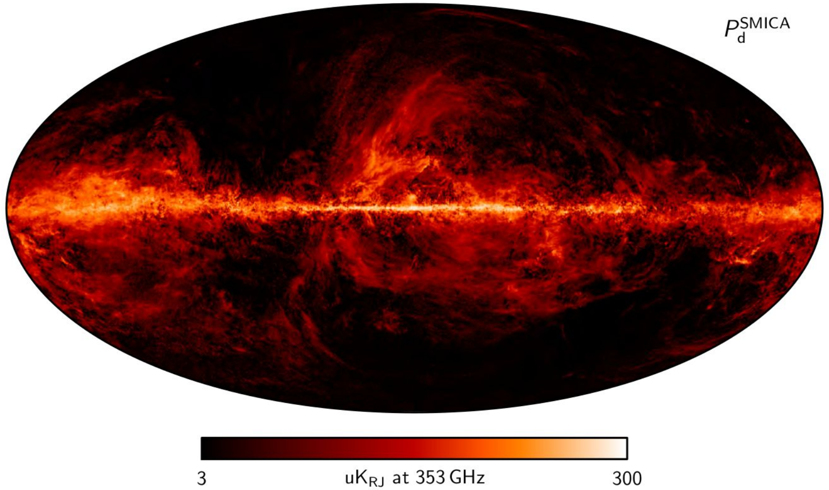


Foregrounds emission and component separation

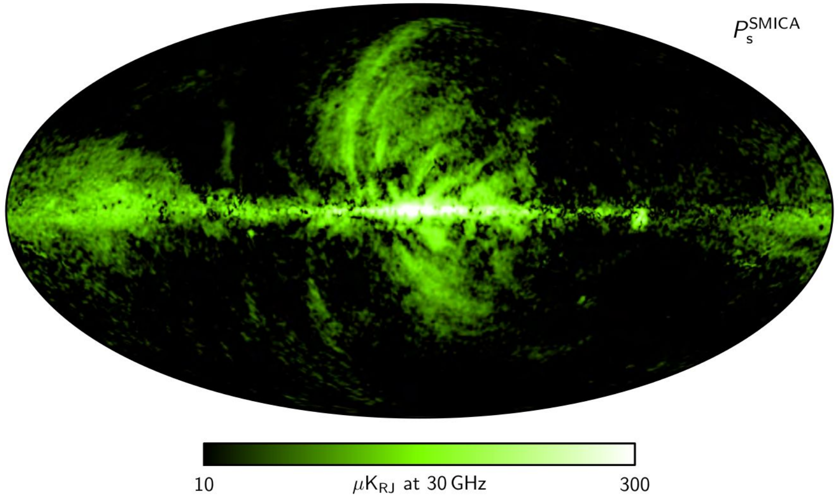
Context: Other sources of contamination are foreground emissions: dust & synchrotron polarization
These are the most important effects when it comes to polarization measurement

⇨ **Component separation method are needed !**

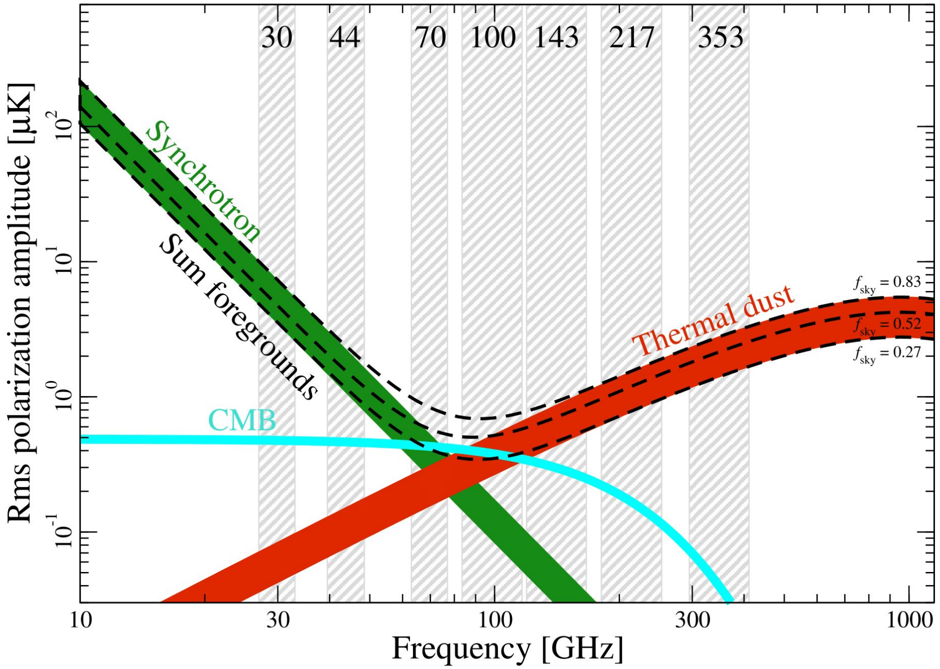
Polarized thermal dust amplitude map @ 353 GHz



Polarized synchrotron amplitude map @ 30 GHz



These methods utilize the fact that we have measurement of the components in multiple different frequencies



Planck collaboration, "Planck 2018 results IV. Diffuse component separation" (2020)

Spectral Matching ICA

$$y_v(\vec{r}) = \sum_{j=1}^{N_c} A_{vj} s_j(\vec{r}) + n_v(\vec{r})$$

- Process & components are Gaussian
- No spatial variation on A
- Components are uncorrelated with each other*

Modeled spectrum:

$$R_Y(\ell) = AR_s(\ell)A^T + R_N(\ell)$$

$\langle YY^T \rangle$ $\quad diag(C_j(\ell))$ $\quad diag(N_v(\ell))$

Observed spectrum:

$$\hat{R}_Y(\ell) = \frac{1}{2\ell+1} \sum_{\ell} \hat{Y}(\ell)\hat{Y}(\ell)^\dagger$$

