BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY
Office of Science

SMICA for Unbiased Primordial Gravitational Wave Inference from the CMB

Alexander Steier, Shamik Ghosh, and **Jacques Delabrouille**

Now starting his Master's in Munich

CMB-France, October 15, 2025

CNRS • Berkeley
Centre Pierre Binétruy

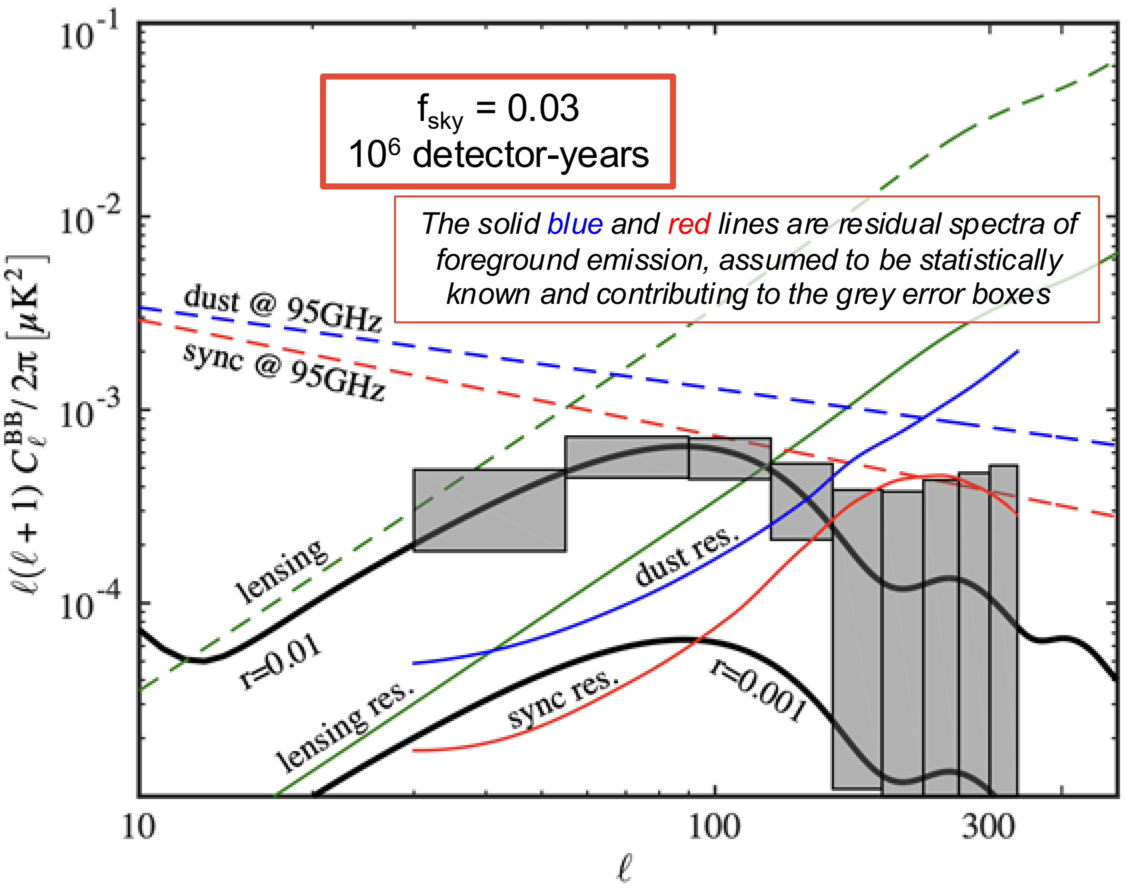
The Problem

CONTEXT: *CMB-S4 forecasts*

Measuring Primordial Gravitational Waves is hard!

