Maximum-likelihood semi-blind component separation for measuring CMB B-mode polarization with a satellite mission

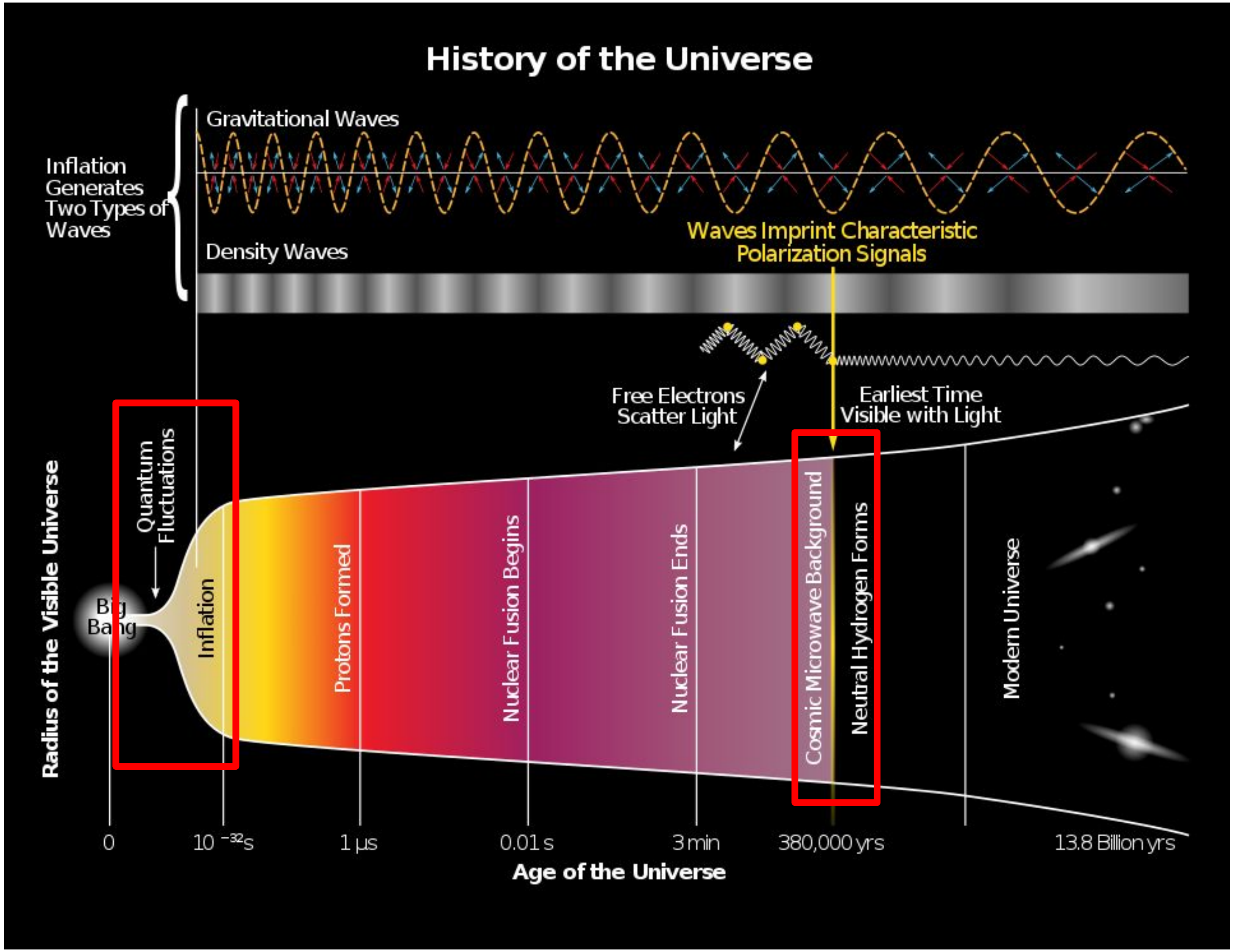
By: Tran Hoang Viet
Supervisor: Guillaume Patanchon - APC, Université Paris Cité
with the help of Michele Citran, Magdy Morshed, Clement Leloup and Benjamin Beringue

Challenges at the CMB Frontier
2nd July 2026

Contents

1. 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

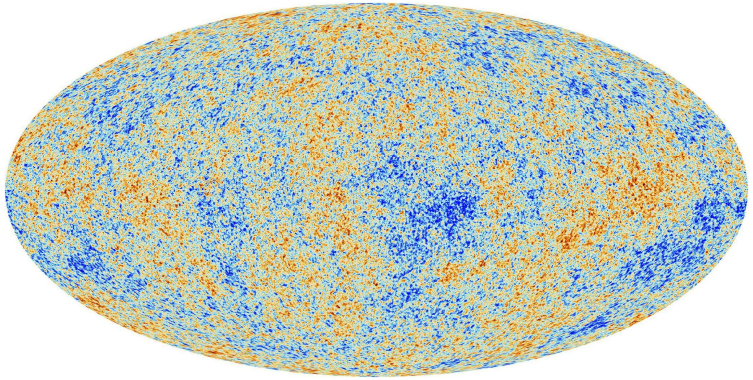
Inflation must have been driven by some process or mechanism not yet seen => **Search for new physics**

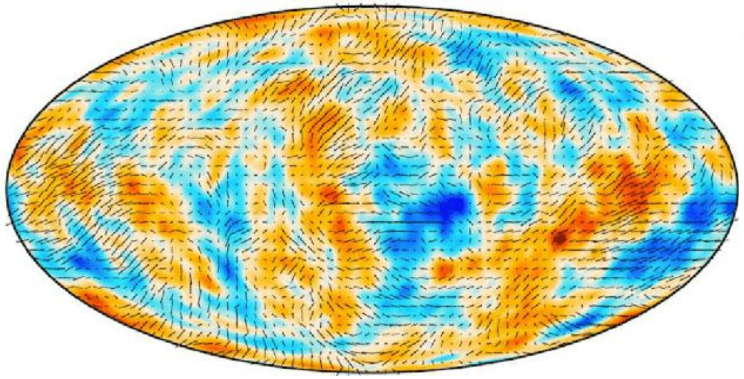
Perturbations evolve till some time after inflation, neutral hydrogen forms, emission of photons from the primordial plasma => **CMB emission**

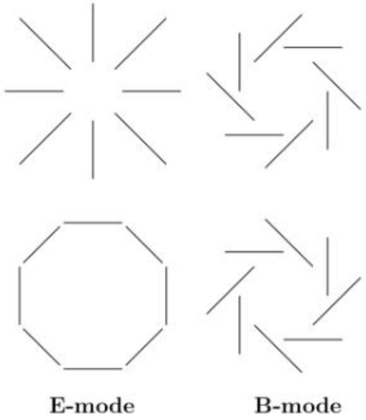
The CMB pattern that we can see today was **imprinted** with the **characteristics of primordial perturbation**

The Cosmic Microwave Background

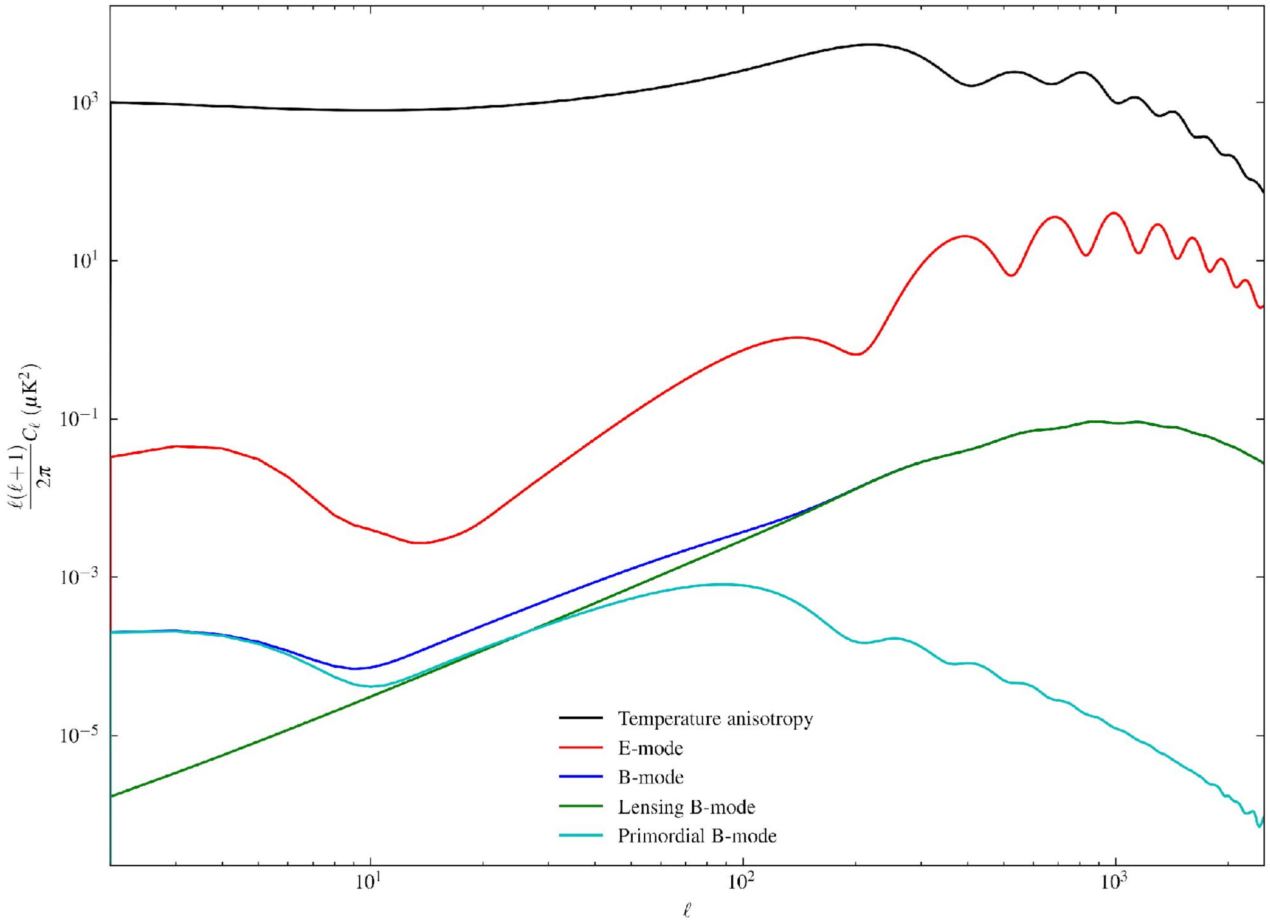
Cosmic Microwave Background: first light of the universe – photons from primordial universe







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

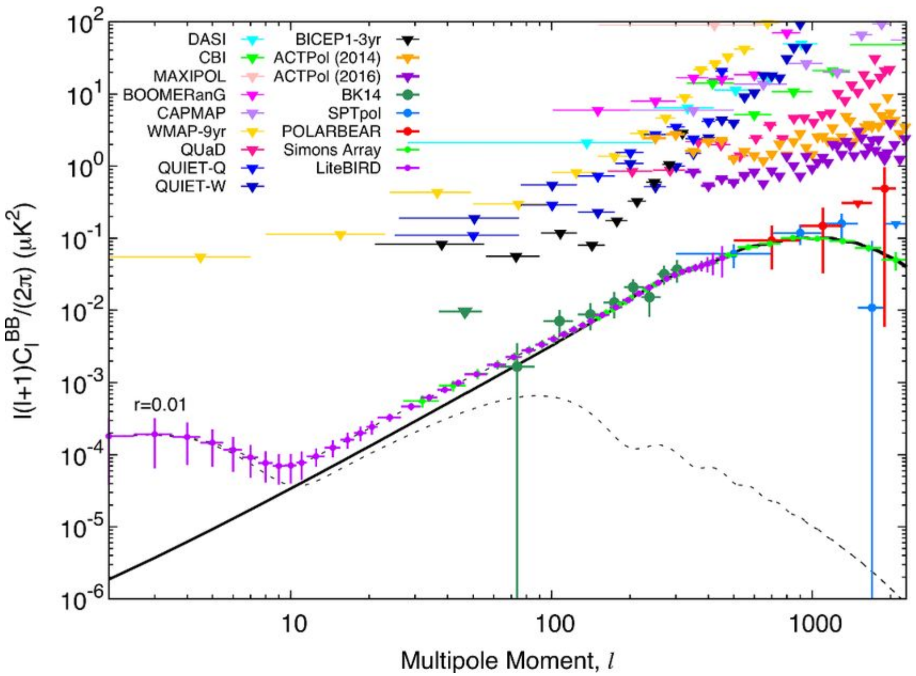
Cosmology and the LiteBIRD mission

Perturbation can be decomposed into: **Scalar** (Density perturbation), **Vector** (Decayed) & **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:

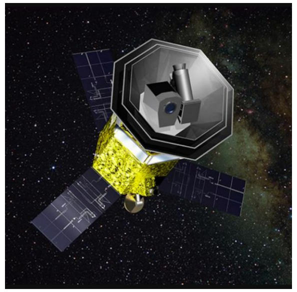
$$r = \frac{\Delta_h^2}{\Delta_k^2} = 16\epsilon$$

r is fully diluted in the intensity & E-mode due to the large amplitude of the scalar perturbation.
=> B-mode is our direct and unique tracer of primordial gravitational wave.



Currently, only an **upper limit** r < 0.032 at 95% confidence (*Tristam et al. 2022*) was established

LiteBIRD is a satellite project led by JAXA under collaboration with CNES, which aims to measure the primordial gravitational waves at **unprecedented precision.**



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

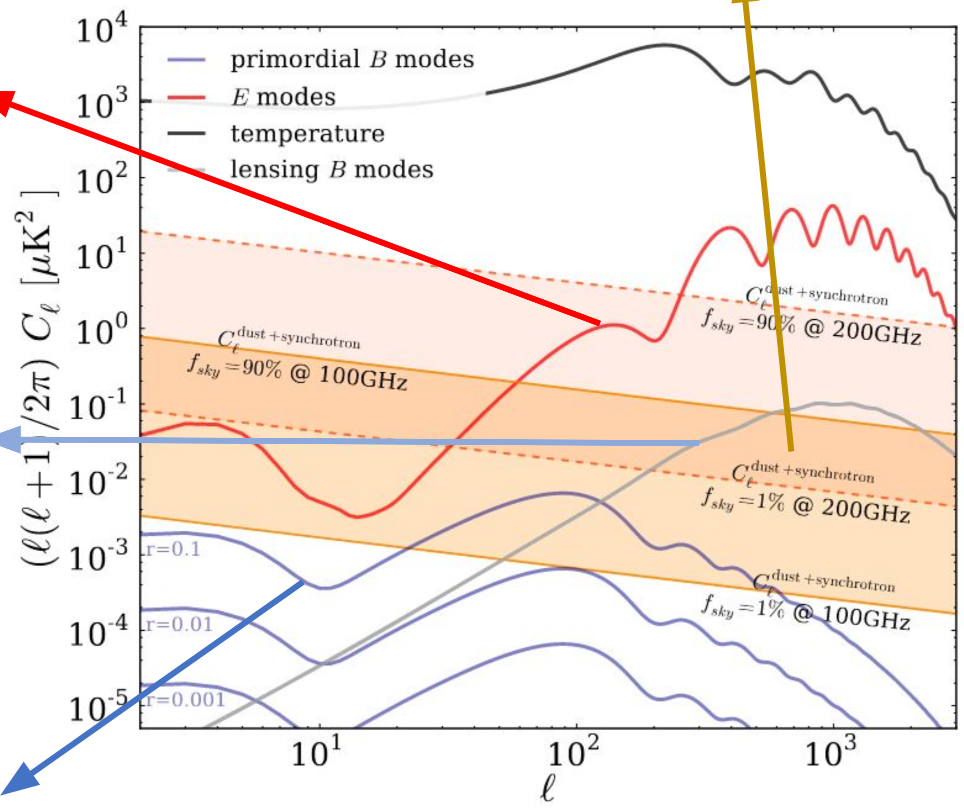
<u>Component separation method are needed !</u>

Foregrounds

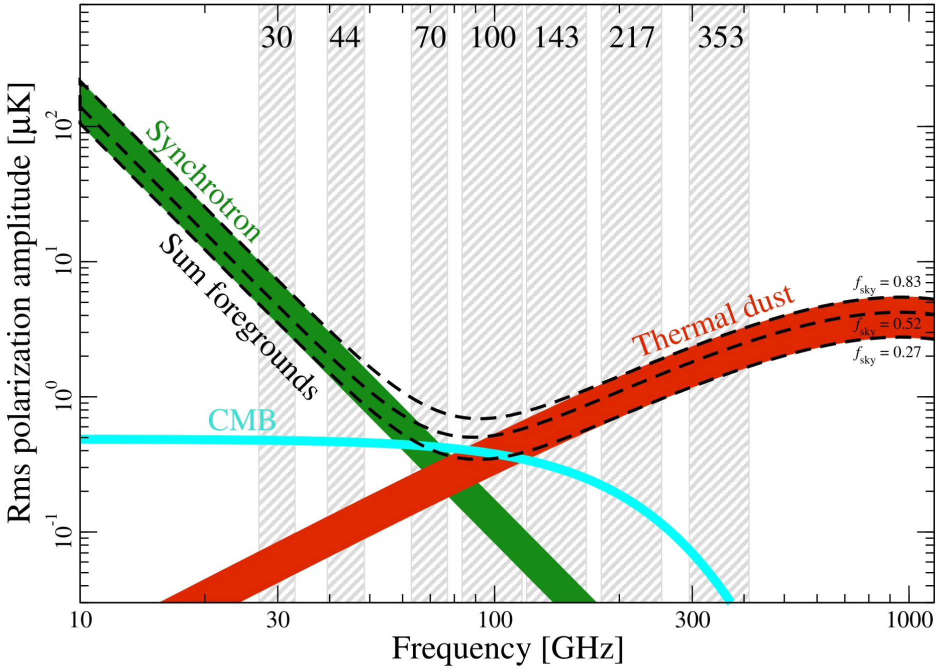
E-mode

Lensing B-mode

Primordial B-mode



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



The landscape of component separation

The data model

$$\mathbf{d} = \mathbf{A}\mathbf{s} + \mathbf{n}$$

Frequency maps

Mixing matrix

Component maps

Noise

Component separation methods must use **assumption(s)** to break the fundamental degeneracy between A and s:

Parametric: explicit model for foreground emissions.