Maximum likelihood leads to minimizing:

$$\phi(\theta) = \sum_{q=1}^{Q} n_q D(\hat{R}_Y(q); R_Y(q,\theta))$$

D is the Kullback-Liebler Divergence:

$$D(R_1, R_2) = \mathrm{trace}(R_1 R_2^{-1}) - \log\left|R_1 R_2^{-1}\right| - m$$

Parameters:

$$\theta = (A, C_j(q), N_v(q))$$

SMICA in the context of component separation

Component separation can be roughly divided into 2 classes:

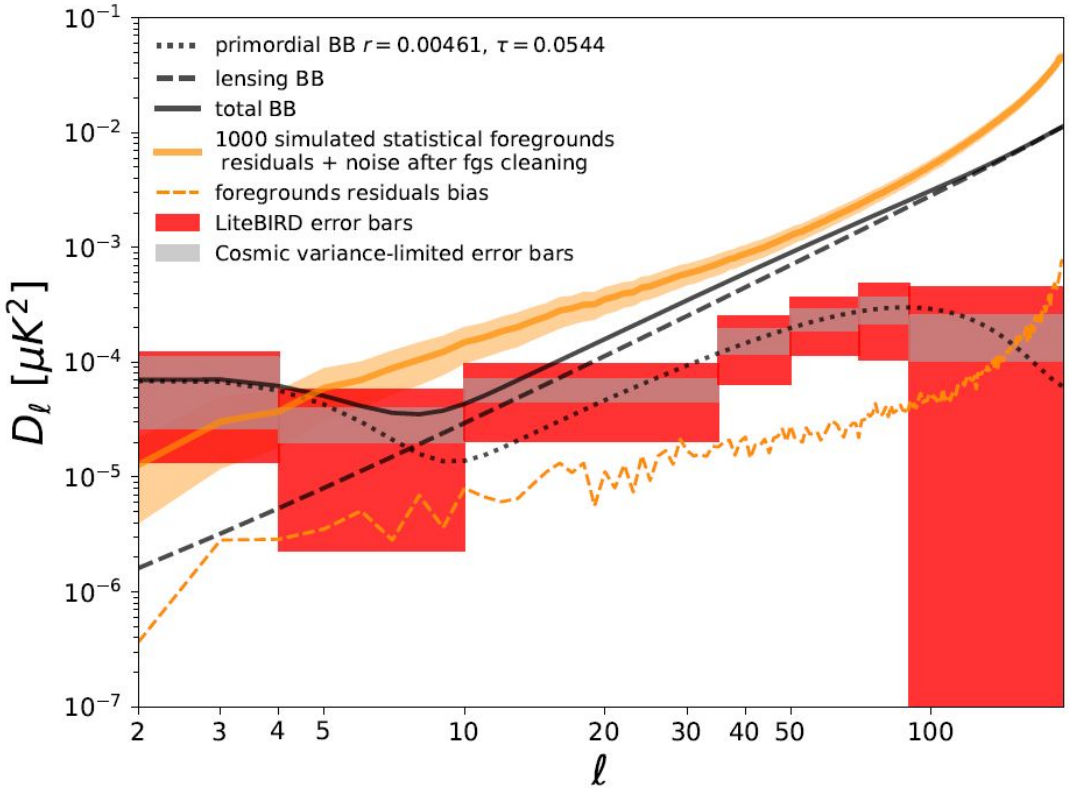
- **Parametric**: makes strong assumption on mixing matrix, doesn't make assumption on the component - FGBuster
- **Blind**: use statistical techniques, no assumption on foregrounds emission - ILC

SMICA lies between these two classes of method

SMICA was developed at APC
Reference method for temperature in Planck

SMICA should be able to **achieve the requirement for component separation** in LiteBIRD

LiteBIRD expected component separation result for B-mode. Method in use was FGBuster – a parametric method



LiteBIRD collaboration, "Probing cosmic inflation with the LiteBIRD cosmic microwave background polarization survey", PTEP 2023

Simulation setup

The simulations (500 random realizations of CMB + noise) are homemade, based on LiteBIRD baseline configuration with 15 frequency bands from 40 - 402 GHz (*LiteBIRD collaboration, PTEP 2023*)

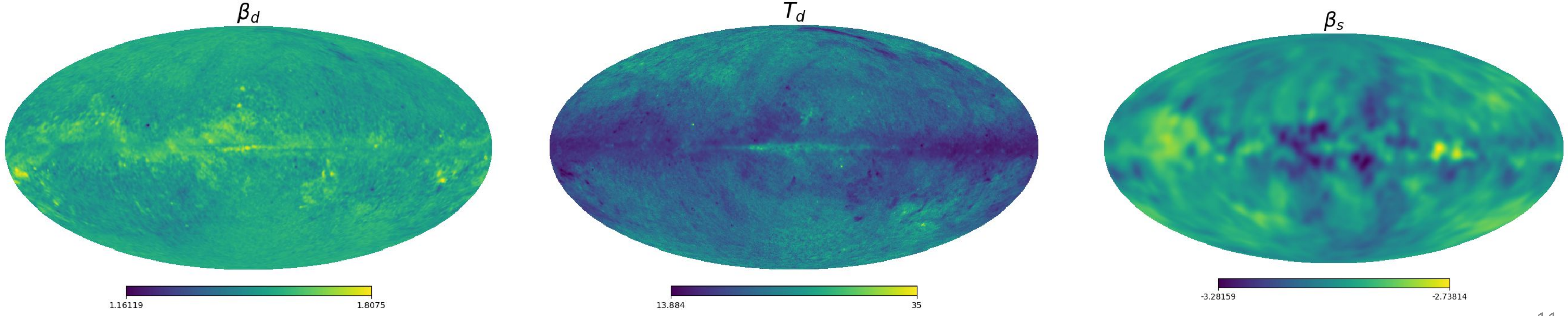
We implement a 2-component (dust + synchrotron) foreground, created from the library PySM

$$y_v(\vec{r}) = \sum_{j=1}^{N_c} A_{vj} s_j(\vec{r}) + n_v(\vec{r})$$

$$A_{dust,\nu}(\beta_d, T_d) \propto \nu^{\beta_d} B(\nu, T_d)$$
$$A_{sync,\nu}(\beta_s) \propto \nu^{\beta_s}$$