- The theorized CMB polarization signal is very small
- Strong foregrounds
- Lensing of CMB
- Detector noise
- Instrumental systematics
- Cosmic variance

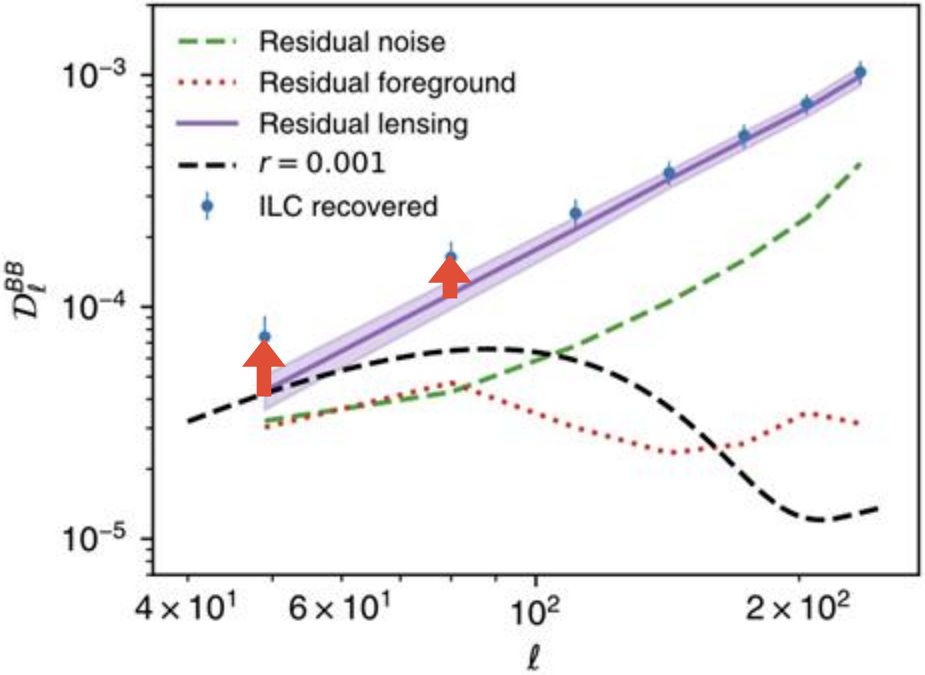
[CMB-S4 Science Book, Figure 7]



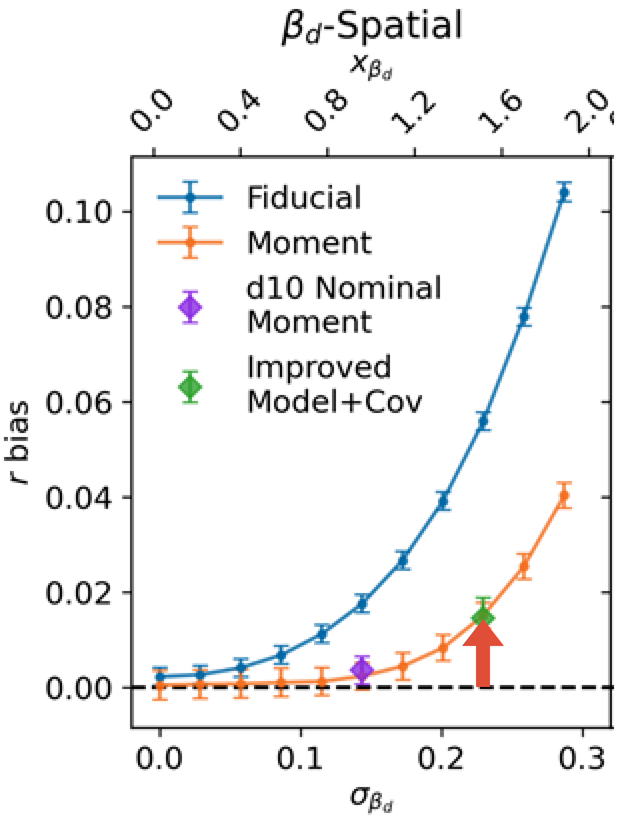
Some existing results

Foreground Bias!

[CMB-S4 Chile Optimization Report (DRAFT)]



SO-like observations, spectral fits
[Liu et al. 25, arXiv 2508.00073, Figure 9]



Why are the forecasts biased?