- **FGBuster** *(Stompor et al. 2009)*
- **Moment Expansion** *(Chluba et al. 2017, Vacher et al. 2023)*
- **Commander** *(Eriksen et al. 2006)*

Blind: statistical assumptions (minimum variance, decorrelation)

- **ILC** *(Bennett et al. 2003)*
- **NILC** *(Delabrouille et al. 2009)*
- **SMICA** *(Delabrouille et al. 2002)*
- **MICS** *(Leloup et al. 2023)*

Realistic foregrounds that we observe exhibits **spatial variability** in their emissions: The **central challenge** to component separation for measuring CMB B-mode.
To counteract this extra degree of complexity, a possible direction is to combine component separation methods with **clustering techniques** *(Carones et al. 2023, Rizzieri et al. 2025, Kabalan et al. 2026)*

Spectral Matching ICA

SMICA (*Delabrouille et al. 2002*) is one of the **reference component separation** of the **Planck** experiment (*Planck Collaboration 2016*)

$$\underbrace{\mathbf{d}}_{\text{Frequency maps}} = \underbrace{\mathbf{A}}_{\text{Mixing matrix}}\underbrace{\mathbf{s}}_{\text{Component maps}} + \underbrace{\mathbf{n}}_{\text{Noise}}$$

SMICA modeled harmonic covariance

$$\mathbf{R}_\ell = \mathbf{A}\mathbf{S}_\ell\mathbf{A}^T + \mathbf{N}_\ell$$

Assumptions

- Data and components are Gaussian process
- Mixing matrix A does not vary with position
- Uncorrelated components

SMICA seeks to match the modeled covariance with the empirical one through minimizing the function

$$\phi(\mathbf{A},\mathbf{S},\mathbf{N}) = \sum_\ell \mathrm{D}[\hat{\mathbf{R}}_\ell, \mathbf{R}_\ell(\mathbf{A},\mathbf{S},\mathbf{N})]$$

with D being the Kullback-Leibler divergence

$$\mathrm{D}[\mathbf{R}_1,\mathbf{R}_2] \propto \mathrm{trace}(\mathbf{R}_1\mathbf{R}_2^{-1}) - \log\det(\mathbf{R}_1\mathbf{R}_2^{-1})$$

<u>Problem: Restricted to harmonic domain and no natural inclusion of spatially varying A</u>

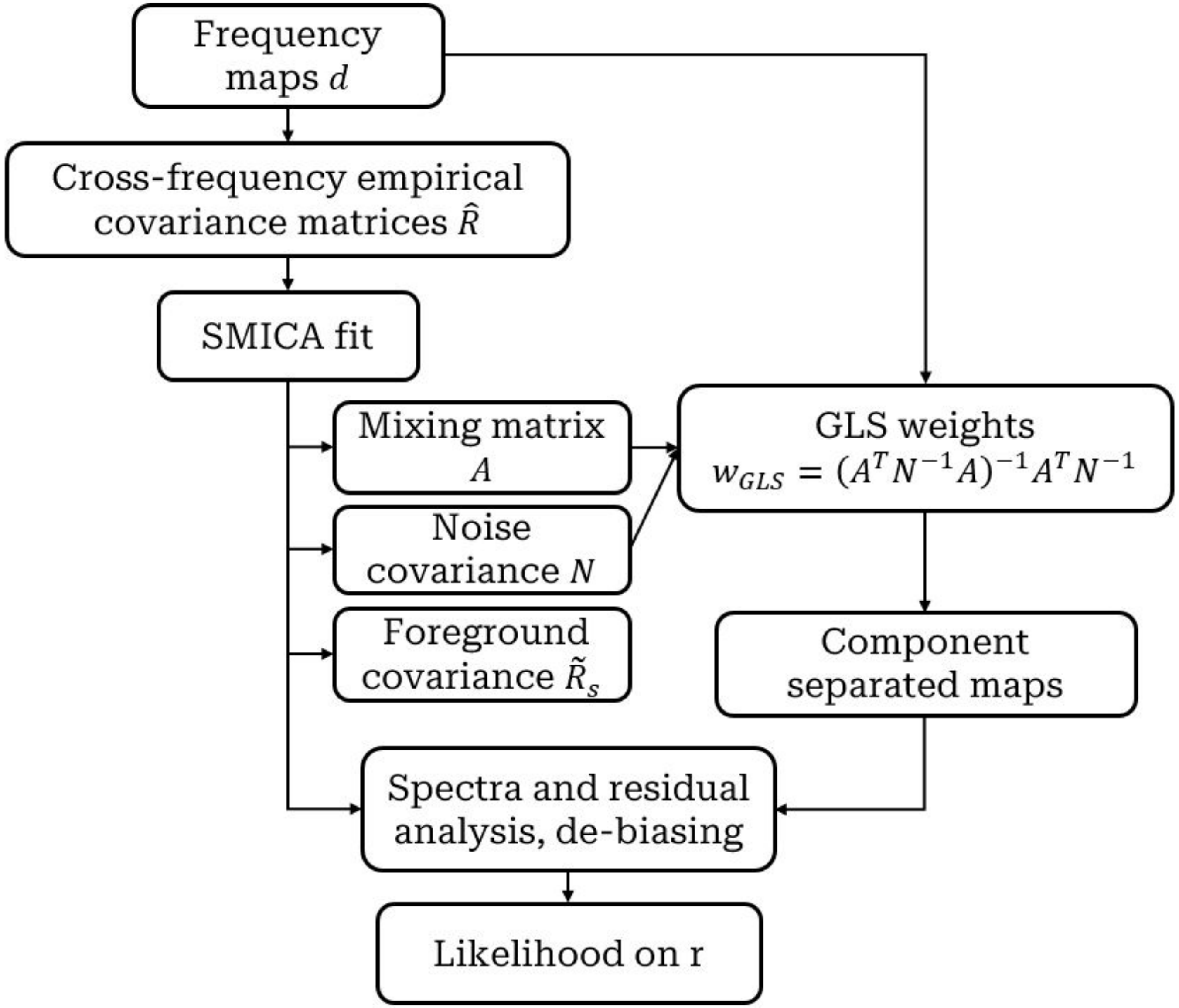
Overall SMICA-based pipeline for B-mode polarization with a satellite mission

We have made attempts to include spatial variation within SMICA by going back and forth between harmonic and pixel domain *(Tran et al. 2026 in prep)*

The pipeline, despite being able to perform comp-sep to a good degree and is used in LiteBIRD comp-sep study, has some notable points that are discussed in the paper:

- The end to end process is not maximum-likelihood motivated
- Inconsistencies between the steps
- The performance relies heavily on external tools such as: pseudo-spectra estimator for partial sky coverage, independent likelihood model to derive constraints…

=> Is there a global, consistent and maximum-likelihood motivated way to optimize SMICA?



2. MISTICA: Map-based Identification of Spectral Templates with Independent Component Analysis

MISTICA

***MISTICA**: **M**ap-based **I**dentification of **S**pectral **T**emplates with **I**ndependent **C**omponent **A**nalysis*

This method is a way to generalize **SMICA** with a natural incorporation of spatial variability and partial-sky coverage by following the **MICS** (*Leloup et al. 2023*) framework and its pixel domain implementation **MICMAC** (*Morshed et al. 2024*).

$$\mathbf{d} = \mathbf{A}\mathbf{s} + \mathbf{n}$$

- Components can be described with Gaussian statistics
- Uncorrelated components

$$-2\ln\mathcal{L}(\mathbf{A}, \mathrm{s}, \mathbf{S}) \propto (\mathbf{d} - \mathbf{A}\mathbf{s})^T\,\mathbf{N}^{-1}\,(\mathbf{d} - \mathbf{A}\mathbf{s}) + \ln|\mathbf{S}| + \mathbf{s}^T\mathbf{S}^{-1}\mathbf{s}$$

Mixing matrix defined in pixel domain for spatial variability

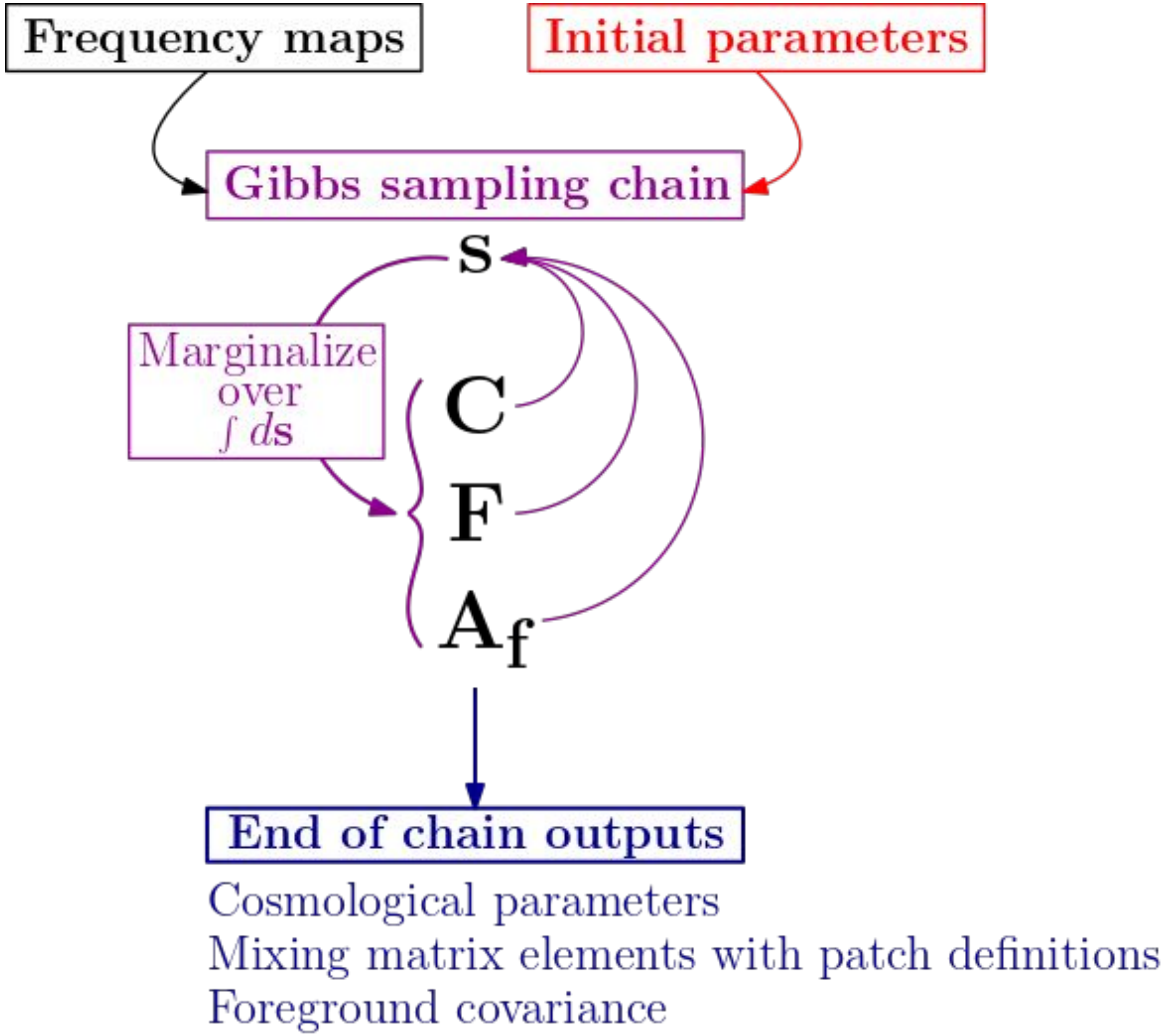
Component maps

Source covariance in harmonic domain featuring the interdependencies of components

$$\mathbf{S} = \begin{pmatrix} \mathbf{C} & 0 \\ 0 & \mathbf{F} \end{pmatrix}$$

Noise covariance defined in pixel domain, possibly including a mask definition for partial/cut sky handling

MISTICA



We go directly from frequency maps to cosmological parameters in a maximum likelihood, consistent framework.

Implemented using the existing structure of the publicly available **MICMAC package,** exploiting **Gibbs sampling** and a **multi-patch approach** (Errard et al. 2018):

- The sampling of the component maps features **in-painting** technique to counteract **partial/cut-sky** without having to use **pseudo-spectra**.
- The mixing matrix elements are defined over an **arbitrary patch distribution** in the **pixel domain** to **capture spatial variability**.