Simplest model: d0s0 - constant SED parameters across the sky (beta_dust = 1.54, T_dust = 20K, beta_sync = -3)

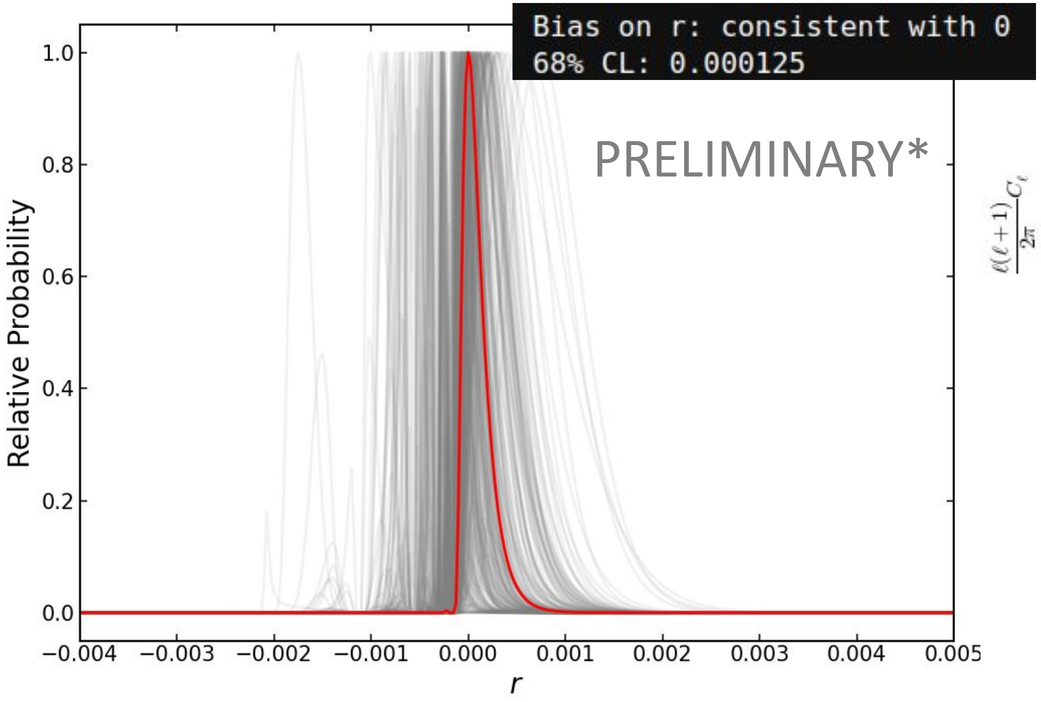
A more realistic model: d1s1 - SED parameters varies smoothly across the sky