- NILC tries to recover the lowest variance map, not the lowest foreground map.
- Spectral fits usually assume a functional form in harmonic space and frequency.

For instance:

$$\frac{I_c(\nu,\ell)}{I_c(\nu_0,\ell_0)} = \left(\frac{\ell}{\ell_0}\right)^{\alpha} \frac{F_c(\nu,T,\beta)}{F_c(\nu_0,T,\beta)}$$

parameters

- Relax these assumptions and let the data speak for themselves: **SMICA !** [1][2][3]

[1] Delabrouille, Cardoso, Patanchon 2003
[2] Cardoso et al. 2008
[3] Steier, Ghosh, Delabrouille 2025 (in prep)

The Method

What is SMICA?

Spectral **M**atching **I**ndependent **C**omponent **A**nalysis[1][2][3]

Component separation tool applied to multivariate power spectra

linear superposition model

signals (vectorized over detectable components)

$$\boldsymbol{d}_{\ell m} = \mathbf{A}\boldsymbol{s}_{\ell m} + \boldsymbol{n}_{\ell m}$$

mixing matrix

beam-corrected noise (vectorized over frequency)

data covariance matrix

de-beamed $a_{\ell m}$ (vectorized over frequency)

$$\hat{\mathbf{C}}_q = \frac{1}{n_q} \sum_{\ell \in \mathcal{D}_q} \sum_m \hat{\mathbf{d}}_{\ell m} \hat{\mathbf{d}}_{\ell m}^\dagger$$

number of modes

ℓ-binning

$$n_q = \left((\ell_{\max}+1)^2 - (\ell_{\min})^2\right) f_{\rm sky}$$

[1] Delabrouille, Cardoso, Patanchon 2003
[2] Cardoso et al. 2008
[3] Steier, Ghosh, Delabrouille 2025 (in prep)

The SMICA Covariance Model

$$\underbrace{\hat{\mathbf{C}}_q}_{\text{Data}} \approx \underbrace{\mathbf{C}_q(\theta) = \mathbf{A}\mathbf{S}_q\mathbf{A}^\dagger + \mathbf{N}_q}_{\text{Model}}$$

Each element in **A** and $\mathbf{S}_q$ is an independent parameter (except CMB power spectrum).

$$C_\ell^{\mathrm{CMB}} = rC_\ell^{\mathrm{tens}\, r=1} + A_{\mathrm{lens}} C_\ell^{\mathrm{lens}}$$

Any part of the model can be parameterized, we implement the most general case.

$\mathbf{C}_q$	=	$\mathbf{A}$	$\mathbf{S}_q$	$\mathbf{A}^T$	+	$\mathbf{N}_q$
Model Covariance Matrix		Mixing Vectors	Auto & Cross Spectra	CMB Vector (uniform) Dust Vector Synchrotron Vector Any Additional Vectors		Noise Covariance Matrix (diagonal)
$(n_f \times n_f)$		$(n_f \times n_c)$	$(n_c \times n_c)$	$(n_c \times n_f)$		$(n_f \times n_f)$