3. Validating the framework

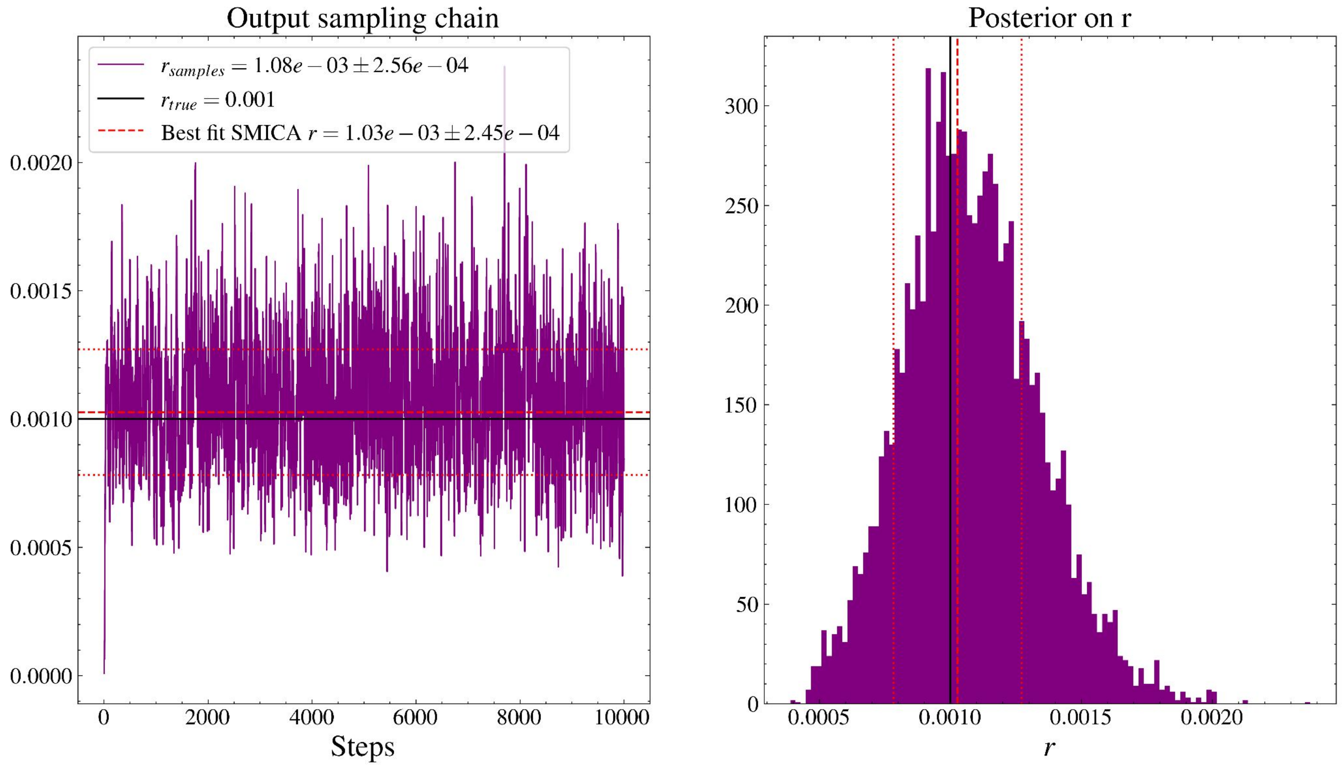
MISTICA validations

We check the performance of our method in the simulated condition of a LiteBIRD-like experiment (*LiteBIRD Collaboration et al. 2022*)

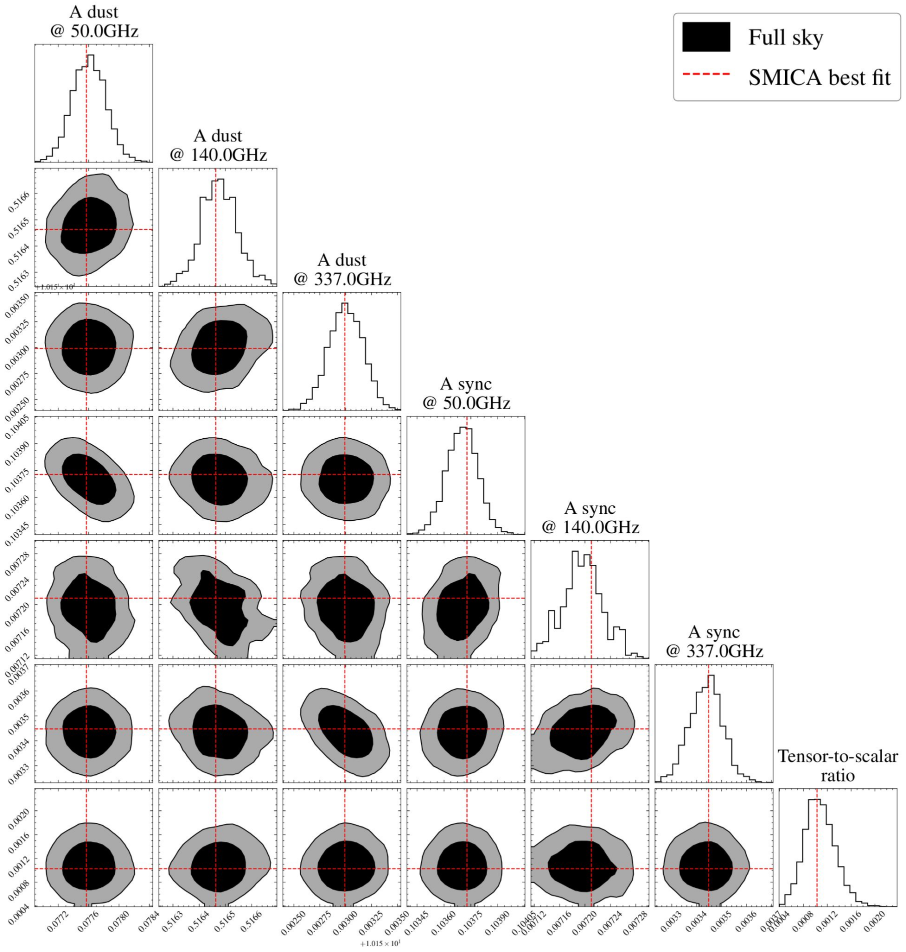
Unbiased recovery on full-sky simple homogeneous foregrounds (d0s0), perfect agreement with standard SMICA

=> Proof that we're working with the same likelihood

Samples of the tensor-to-scalar ratio

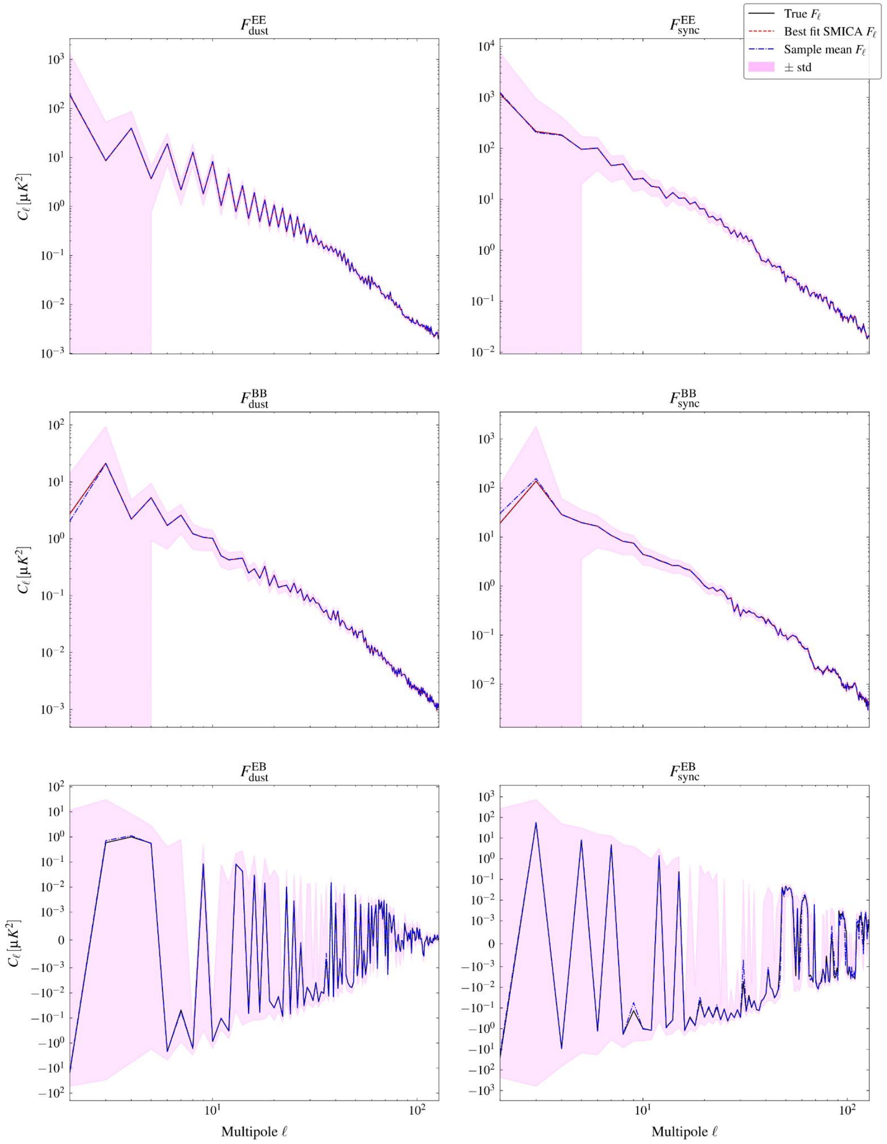


Samples of the mixing matrix



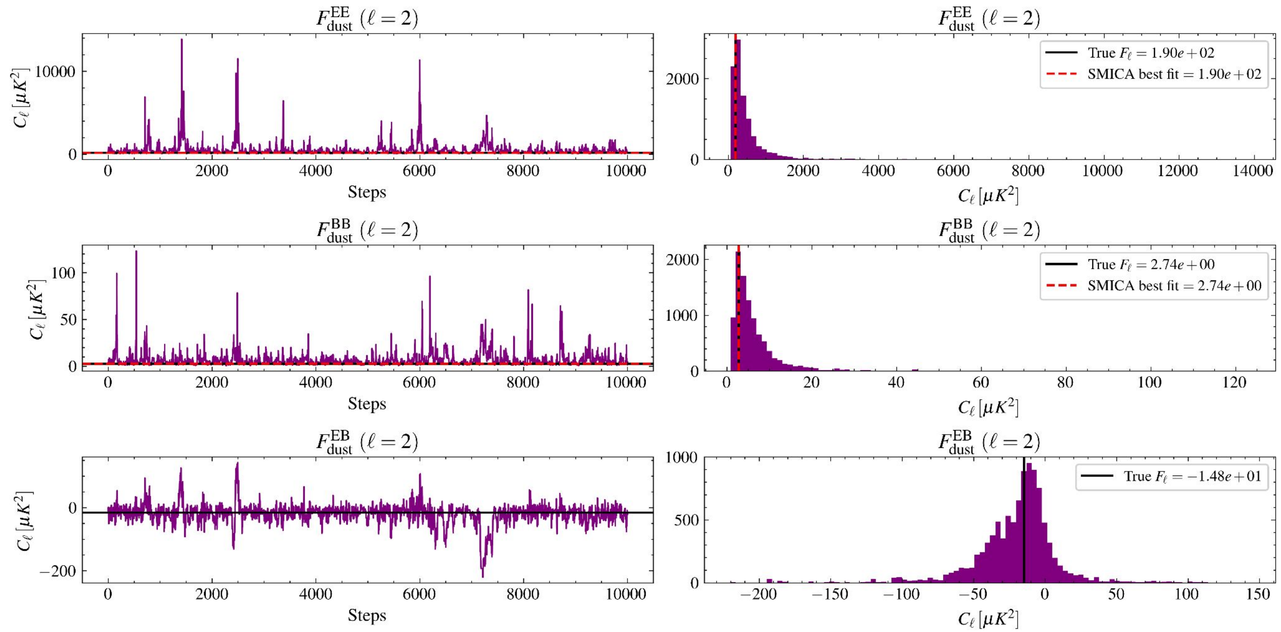
MISTICA validations

Unbiased recovery on d0s0, perfect agreement with standard SMICA
=> Proof that we're working with the same likelihood



Sampling of the foreground covariance is possible !

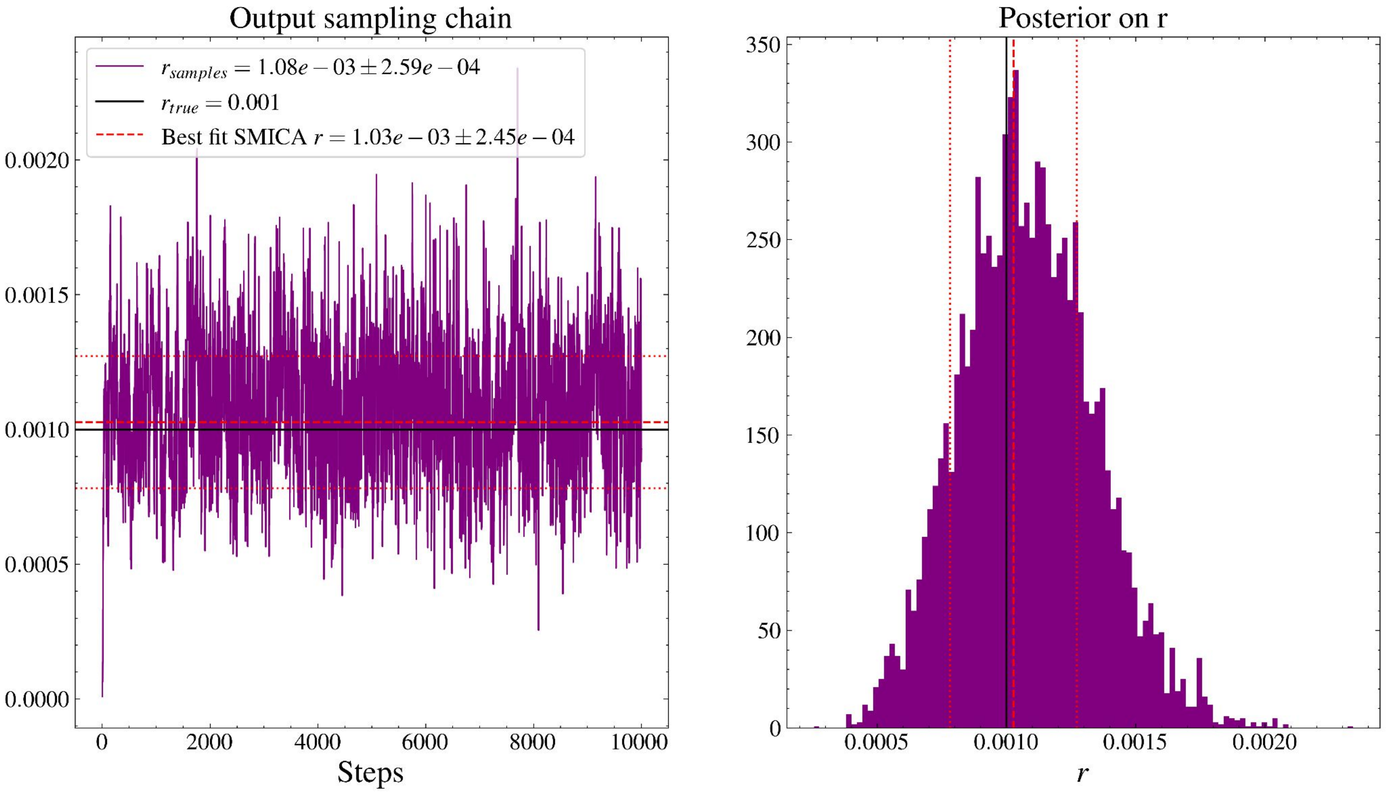
Even the EB part (not included in standard SMICA) is perfectly reconstructed
=> Promising for birefringence/polarization angle study !



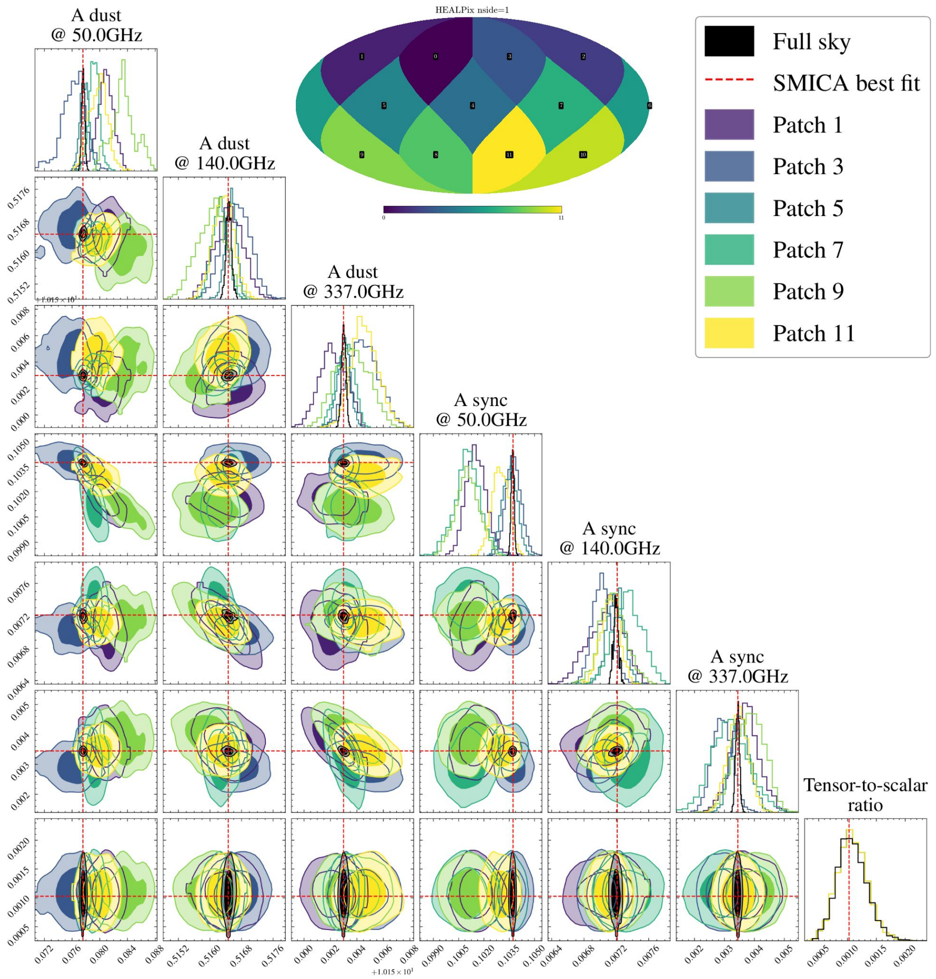
MISTICA validations

Multi-patch capability: testing on homogeneous foregrounds => No bias

Samples of the tensor-to-scalar ratio: Does not deviate from the case with no patches!