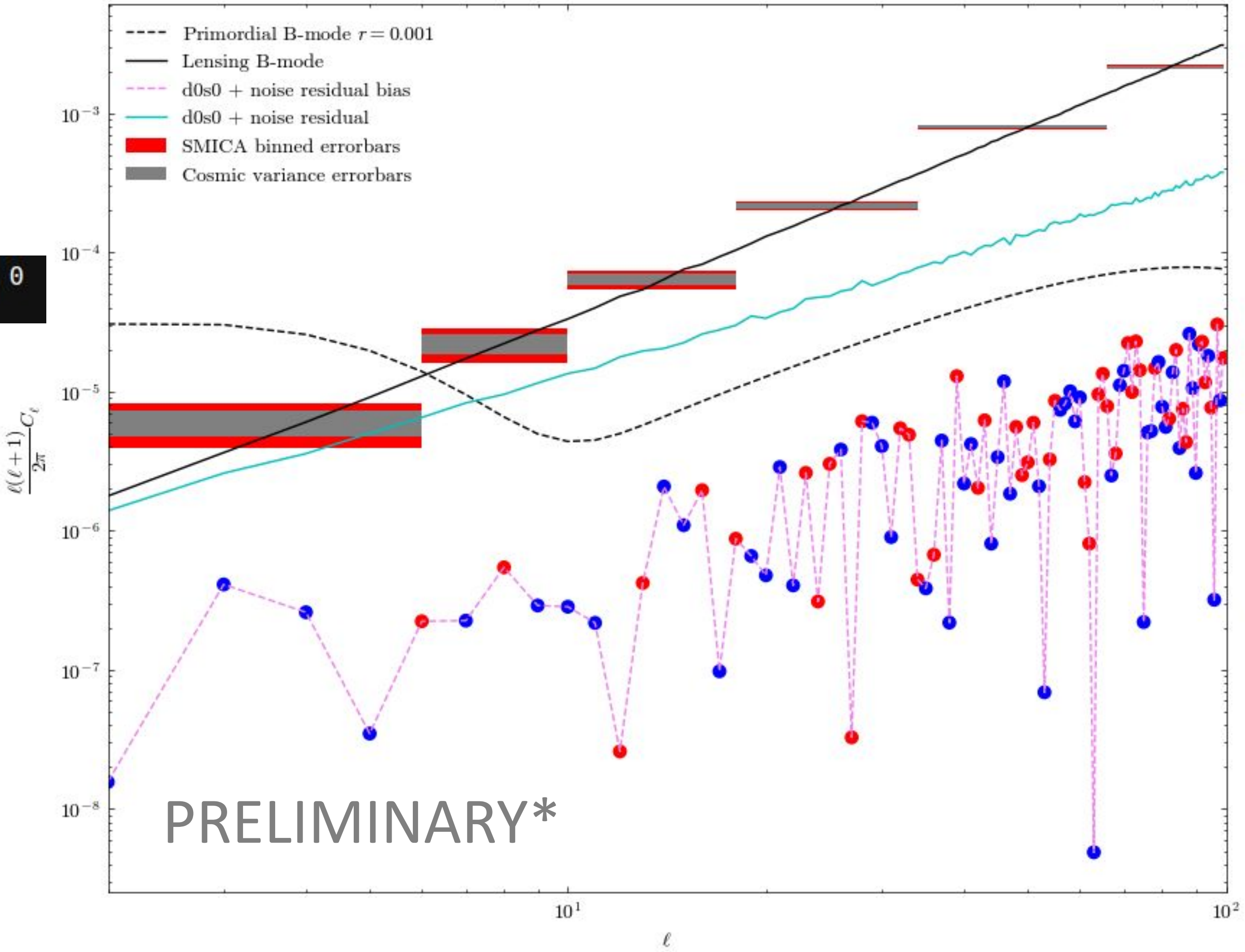


Simple foregrounds d0s0 case

Assumptions: fixed CMB emission, white noise

Likelihood (LoLLiPoP-like) on r for this residual





as the instrument's configuration is under rescope

2. Strategies for semi-blind component separation

SMICA for complex foregrounds

SMICA doesn't have a framework to account for spatial variation => Two strategies to counteract this:

+ Fitting for effective component with SMICA
+ Better modelling of the foregrounds: SMICA with Healpix patches (or some other clustering schemes) => Allow for variation of the mixing matrix A

SMICA with independent components

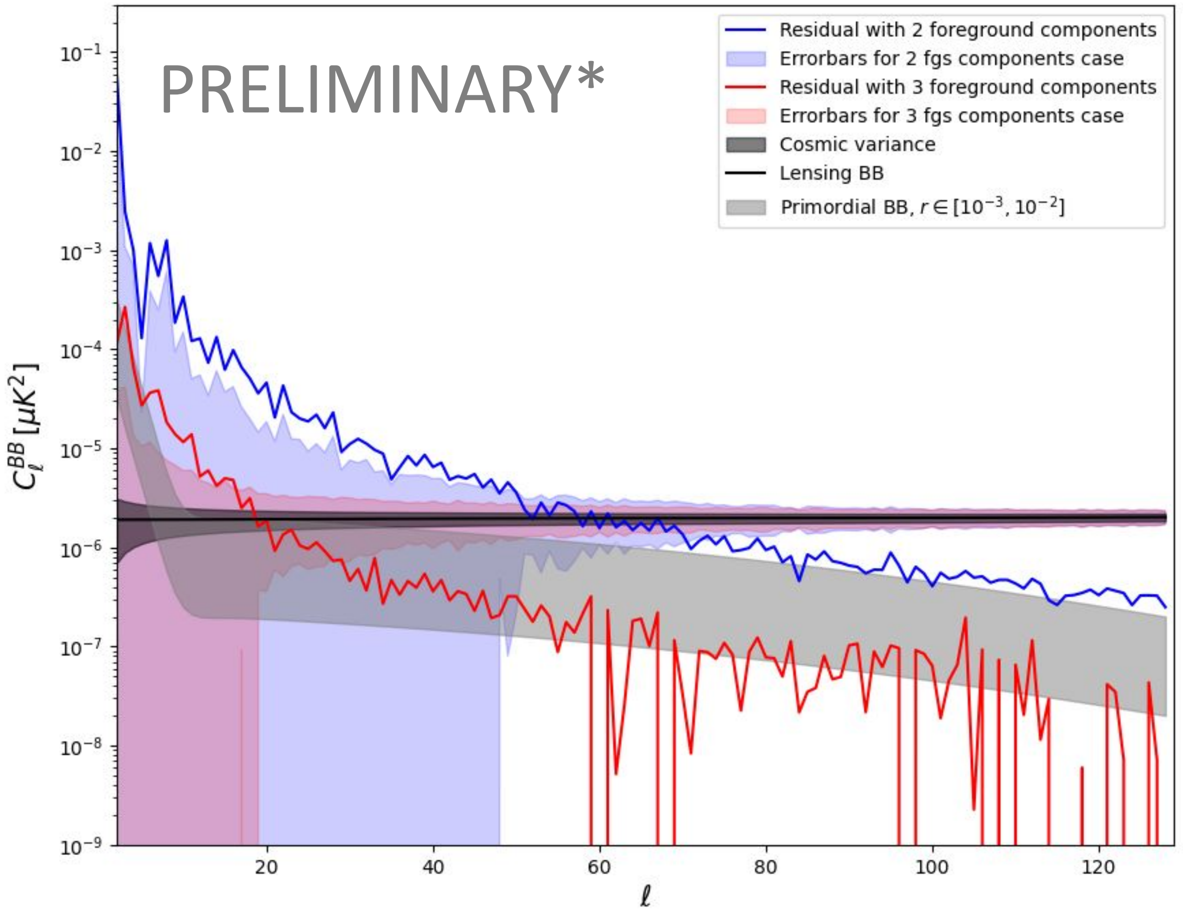
Complex foregrounds can be decomposed into a number of modes (depending on the sensitivity & the complexity of the sky)
=> Adding more components into SMICA can solve the spatial variation

With LiteBIRD sensitivity + d1s1 foregrounds: Across the concerned multipole range (2-128), fitting for 3 foregrounds components should absorb somewhat the bias coming from spatial variation