For Q ℓ-bins with n_f frequency channels & n_c components

Implementation, tests and forecasts

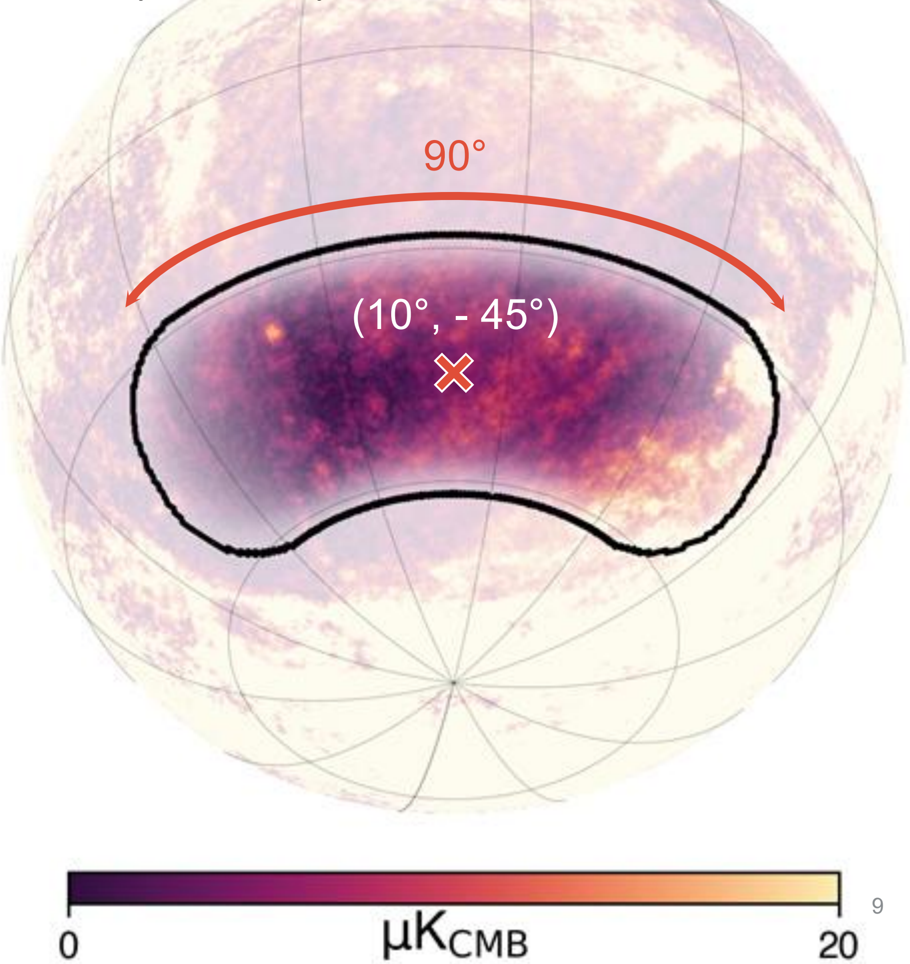
Sky Patch

Hit map based on constant dec SAT scans

Small Sky fraction!

$$f_{\mathrm{sky}} = \frac{1}{n_{\mathrm{pix}}} \sum_p w_p^2 = 2.5\%$$

PySM3 high-complexity polarized intensity at 270 GHz
dominated by d12 multilayer dust, Martinez-Solaeche et al, 2018