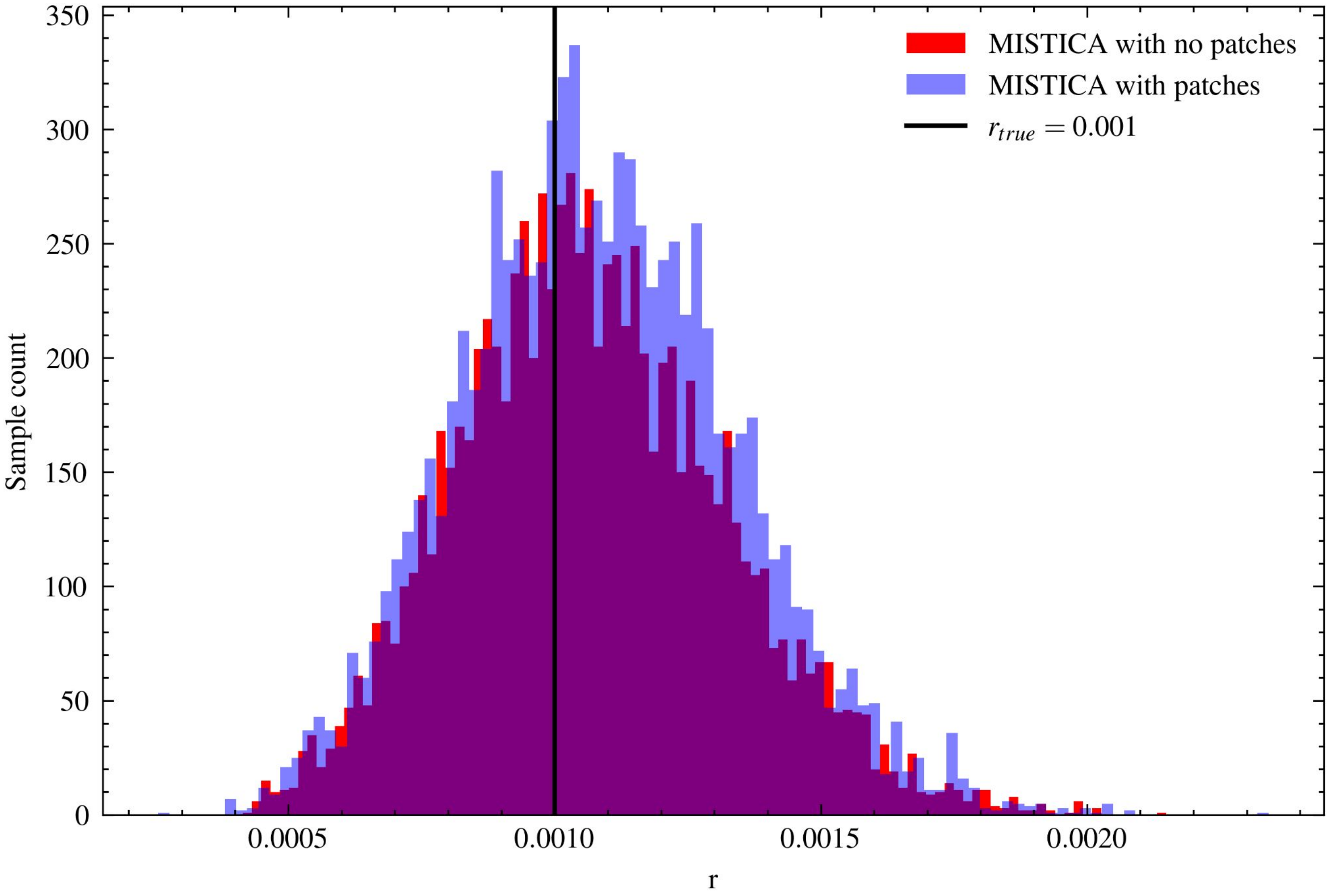
Samples of the mixing matrix with elements defined on patches!



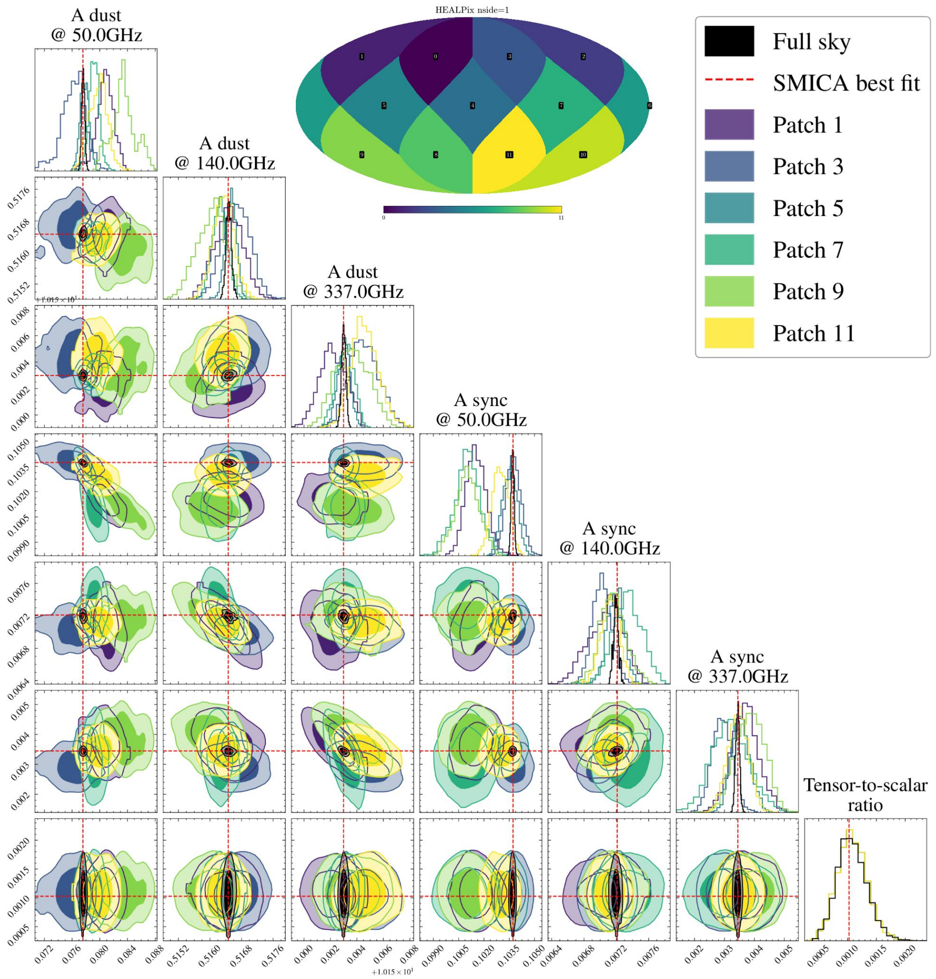
MISTICA validations

Multi-patch capability: testing on homogeneous foregrounds => No bias

Samples of the tensor-to-scalar ratio: Does not deviate from the case with no patches!



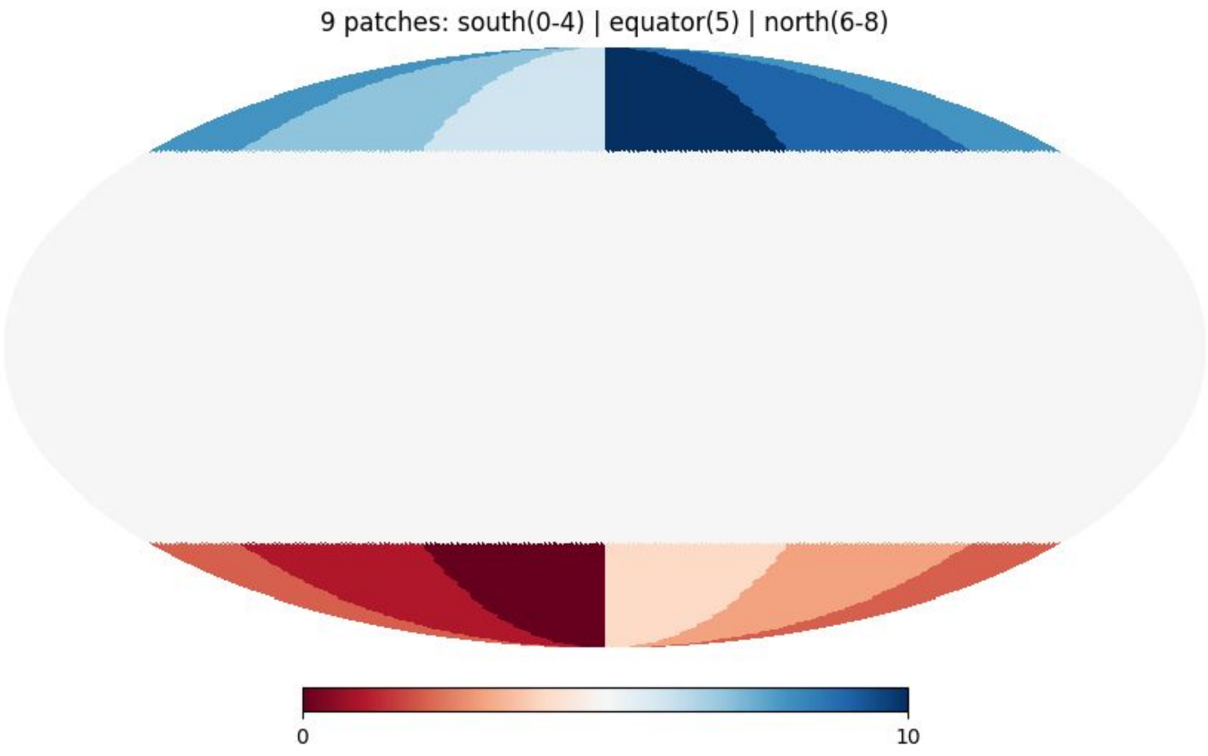
Samples of the mixing matrix with elements defined on patches!



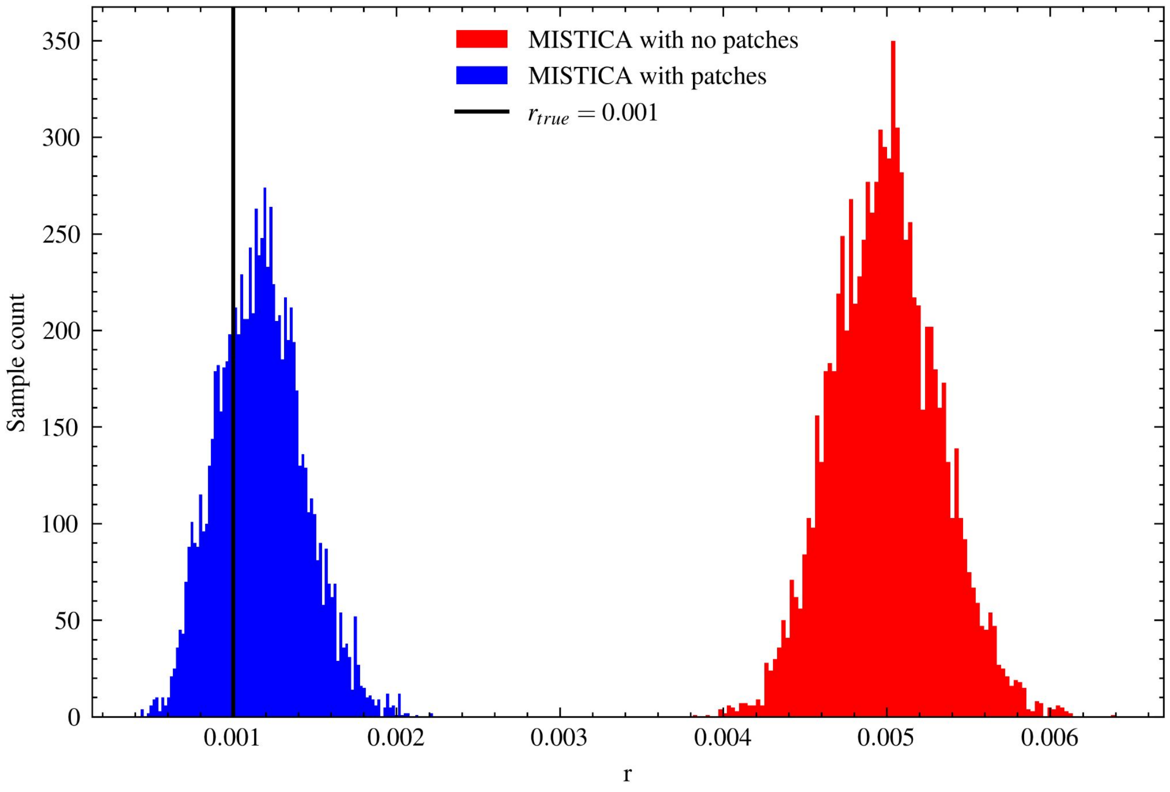
MISTICA validations

Multi-patch capability: custom complex foregrounds

The spatial variability of the mixing matrix is artificially created with the following template



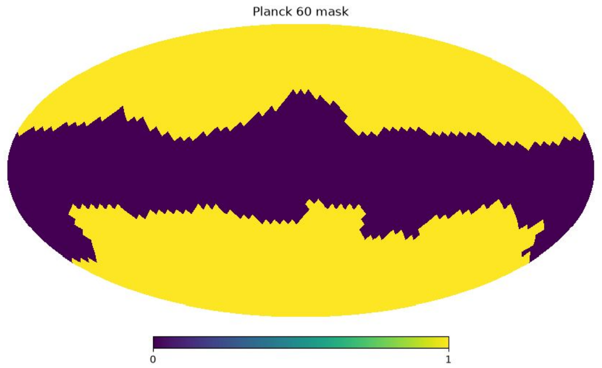
Create bias on r in the standard version that is corrected once multi-patch was included in the framework



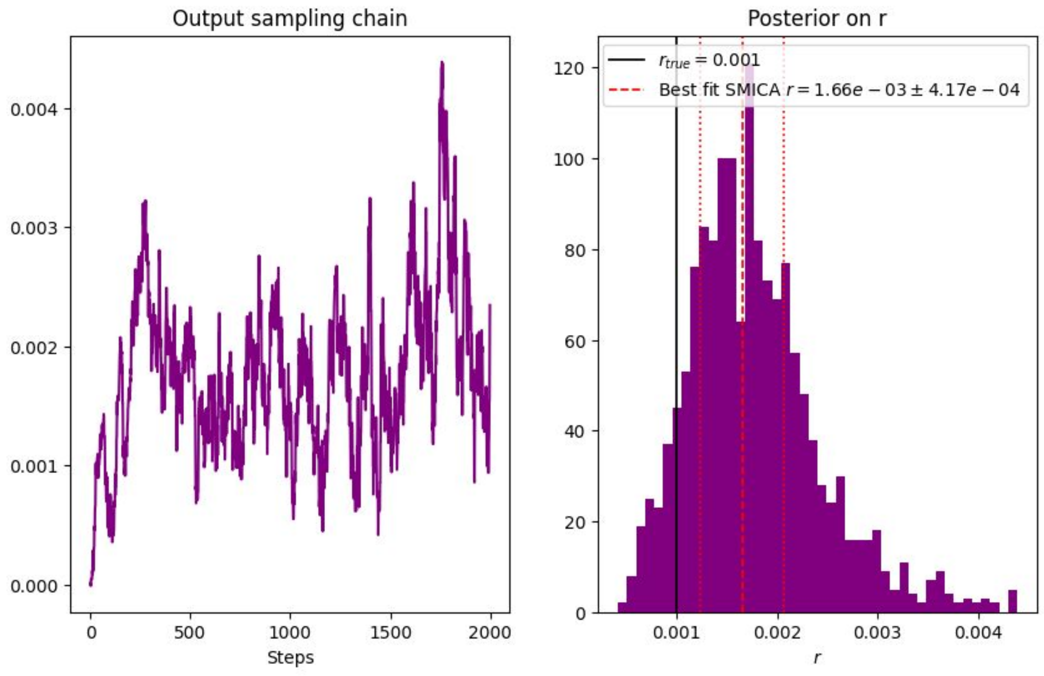
The case is simple but serve well the purpose of validation, applications on actual complex foregrounds are saved for future work.