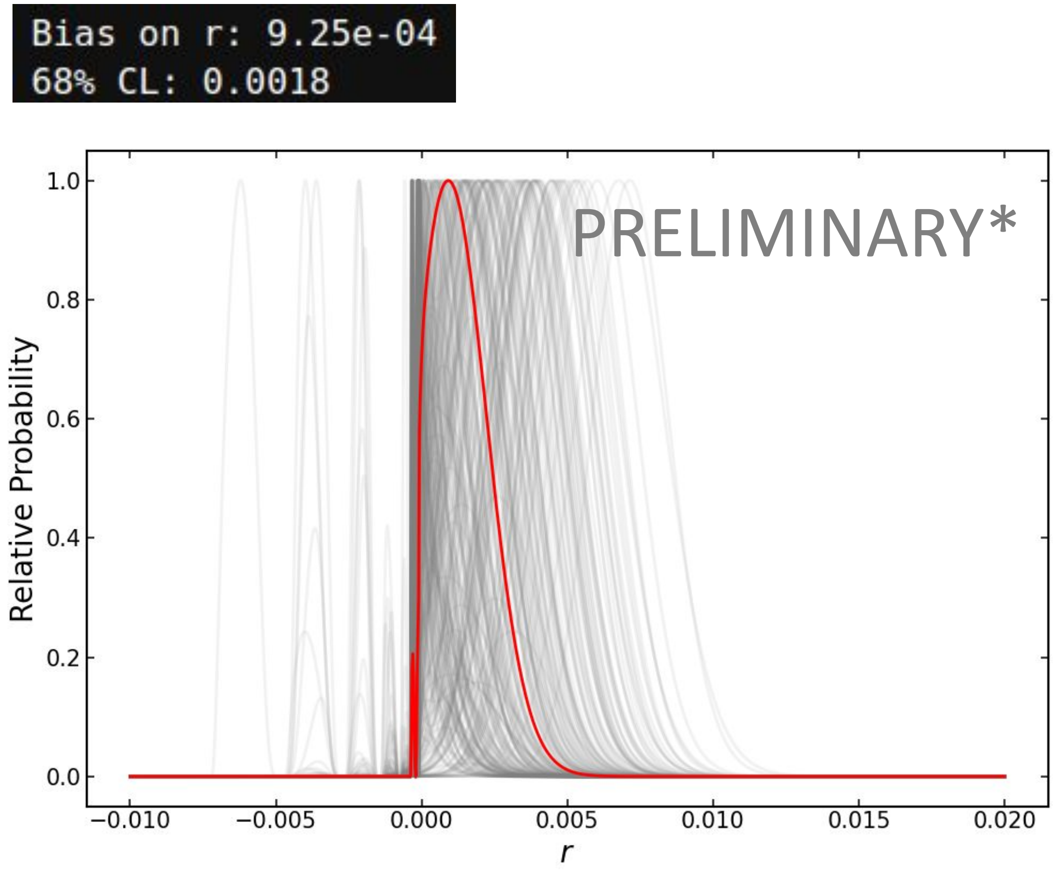
A.Carones. "Optimization of foreground moment deprojection for semi-blind CMB polarization reconstruction" https://arxiv.org/pdf/2402.17579

SMICA with independent components

Applying to a complex sky example (d1s1), featuring some spatial variation of the mixing matrix A on the sky (deviating away from SMICA assumption)



SMICA fitted spectrum (r = 0 in simulation), large bias at low ell is reduced when an additional component is introduced



Likelihood on r, derived from the SMICA fitted spectrum for 3 foreground components.

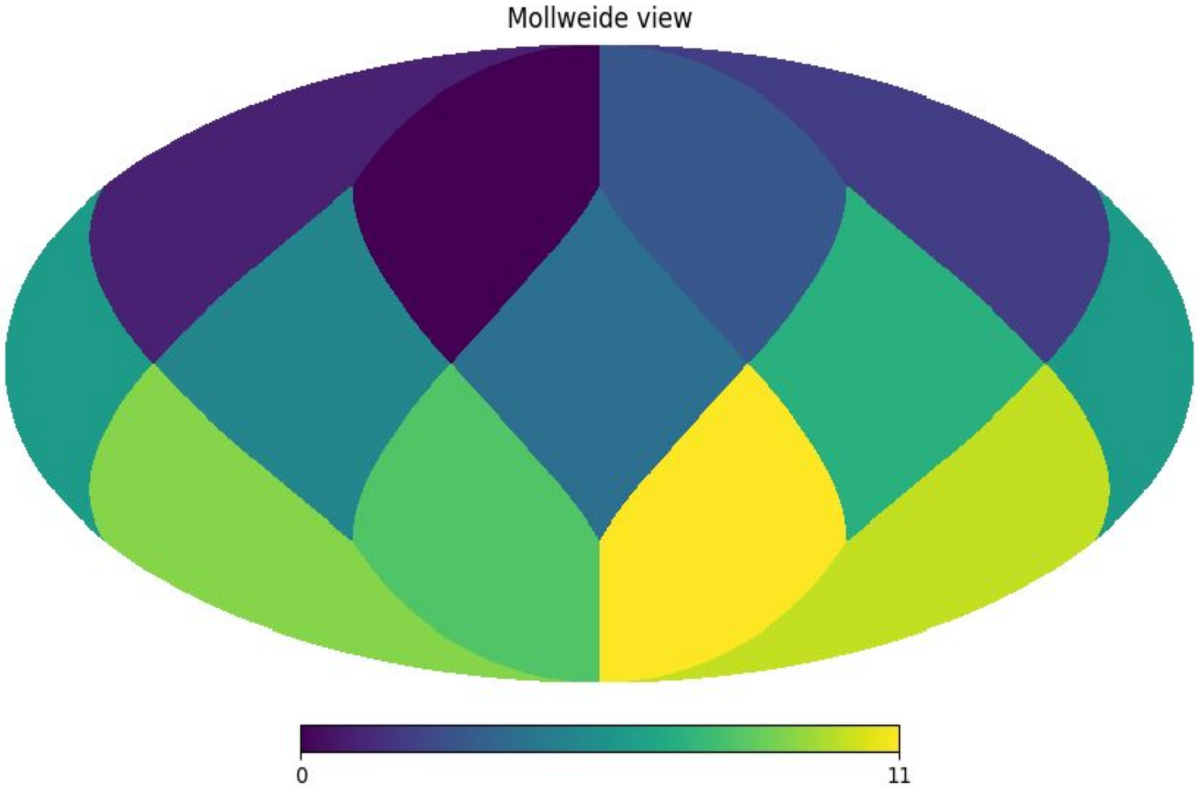
**as the instrument's configuration is under rescope*