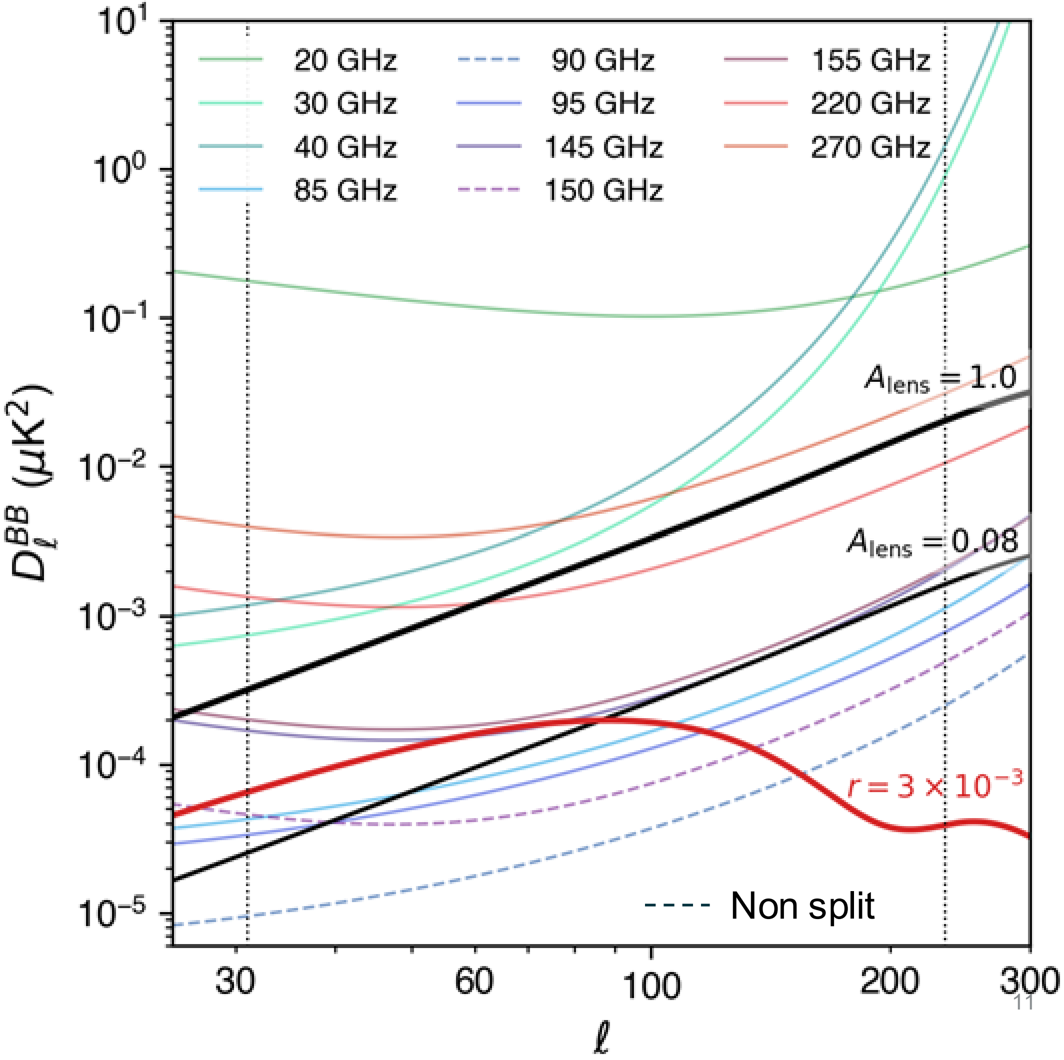
Experiment Configuration (Stage 4)

Frequency (GHz)	θ_{FWHM} (arcmin)	Noise Δ_P (μK-arcmin)	ℓ_{knee}	α_{knee}	Configuration
20	11	13.6	150	2.7	Both
30	73	3.53	60	1.7	
40	73	4.46	60	1.7	
85	26	0.88	60	1.7	Split
95	23	0.78	60	1.7	
145	26	1.23	60	3.0	
155	23	1.34	60	3.0	
90	26	0.42	60	1.7	Non-Split
150	23	0.64	60	3.0	
220	13	3.48	60	3.0	Both
270	13	5.97	60	3.0	