MISTICA validations

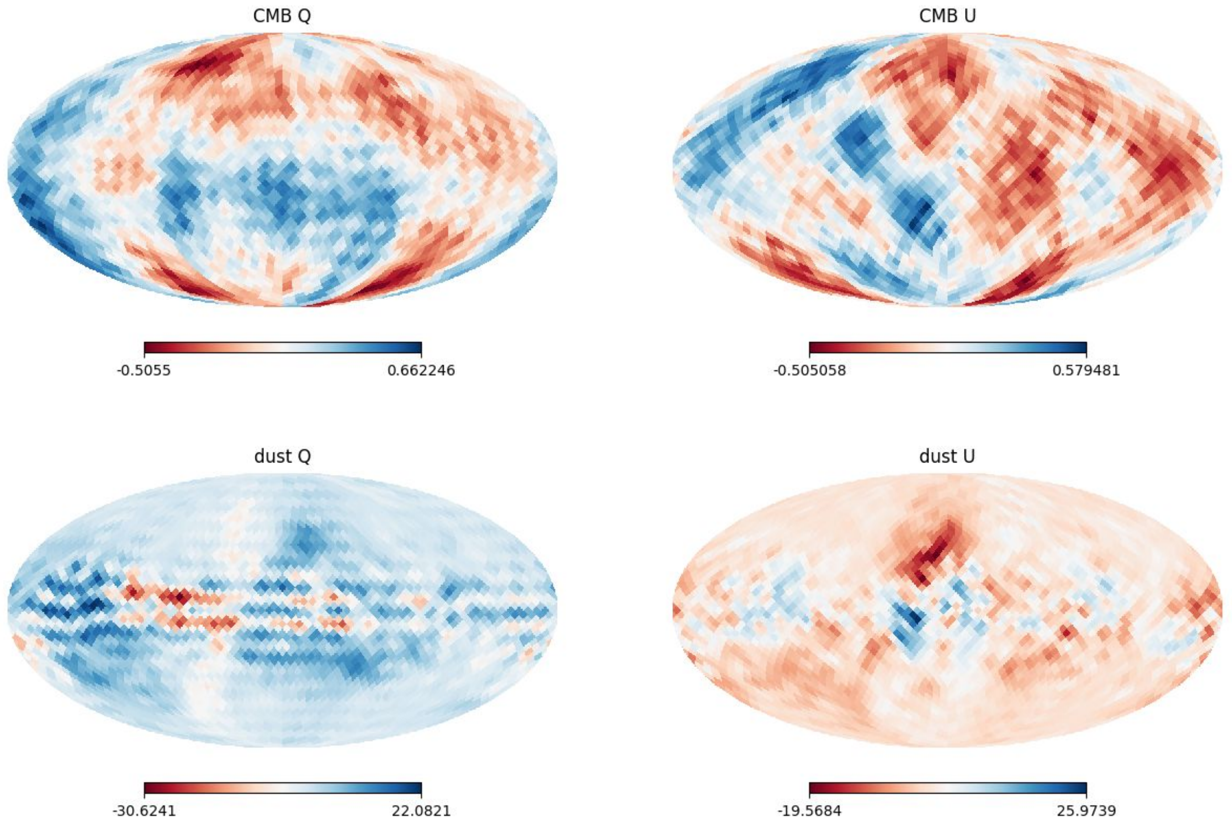
Partial sky coverage capability (work in progress)



We're working only with CMB + simple dust at very low resolution so the numbers are slightly different



In-painted sample maps !



4. Summary and outlooks

Developments…

SMICA:

- Can be optimized for more complex conditions with some caveats: *Tran et al. 2026 in prep*
- Featured in the general LiteBIRD comp-sep analysis.

MISTICA:

- Global analysis, well motivated by maximum-likelihood component separation method that allows for natural incorporation of spatial variability, partial/cut sky.
- Validations of the framework are mostly in place now.
- Preparing for a first draft of the methodology paper. (Coming out before the end of the year)

Plan for development

- Testing MISTICA in more complex conditions: d1s1, d10s5 foregrounds.
- Exploring the effect of systematics on comp-sep: beam far-sidelobes, polarization angle miscalibration
- Early 2027: Applying MISTICA on Planck data

THANKS FOR LISTENING

BACKUP SLIDES

2. Building a full SMICA-based comp-sep pipeline for complex foregrounds

MISTICA

To sample the full complex likelihood, we break it into conditional probabilities and utilize Gibbs sampling

$$\mathcal{P}(\mathbf{s}|....) \propto \left(\mathbf{s} - \hat{\mathbf{s}}_{\mathbf{WF}}\right)^T \left(\mathbf{S}^{-1} + \hat{\mathbf{N}}^{-1}\right) \left(\mathbf{s} - \hat{\mathbf{s}}_{\mathbf{WF}}\right)$$

$$\mathcal{P}(\mathbf{C}|....) \propto \mathbf{s_c}^T \mathbf{C}^{-1} \mathbf{s_c} + \ln|\mathbf{C}|$$

$$\mathcal{P}(\mathbf{F}|....) \propto \mathbf{s_f}^T \mathbf{F}^{-1} \mathbf{s_f} + \ln|\mathbf{F}|$$

$$\mathcal{P}(\mathbf{A_f}|....) \propto (\mathbf{d} - \mathbf{A_c s_c} - \mathbf{A_f s_f})^T \mathbf{N}^{-1} (\mathbf{d} - \mathbf{A_c s_c} - \mathbf{A_f s_f})$$

The sample variables are:
- The component maps: s
- The component spectra: C and F
- The foreground mixing matrix: A_f

$$\hat{\mathbf{s}}_{\mathbf{ML}} \equiv \hat{\mathbf{N}}\mathbf{A}^T\mathbf{N}^{-1}\mathbf{d}$$

$$\hat{\mathbf{N}} \equiv \left(\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A}\right)^{-1}$$

$$\hat{\mathbf{s}}_{\mathbf{WF}} \equiv (\hat{\mathbf{N}}^{-1} + \mathbf{S}^{-1})^{-1}\hat{\mathbf{N}}^{-1}\hat{\mathbf{s}}_{\mathbf{ML}}$$

MISTICA

To sample the full complex likelihood, we break it into conditional probabilities and utilize Gibbs sampling

$$\mathbf{s} = \hat{\mathbf{s}}_{\mathbf{WF}} + \zeta$$

Sample mean

$$\begin{aligned}(\mathbf{S}^{-1} + \hat{\mathbf{N}}^{-1})\hat{\mathbf{s}}_{\mathbf{WF}} &= \hat{\mathbf{N}}^{-1}\,\hat{\mathbf{s}}_{\mathbf{ML}} \\ (\mathbb{1} + \mathbf{S}^{1/2}\hat{\mathbf{N}}^{-1}\mathbf{S}^{1/2})\mathbf{S}^{-1/2}\hat{\mathbf{s}}_{\mathbf{WF}} &= \mathbf{S}^{1/2}\hat{\mathbf{N}}^{-1}\,\hat{\mathbf{s}}_{\mathbf{ML}}\end{aligned}$$

The mean is data dependent and can be cut/partial sky

Sample variance

$$(\mathbb{1} + \mathbf{S}^{1/2}\hat{\mathbf{N}}^{-1}\mathbf{S}^{1/2})\mathbf{S}^{-1/2}\zeta = \xi + \mathbf{S}^{1/2}\mathbf{A}^T\mathbf{N}^{-1/2}\chi$$

Eta and chi are random vectors and weighted to have the correct variance. These are always generated as full sky so they inpaint the masked part

$$\begin{aligned}\hat{\mathbf{s}}_{\mathbf{ML}} &= \hat{\mathbf{N}}\mathbf{A}^T\mathbf{N}^{-1}\mathbf{d} \\ \hat{\mathbf{N}} &\equiv \left(\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A}\right)^{-1} \\ \hat{\mathbf{s}}_{\mathbf{WF}} &\equiv (\hat{\mathbf{N}}^{-1} + \mathbf{S}^{-1})^{-1}\hat{\mathbf{N}}^{-1}\hat{\mathbf{s}}_{\mathbf{ML}}\end{aligned}$$

To sample the full complex likelihood, we break it into conditional probabilities and utilize Gibbs sampling

$$\mathrm{P}(\mathbf{C}_\ell|\mathbf{s_c}) \propto \prod_\ell \frac{1}{\sqrt{|\mathbf{C}_\ell|^{2\ell+1}}} \exp\left(-\frac{1}{2}\,\mathrm{tr}\,\boldsymbol{\sigma}_\ell \mathbf{C}_\ell^{-1}\right), \quad (24)$$

where $\boldsymbol{\sigma}_\ell \equiv \sum_{m=-\ell}^{\ell} \mathbf{a}_{\ell m}\mathbf{a}_{\ell m}^{\dagger}$, and $\mathbf{a}_{\ell m}$ are the spherical harmonic coefficients of $\mathbf{s_c}$, given as:

$$\mathbf{a}_{\ell m}^{\mathrm{T}} = \left[\mathbf{a}_{\mathbf{E},\ell m}^{\mathrm{T}},\ \ \mathbf{a}_{\mathbf{B},\ell m}^{\mathrm{T}}\right], \quad (25)$$

and $\mathbf{C}_\ell$ given by,

$$\mathbf{C}_\ell = \begin{pmatrix} \mathbf{C}_\ell^{\mathbf{EE}} & \mathbf{C}_\ell^{\mathbf{EB}} \\ \mathbf{C}_\ell^{\mathbf{BE}} & \mathbf{C}_\ell^{\mathbf{BB}} \end{pmatrix}. \quad (26)$$

where $\mathbf{C}_\ell^{\mathbf{EB}} = \mathbf{C}_\ell^{\mathbf{BE}}$.

Spectral Matching ICA

SMICA (*Delabrouille et al. 2002*) is one of the **reference component separation** of the **Planck** experiment (*Planck Collaboration 2016*)

$$y_\nu(\vec{r}) = \sum_{j=1}^{N_c} A_{\nu j} s_j(\vec{r}) + n_\nu(\vec{r})$$

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

Modeled spectrum:

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

$$\langle YY^T \rangle \qquad diag(C_j(\ell)) \qquad diag(N_\nu(\ell))$$

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:

$$\theta = (A, C_j(q), N_\nu(q))$$

From these parameters we can build the set of comp sep weights

SMICA with effective components

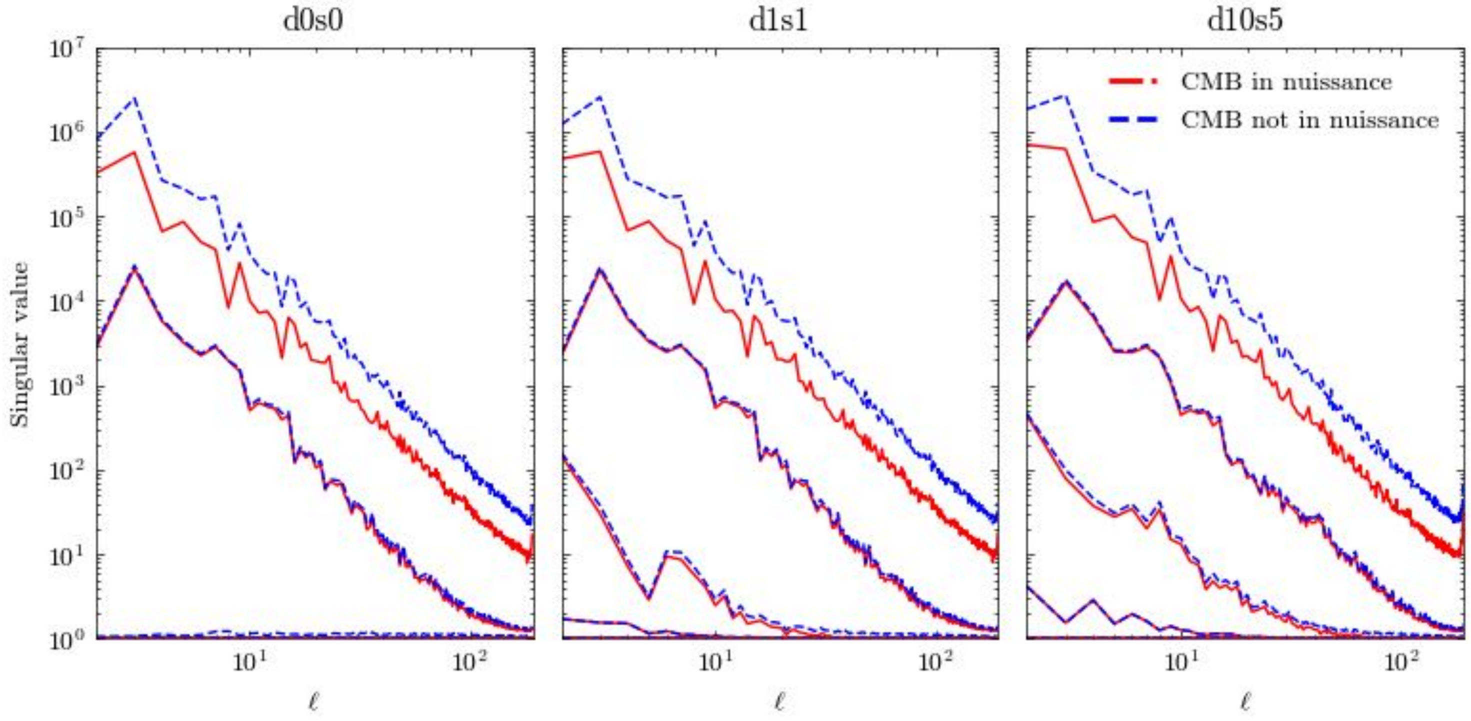
We can decompose the covariance into the CMB, foregrounds and noise part

$$\mathbf{A} = \left(\mathbf{a_c}\ \mathbf{A_f}\right),$$
$$\mathbf{R_S} = \begin{pmatrix} \mathbf{C} & 0 \\ 0 & \mathbf{P_f} \end{pmatrix},$$
$$\mathbf{R_Y} = \mathbf{a_c}\mathbf{a_c}^T\mathbf{C} + \mathbf{A_f}\mathbf{P_f}\mathbf{A_f}^T + \mathbf{R_N}.$$

$$\mathbf{R_C} = \mathbf{a_c}\mathbf{a_c}^T\mathbf{C} \qquad \mathbf{R_f} = \mathbf{A_f}\mathbf{P_f}\mathbf{A_f}^T$$

Construct whitened covariance

$$\tilde{\mathbf{R}}_\mathbf{Y} = (\mathbf{R_C}+\mathbf{R_N})^{-1/2}\mathbf{R_Y}(\mathbf{R_C}+\mathbf{R_N})^{-1/2}$$
$$= (\mathbf{R_C}+\mathbf{R_N})^{-1/2}\mathbf{R_f}(\mathbf{R_C}+\mathbf{R_N})^{-1/2} + \mathbb{I},$$