SMICA with clusters

Instead of assuming constant SED across the entire sky, we just force it to be constant within certain patches (e.g. Healpix pixel)
=> Run SMICA locally on each patches

Using the SMICA-fitted parameters: mixing matrix, component spectra, noise spectra, we can reconstruct the CMB signal on the using a filtering operation

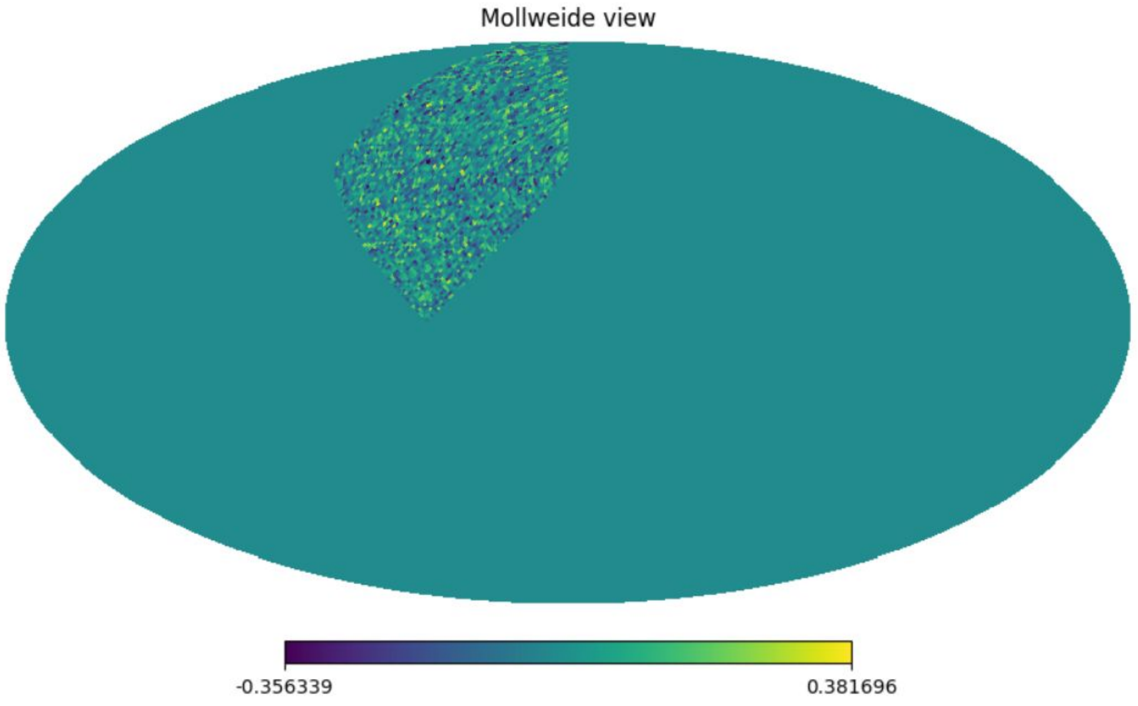
$$s_{WF} = (A^T R_N^{-1} A + R_s^{-1})^{-1} A^T R_N^{-1} d$$



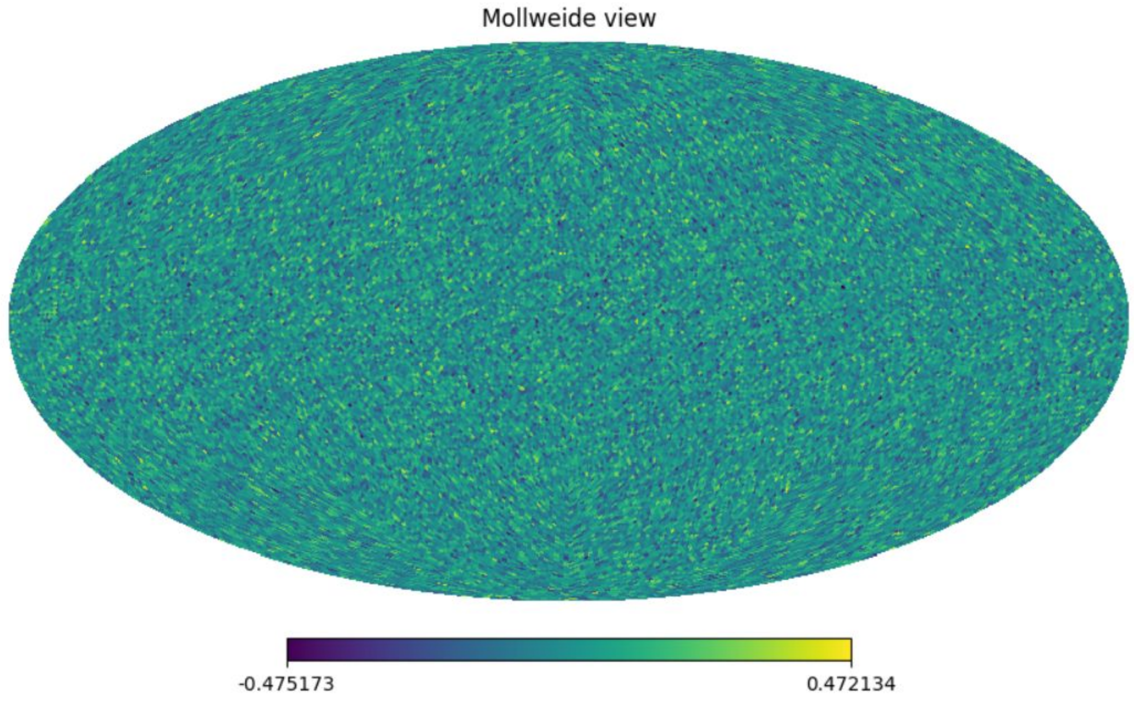
SMICA with clusters

A proof of concept with simple d0s0 foreground: apply the procedure for each patch & combine to get the final map