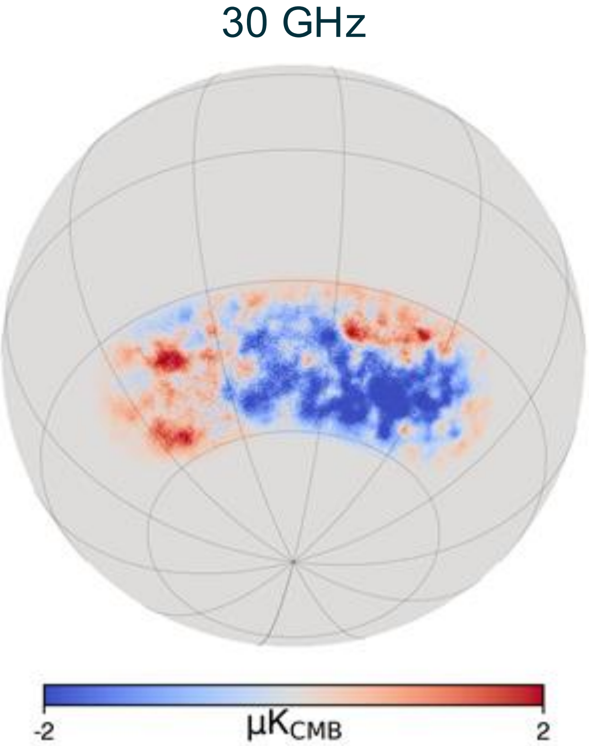
Inputs to the Pipeline

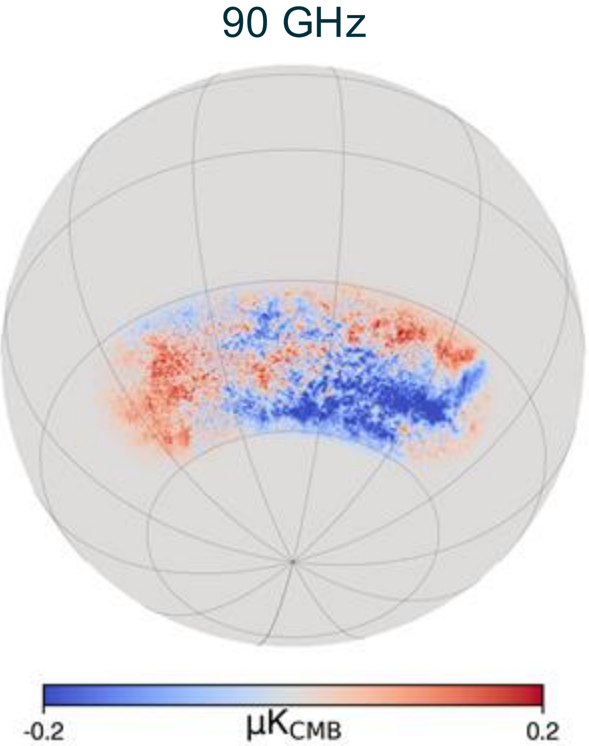
96 independent realizations of IQU maps:

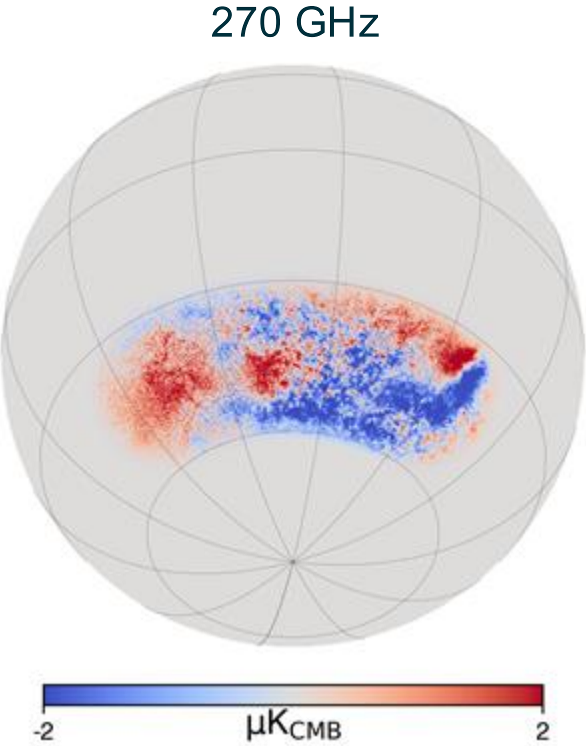
- CMB lensing (and tensor) modes
- PySM3 Foregrounds (low, medium and high complexity)
- Spatially anisotropic white noise
- "Split bands" and "non-split bands" configurations
- For simplicity we use delta passbands in ν, but this has little incidence on the method



Some B-mode Input Maps (high complexity, r = 0.003, A_{lens}=0.08)