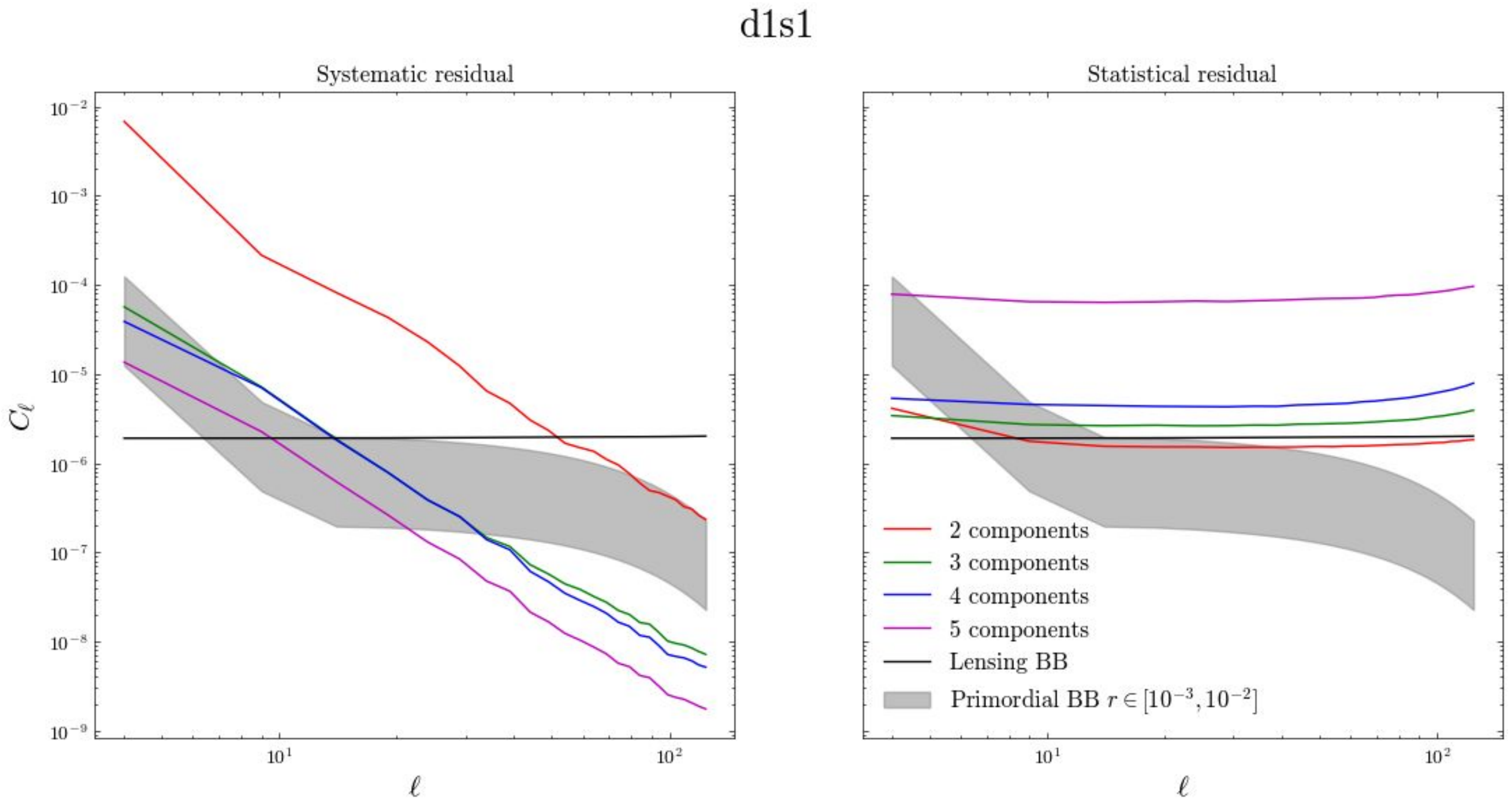
SVD on the whitened covariance in 2 cases:

- CMB in nuissance (with simulation)
- CMB not in nuissance (realistic)

3 components seem to be optimal for complex foregrounds

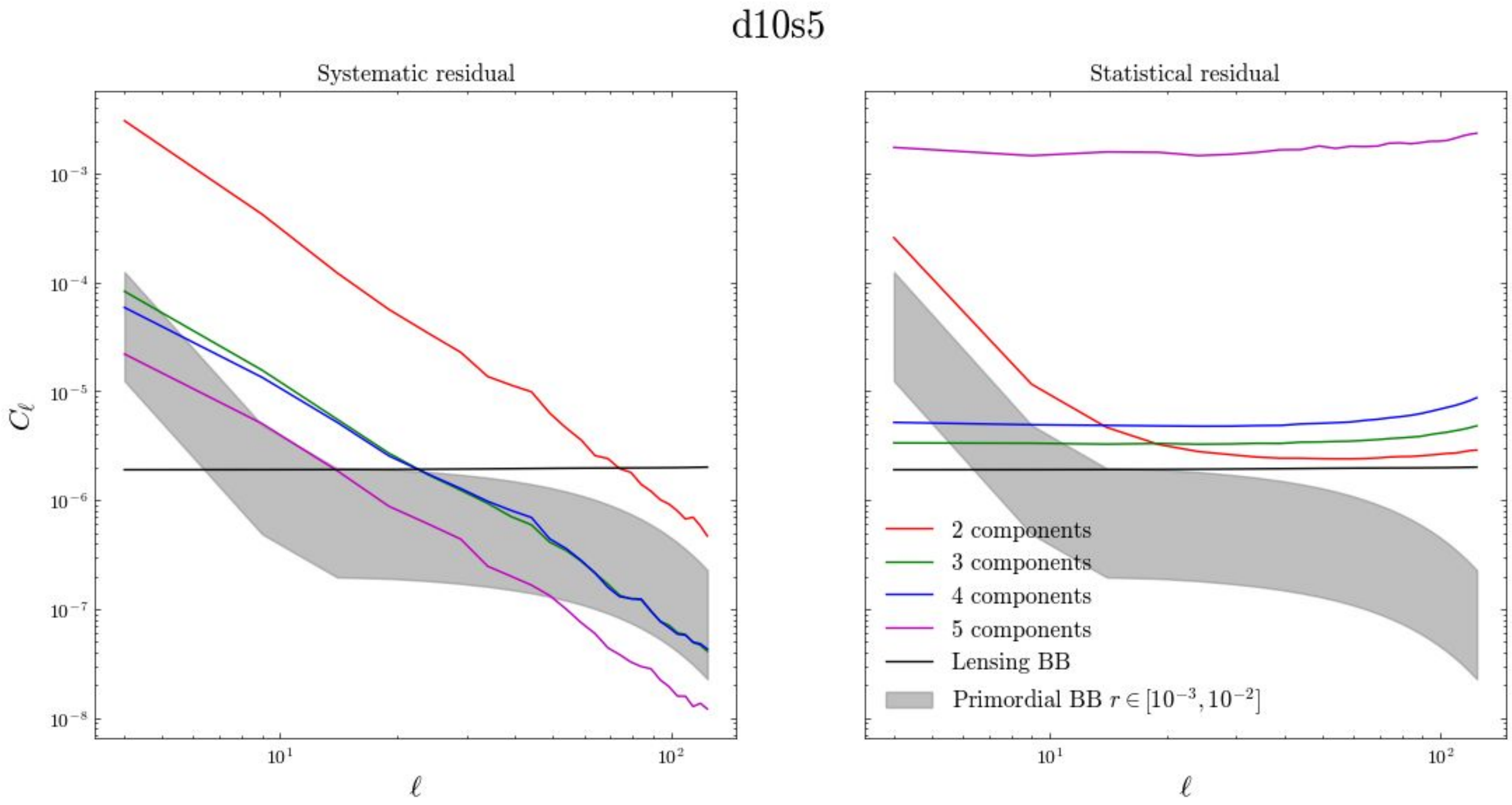
SMICA with effective components

Complex foregrounds can be decomposed into a number of modes (depending on the sensitivity & the complexity of the sky)
=> Simply adding more components into SMICA can help reduce the bias coming from spatial variability of the mixing matrix



SMICA with effective components

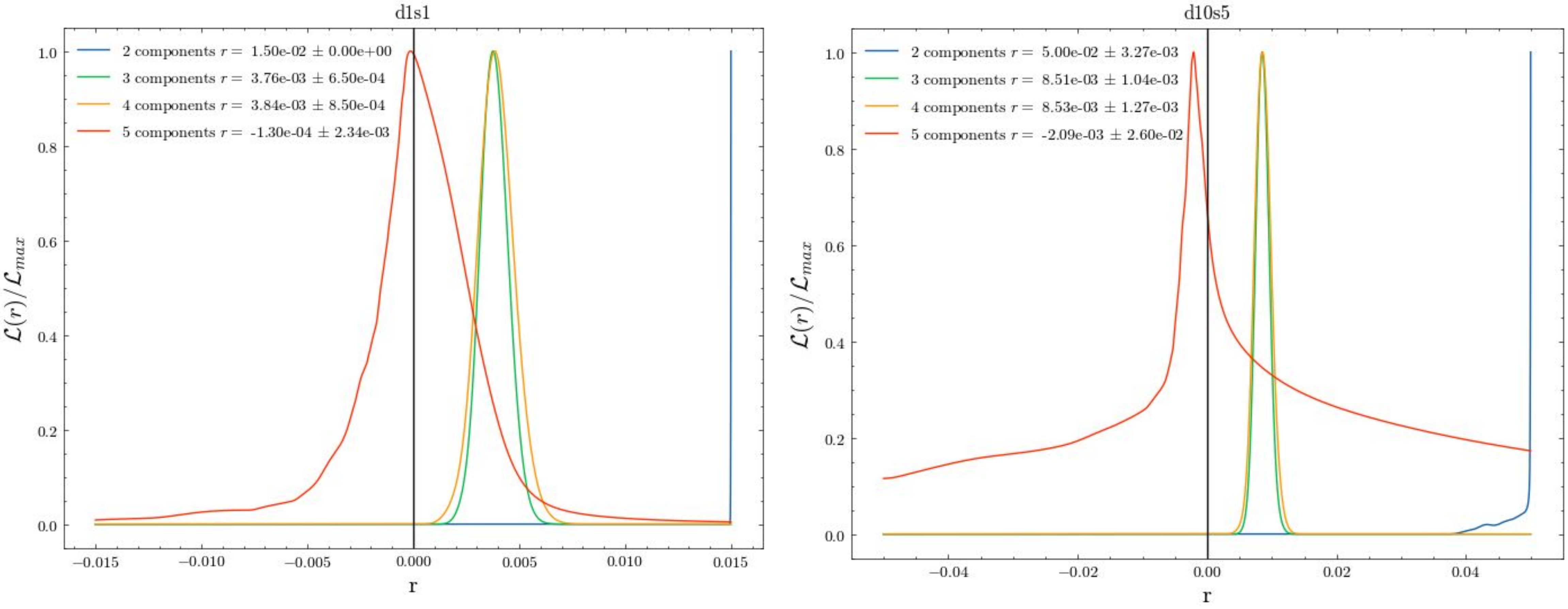
Complex foregrounds can be decomposed into a number of modes (depending on the sensitivity & the complexity of the sky)
=> Simply adding more components into SMICA can help reduce the bias coming from spatial variability of the mixing matrix



SMICA with effective components

Even in our best performing cases, a bias was found for the complex foregrounds…

Average posteriors for 200 realizations:



SMICA with clusters

Frequency maps

↓

Divide into Healpix patches
Calculate covariance matrices on each patches

↓

Run SMICA on each patches independently
Fit for A, N

↓

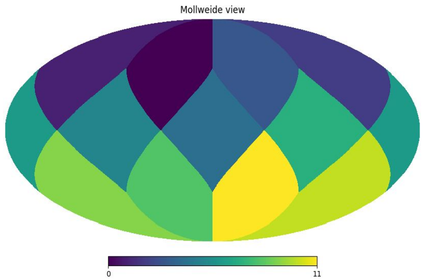
Filter to build map for each patches using fitted A, Rs, RN

↓

Sum up the results to create a full sky map

↓

Cl and r analysis



+ Masking (none for d0s0)

Assumptions: White noise, unit vector for CMB
2 foreground components for d0s0
3 foreground components for d1s1 and d10s5

$$\hat{\mathbf{s}} = (\mathbf{A}^T \mathbf{R_N}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{R_N}^{-1}\mathbf{d} = \mathbf{W}\mathbf{d}.$$

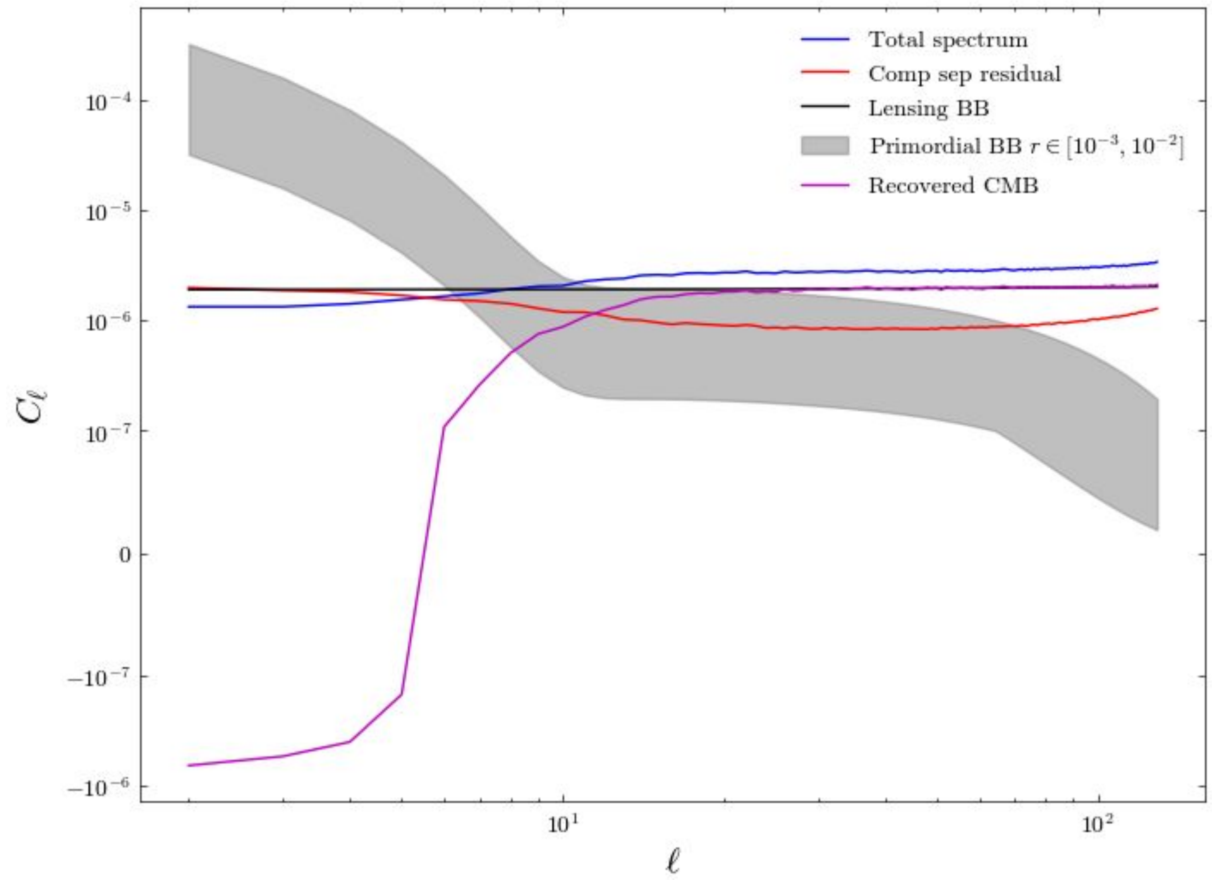
This approach of going back and forth between pixel and harmonic domain is artificial and does not correspond to maximizing any likelihood.