Cleaned CMB map on an example patch

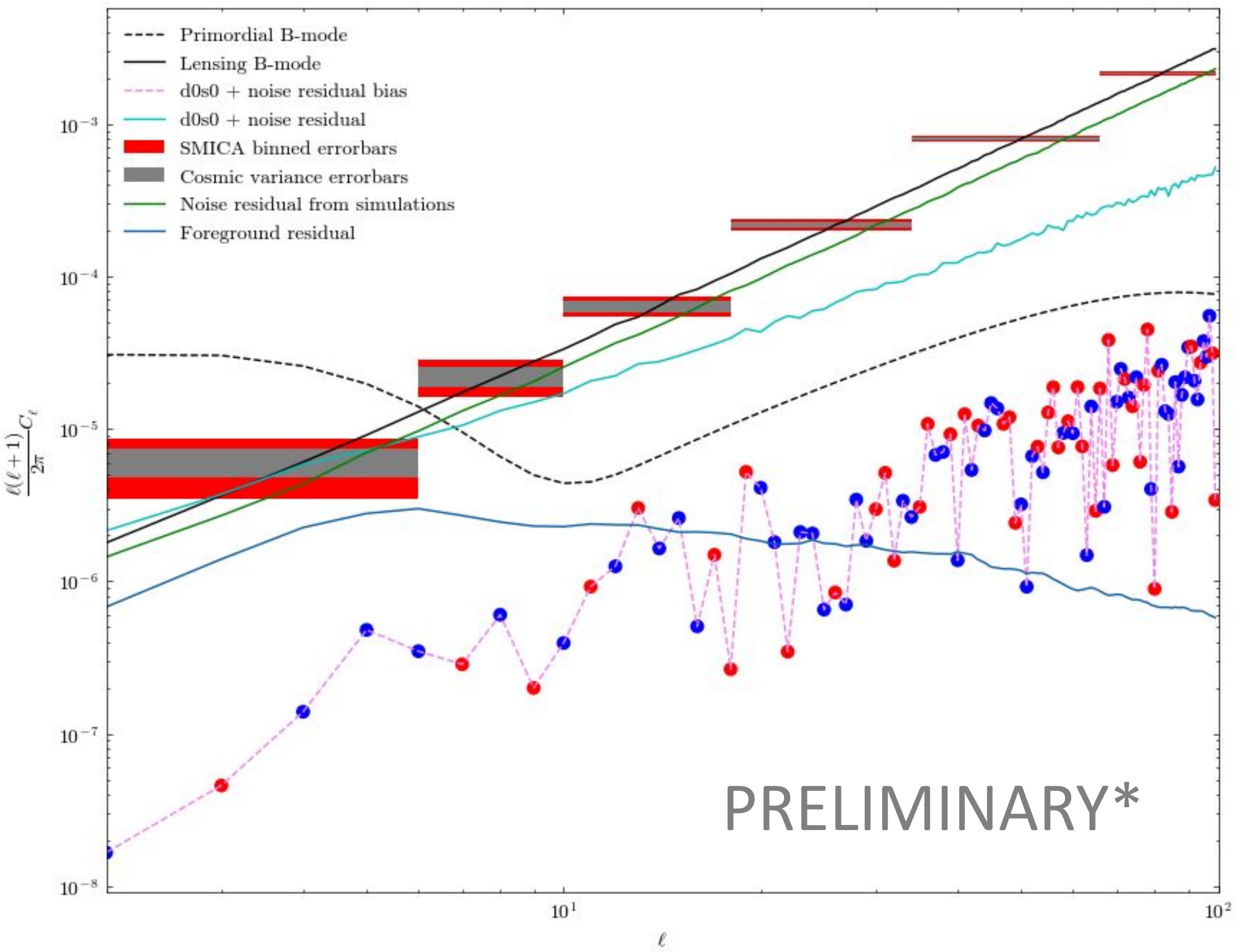


Combined cleaned CMB map

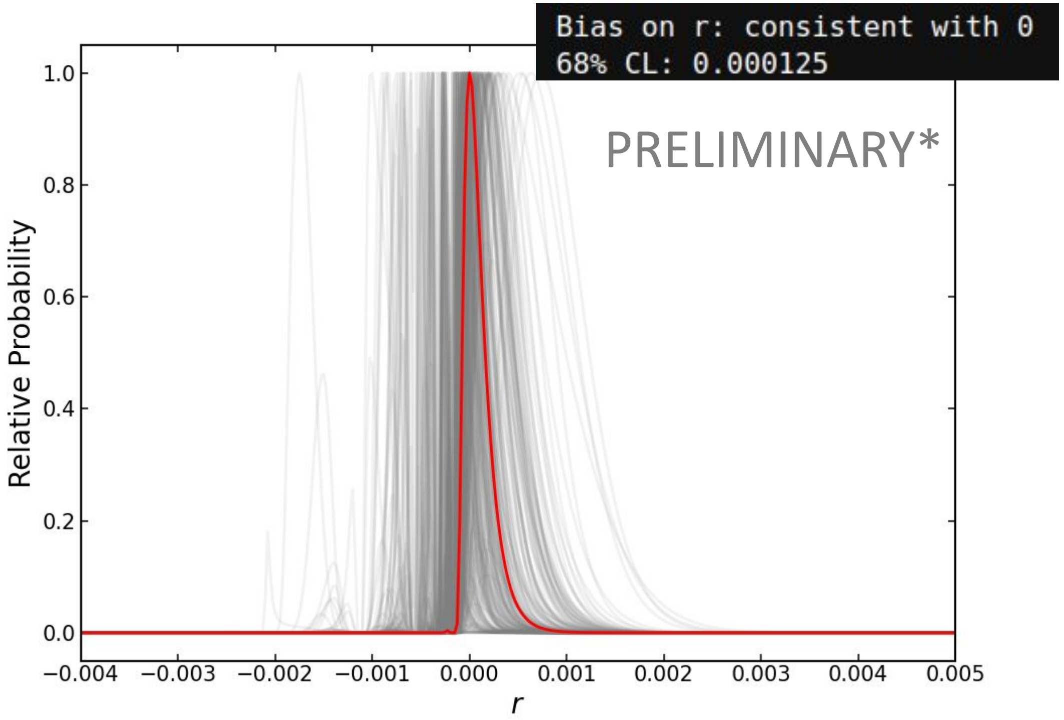


SMICA with clusters

Evaluating the power spectrum of the cleaned map



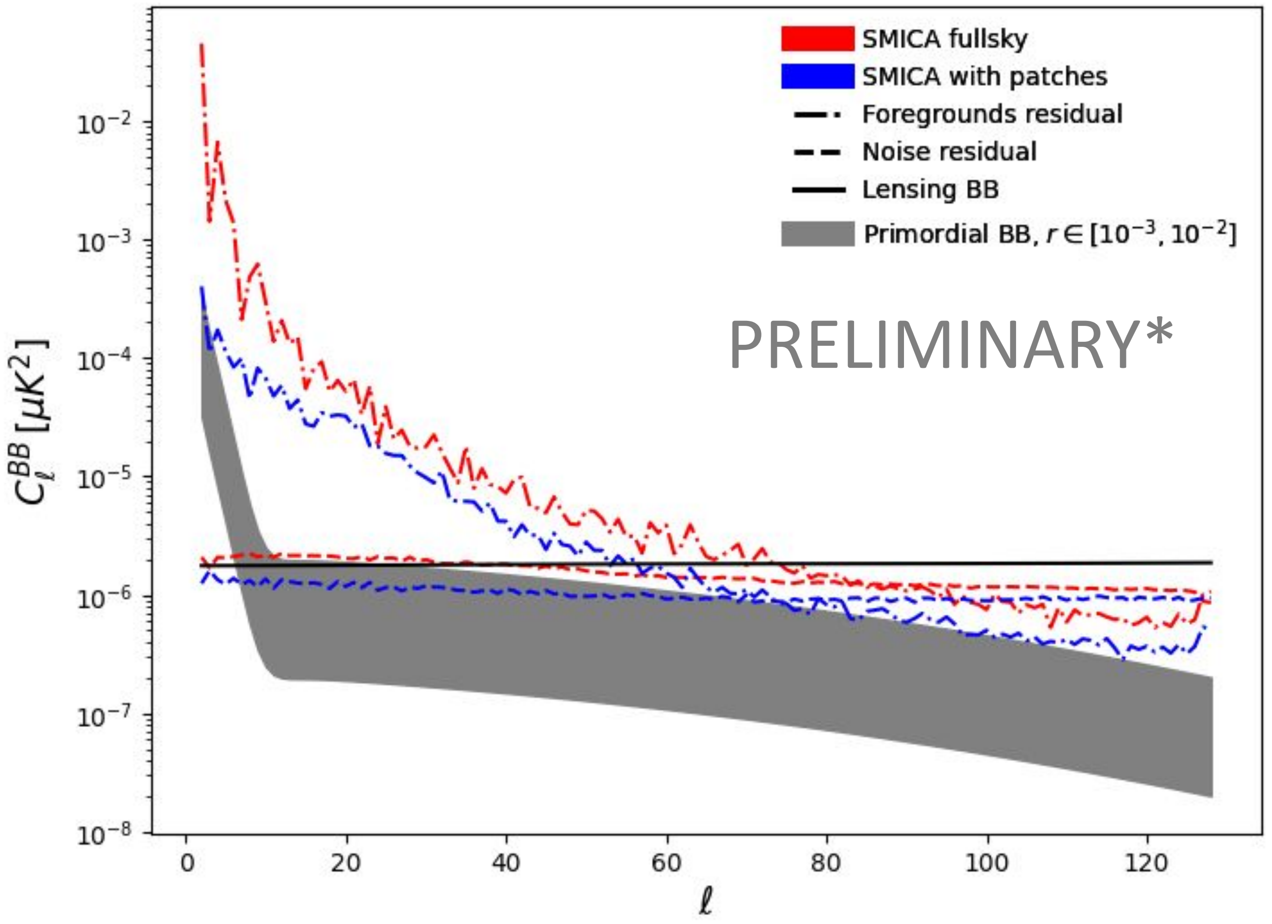
Similar constraints to the full sky d0s0: in this case, the operation does not have an impact on the recovery of r



*as the instrument's configuration is under rescope

SMICA with clusters

<u>Work in progress:</u> Push this pipeline to d1s1 simulations, seeing some improvements compared to the full sky model
We're seeing a reduction in both foregrounds and noise residual



**as the instrument's configuration is under rescope*

3. Summary

Summary

- **SMICA with effective foreground components:**
  - Testing with varying number of components with respect to ell
- **SMICA with clustering schemes:** optimization to be done
  - Pushing to complex foregrounds
  - Different, more optimized clustering scheme (K-means, realistic clusters built from data…)

Achieve requirements for component separation and contribute to the data analysis pipeline of LiteBIRD

- **Other possible developments:**
  - Sampling the SMICA likelihood (similar to MICS)
  - Joint-EB fit with SMICA: To be developed theoretically first
  - Non-gaussian SMICA

THANKS FOR LISTENING