Evaluating the SMICA Likelihood

$$\ln \mathcal{L}(\theta) = -\frac{1}{4} \sum_{q=1}^{Q} n_q D(\hat{\mathbf{C}}_q, \mathbf{C}_q(\theta))$$

https://healpix.sourceforge.io

https://github.com/jax-ml/jax

where the Kullback-Leibler divergence for two *n* x *n* matrices is

$$D(\mathbf{R}_1, \mathbf{R}_2) = \mathrm{tr}\left(\mathbf{R}_1 \mathbf{R}_2^{-1}\right) - \ln\det\left(\mathbf{R}_1 \mathbf{R}_2^{-1}\right) - n$$

and n_q is the number of independent modes in bin q

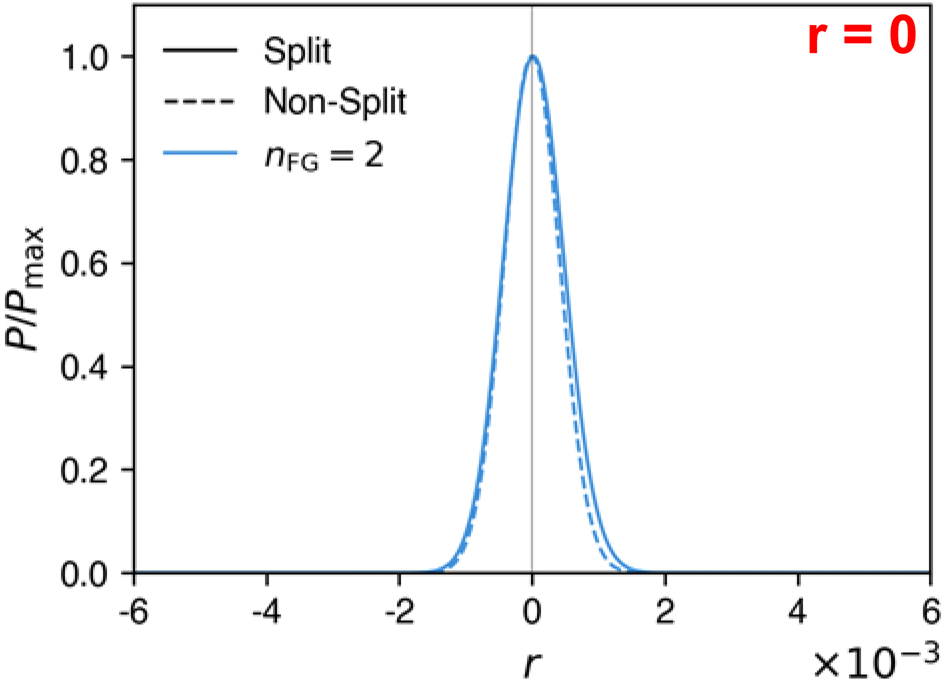
$$n_q = \left((\ell_{\max} + 1)^2 - (\ell_{\min})^2\right) f_{\mathrm{sky}}$$

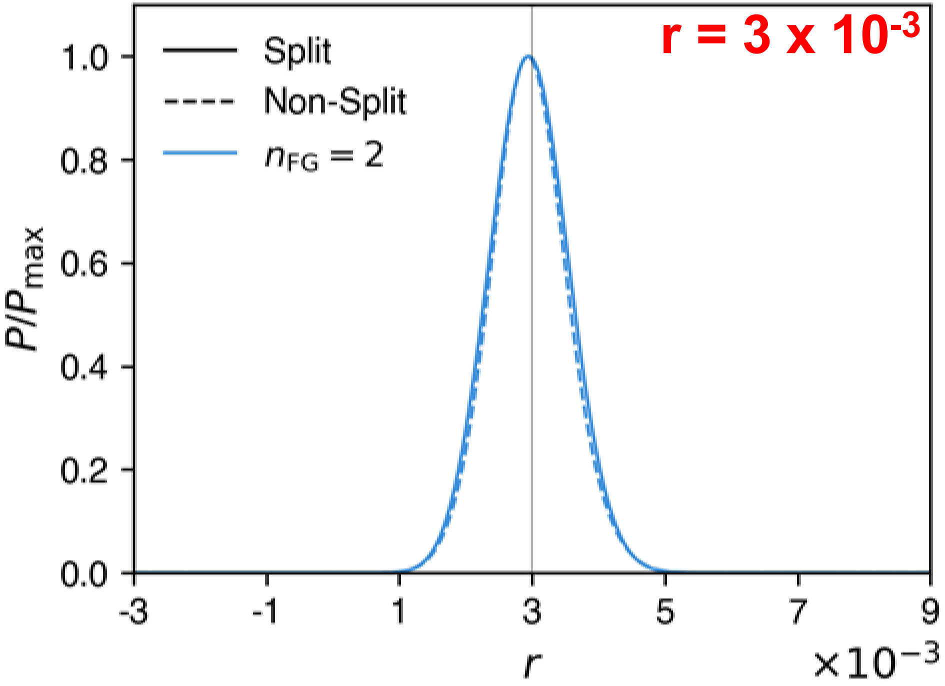
https://blackjax-devs.github.io/blackjax/