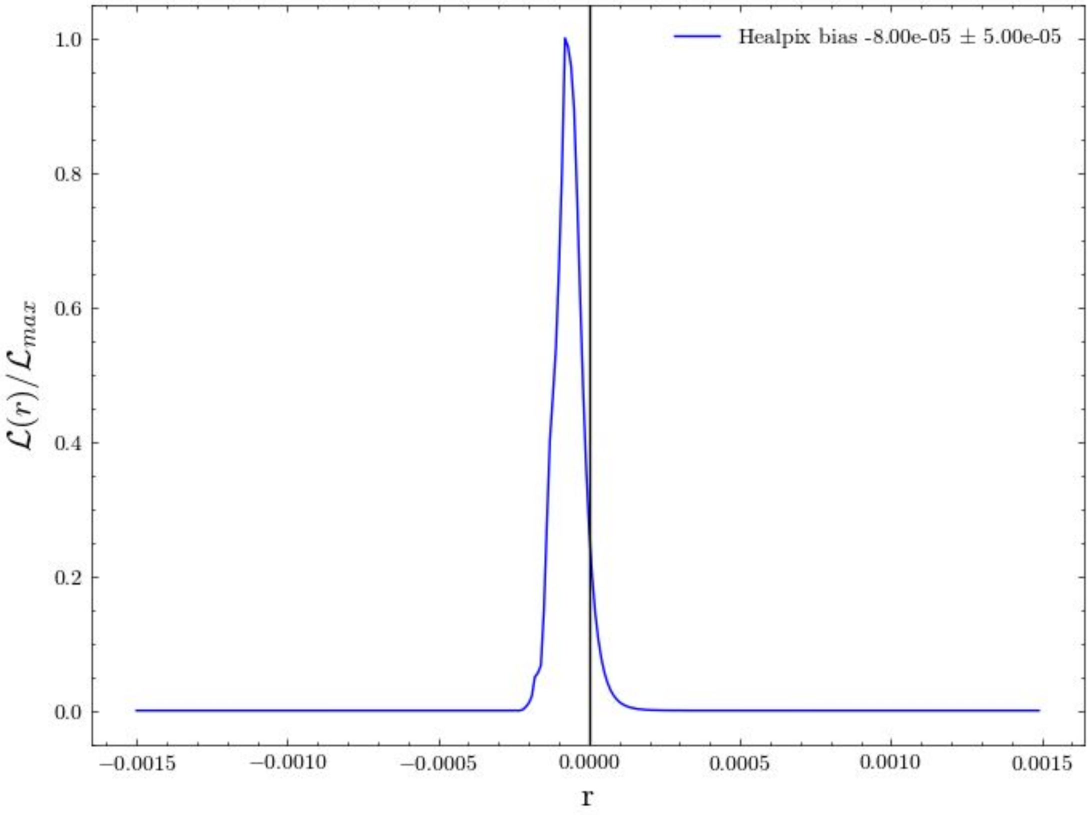
=> It leads to a bias that we have to correct for!

SMICA with clusters

ILC-bias-like issue: If the same data was used for computing covariance and weighting, a suppression of the output CMB signal appears

=> Negative bias on r

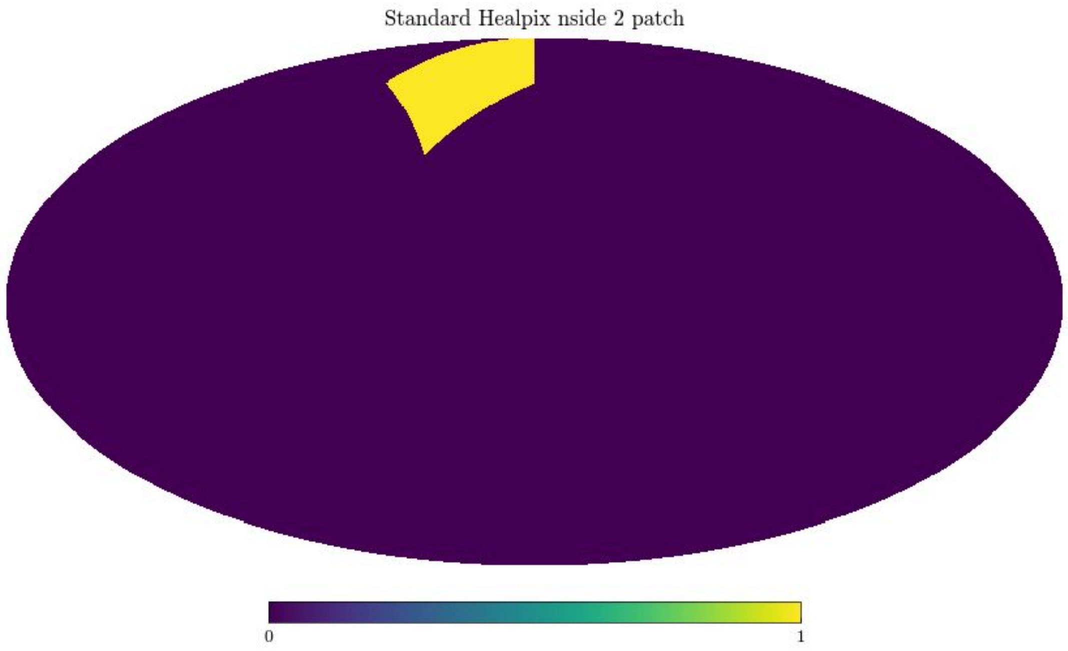


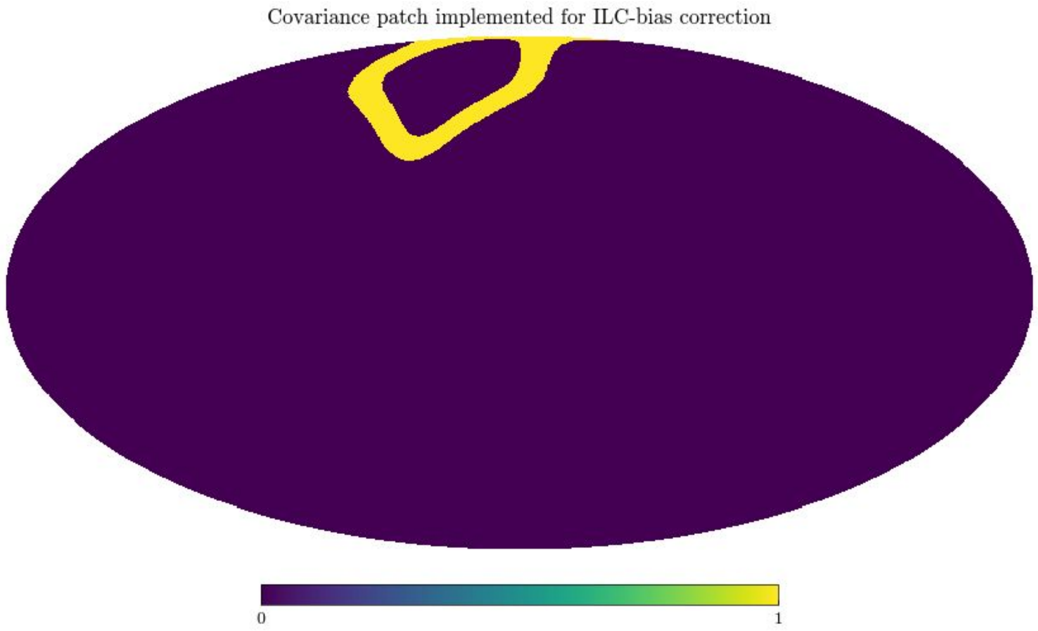


SMICA with clusters

We implement a correction inspired by Coulton et al. 2023 https://arxiv.org/abs/2307.01258

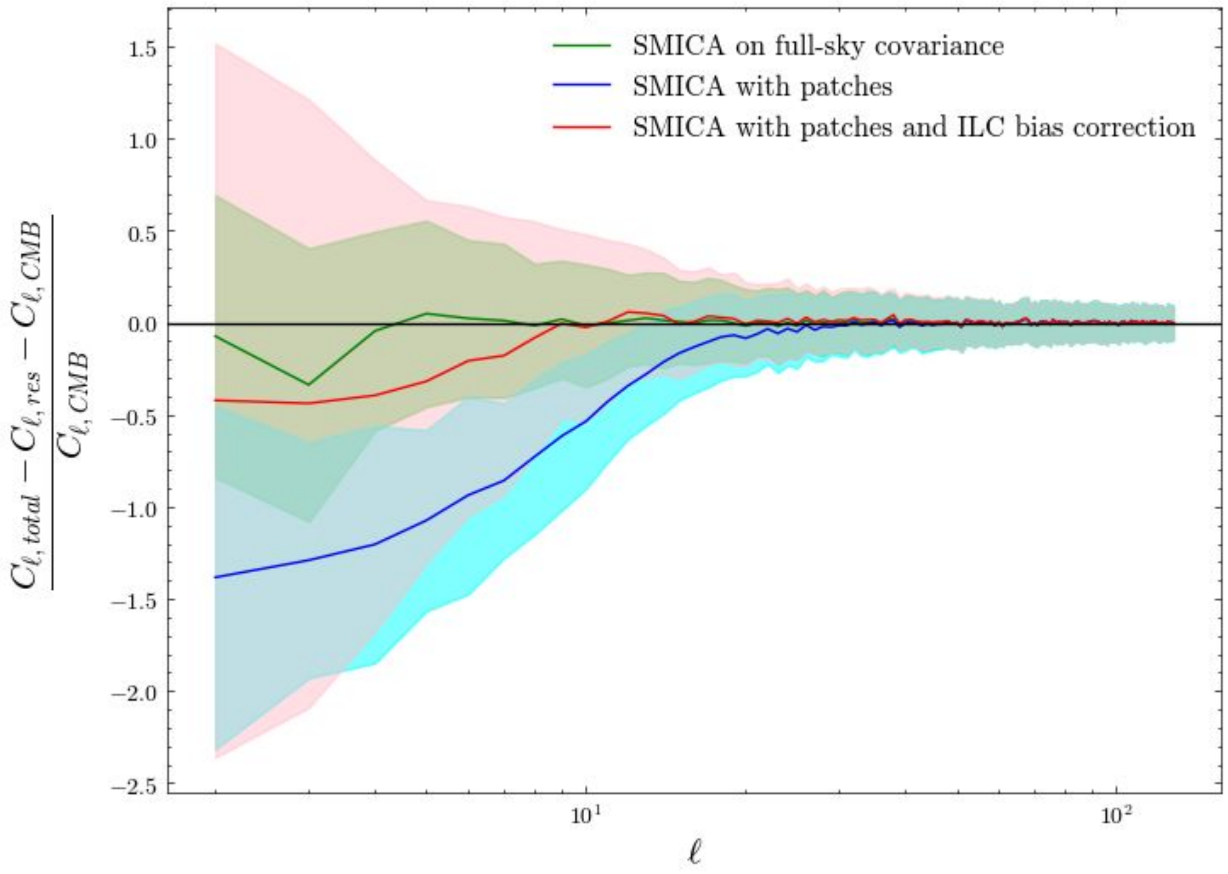
The covariance is estimated on an extended region that excludes the subset itself to avoid a double use of the data that leads to the usual ILC-bias



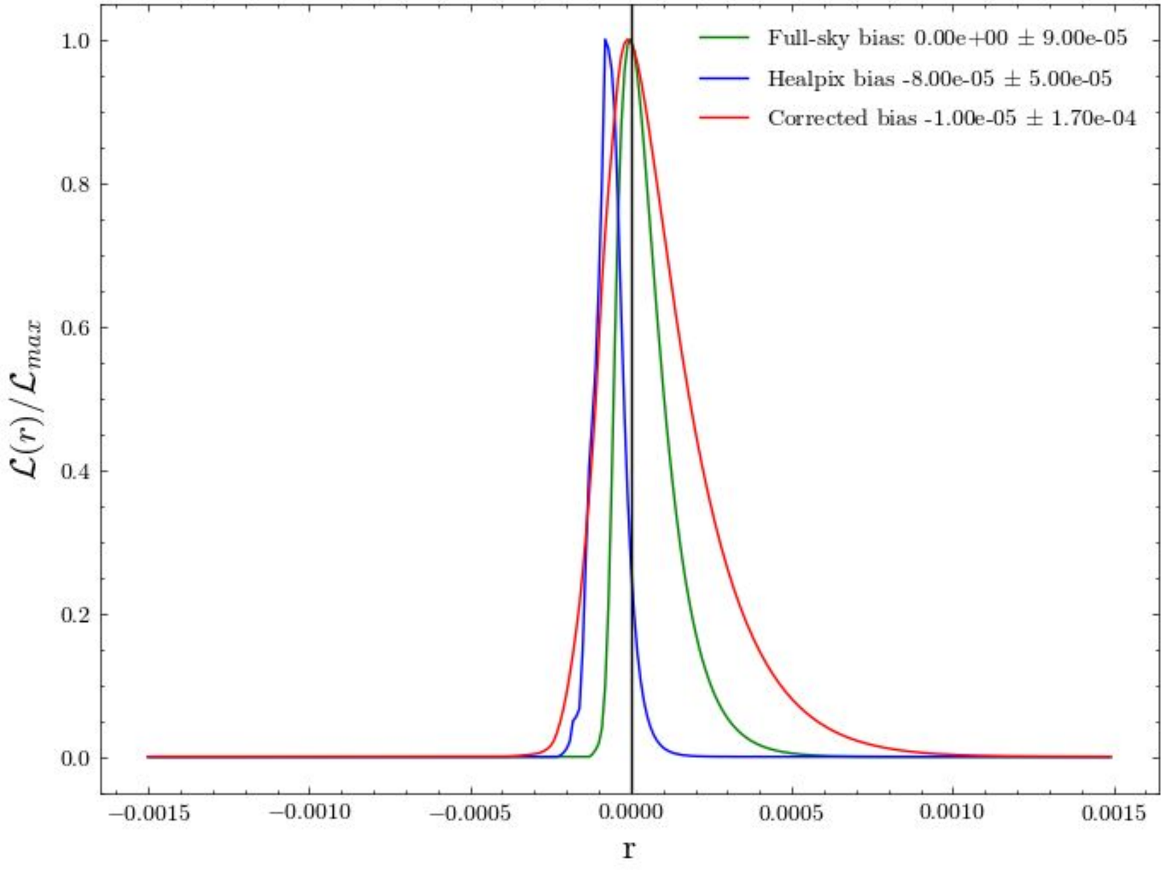


SMICA with clusters

By implementing this simple correction, we reduced the effect of the ILC bias with a cost of increasing the residuals.



Note that this is not the actual error bar on r, this is the uncertainty of the ILC-bias signal

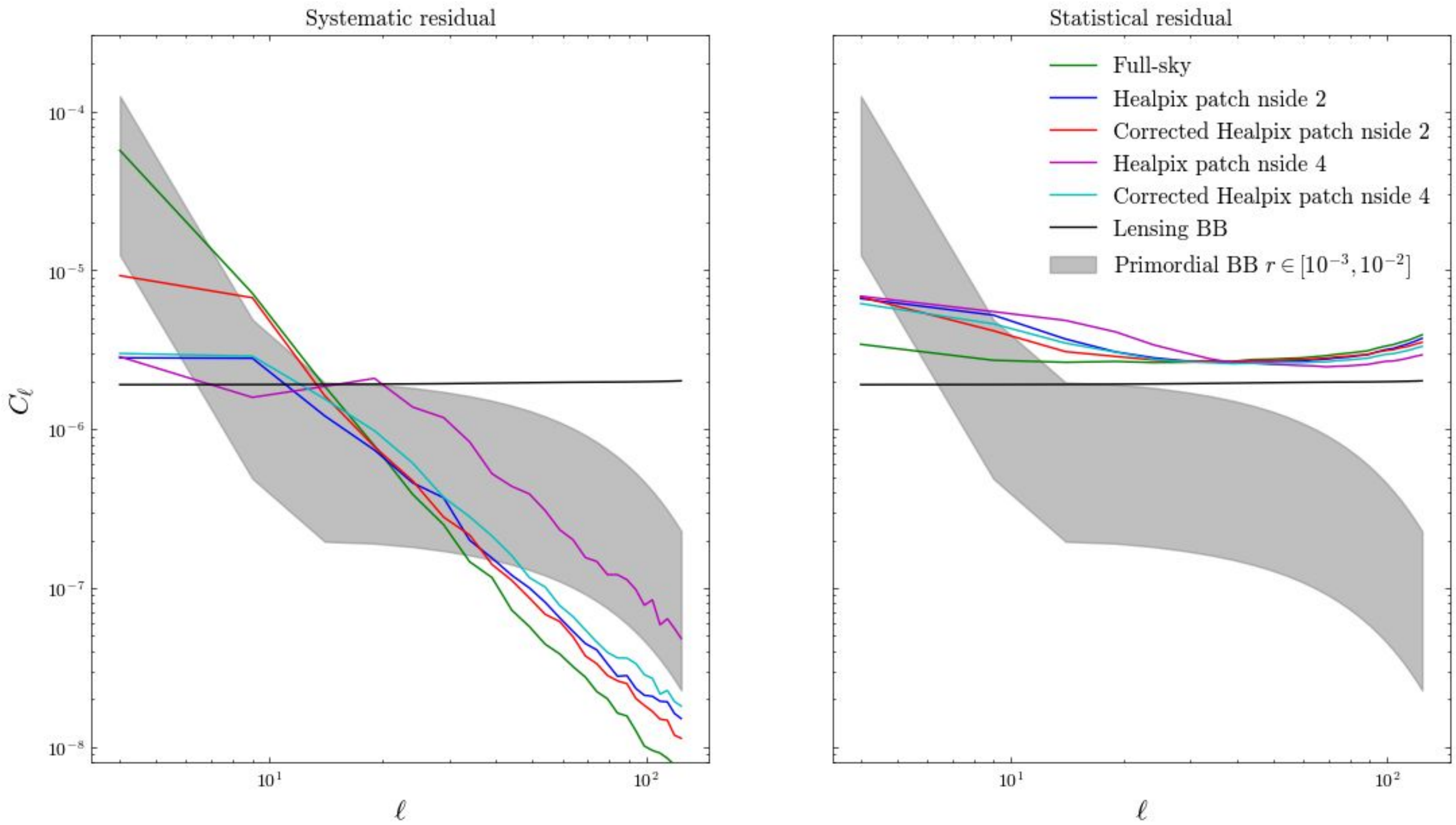


SMICA with clusters

Performance with d1s1 (3 foreground components)

The optimization of including patches help reduce a lot the systematic residuals, at a cost of increasing the statistical residual

The correction scheme reduces the ILC bias but increases our systematic residuals.

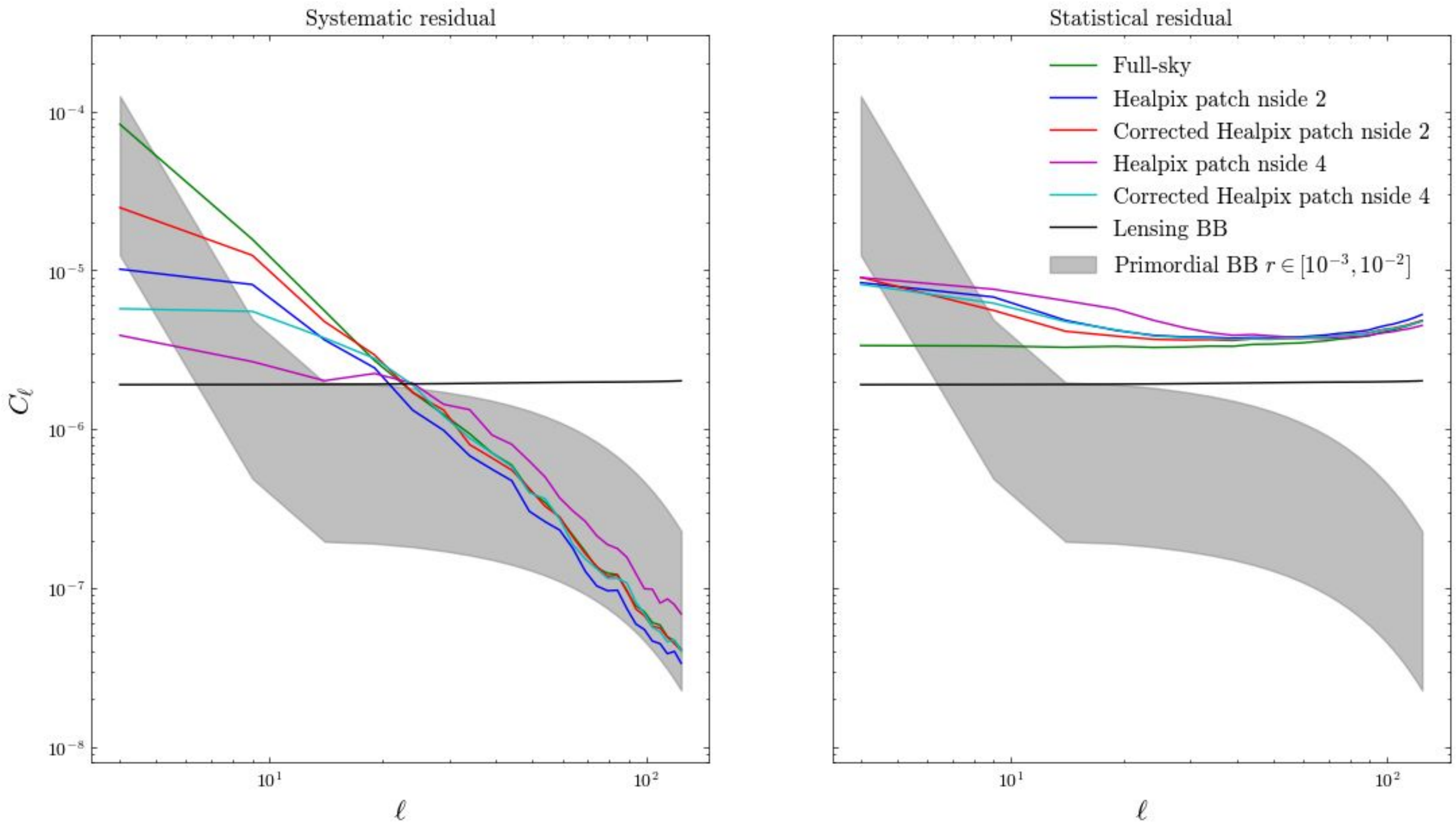


SMICA with clusters

Performance with d10s5 (3 foreground components)

The optimization of including patches help reduce a lot the systematic residuals, at a cost of increasing the statistical residual

The correction scheme reduces the ILC bias but increases our systematic residuals.



SMICA with clusters

Conclusion

Many type of residuals and effects are contributing to our final results with this pipeline. There are a few notable things:

- The best option in terms of minimizing residual is the **uncorrected Healpix case**. However, we are heavily contaminated by the ILC-bias in the output signal and it may lead to **underestimation** in terms of both **bias and uncertainty.**
- The **bias correction** always degrade the performance of the method. Despite this, because of the nature of the ILC bias (the suppression is dependent on foregrounds model, comp-sep performance…), **in the case of real data, this is the safer option to take.**
- The remaining systematic residual can be reduced through **masking strategy** (the current one is slightly pessimistic) and **template marginalization** at the likelihood level (WIP)

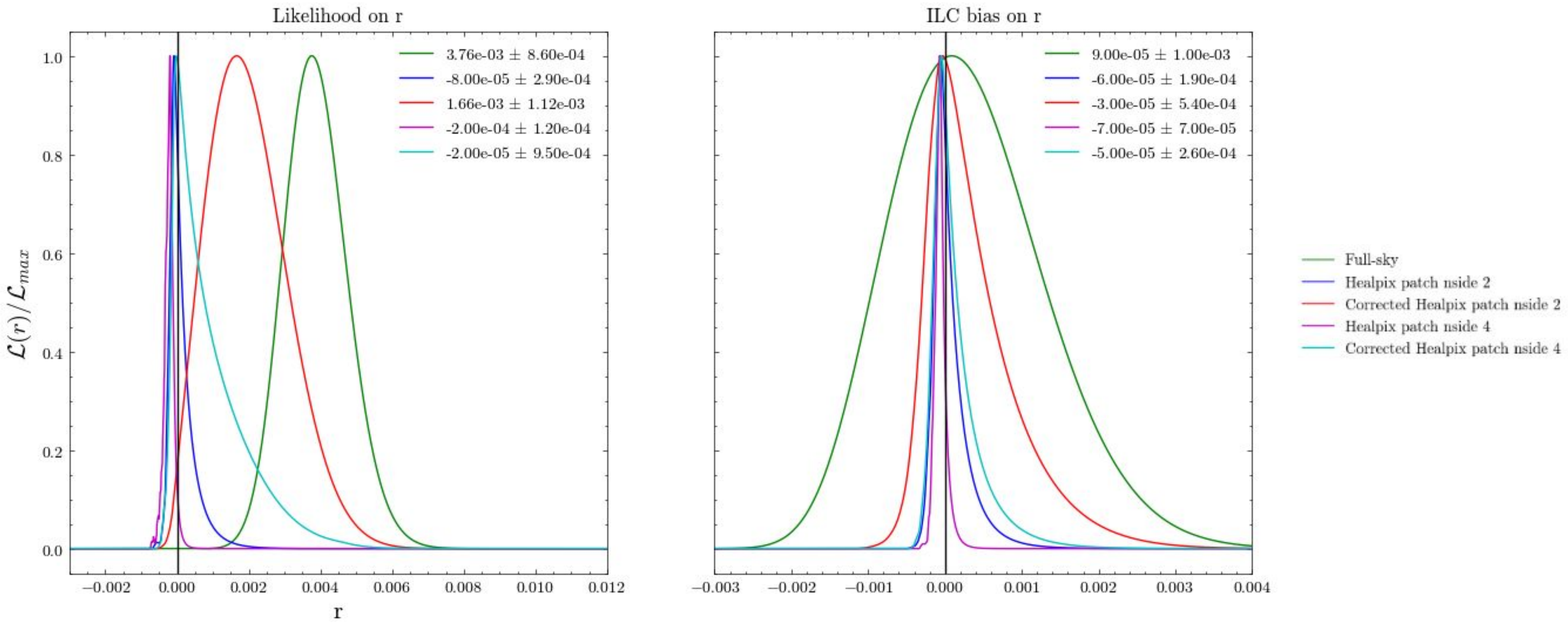
The method works and can perform comp-sep to a good level. However, the pipeline is full of nuances and inconsistencies leading to biases…

=> There is motivation for a global, consistent maximum likelihood framework

SMICA with clusters

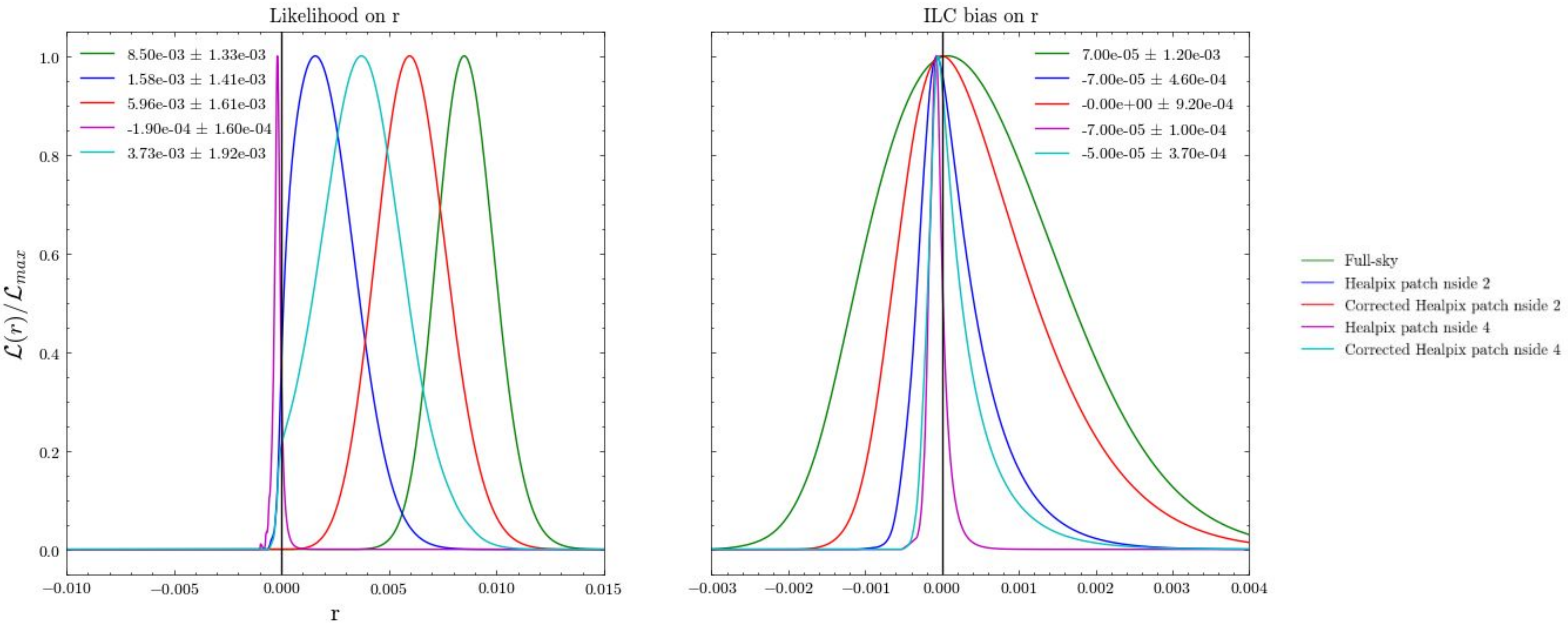
Performance with d1s1

The optimization help reduce a lot the systematic residuals, at a cost of increasing the statistical residual



SMICA with clusters

Performance with d10s5
Similar to d1s1, reduction in systematic residual, increase in statistical residual



Why SMICA?

SMICA was developed at APC
Reference method for temperature in Planck

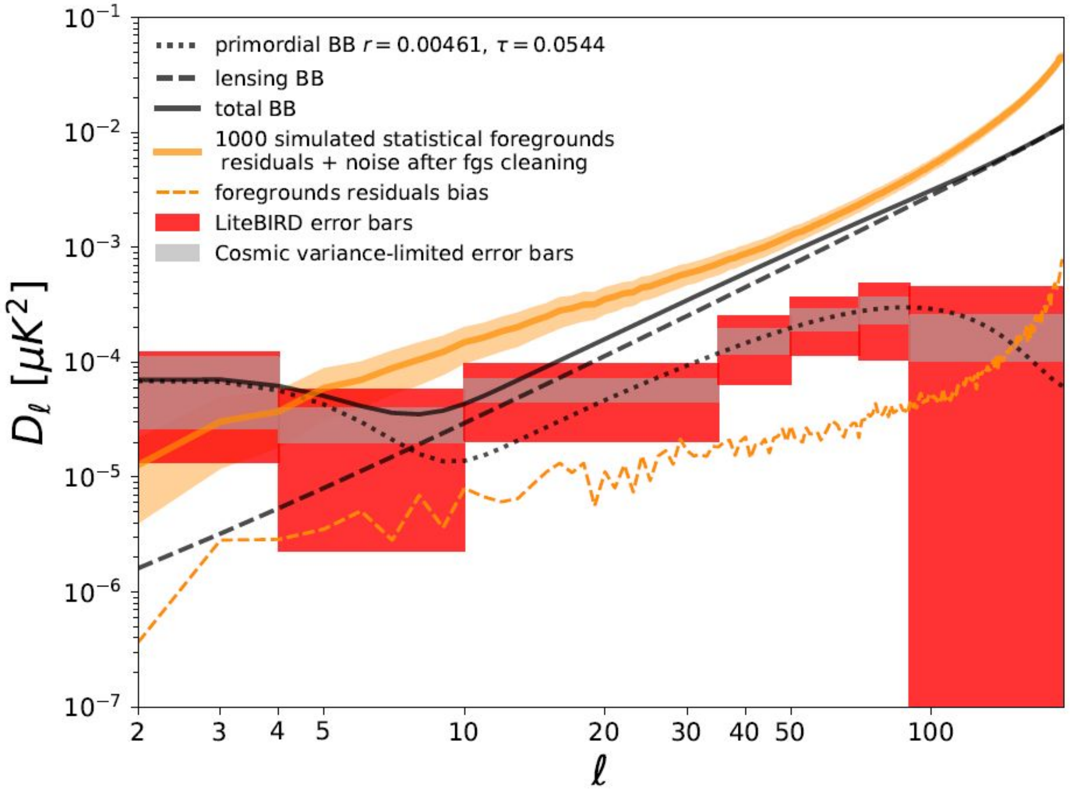
Component separation can be roughly divided into 2 classes:

- **Parametric**: makes strong assumption on mixing matrix, doesn't make assumption on the component
- **Blind**: use statistical techniques (components are assumed to be independent)

SMICA belongs to the second class

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



Starting point

Simulations: Produced locally, using PySM for fgs and PTEP sensitivity for noise

SMICA implementation: Planck’s SMICA by Maude, toolbox currently on Gitlab => Adapt for LiteBIRD polarization measurement context