mock observations: 96 independent IQU maps → healpy anafast: covariance matrices → JAX framework: model is differentiable → BlackJAX: MC sampling in parallel → Results: Posteriors for all parameters

(96 realizations) (5 spectral bins) No-U-Turn-Sampler (NUTS)

r-Posterior - Low Complexity ("perfect" test case)

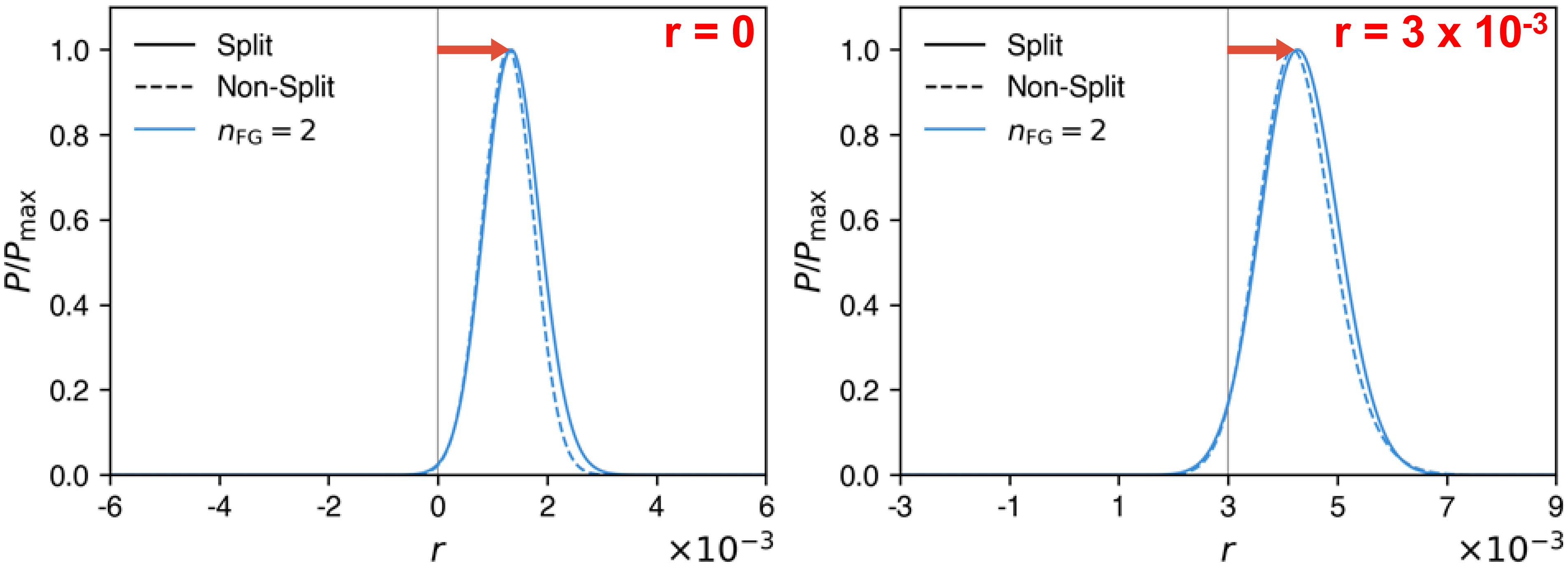




Split:	0.0 ± 0.3	2.9 ± 0.5
Non-Split:	0.0 ± 0.3	2.9 ± 0.5

r-Posterior - Medium Complexity

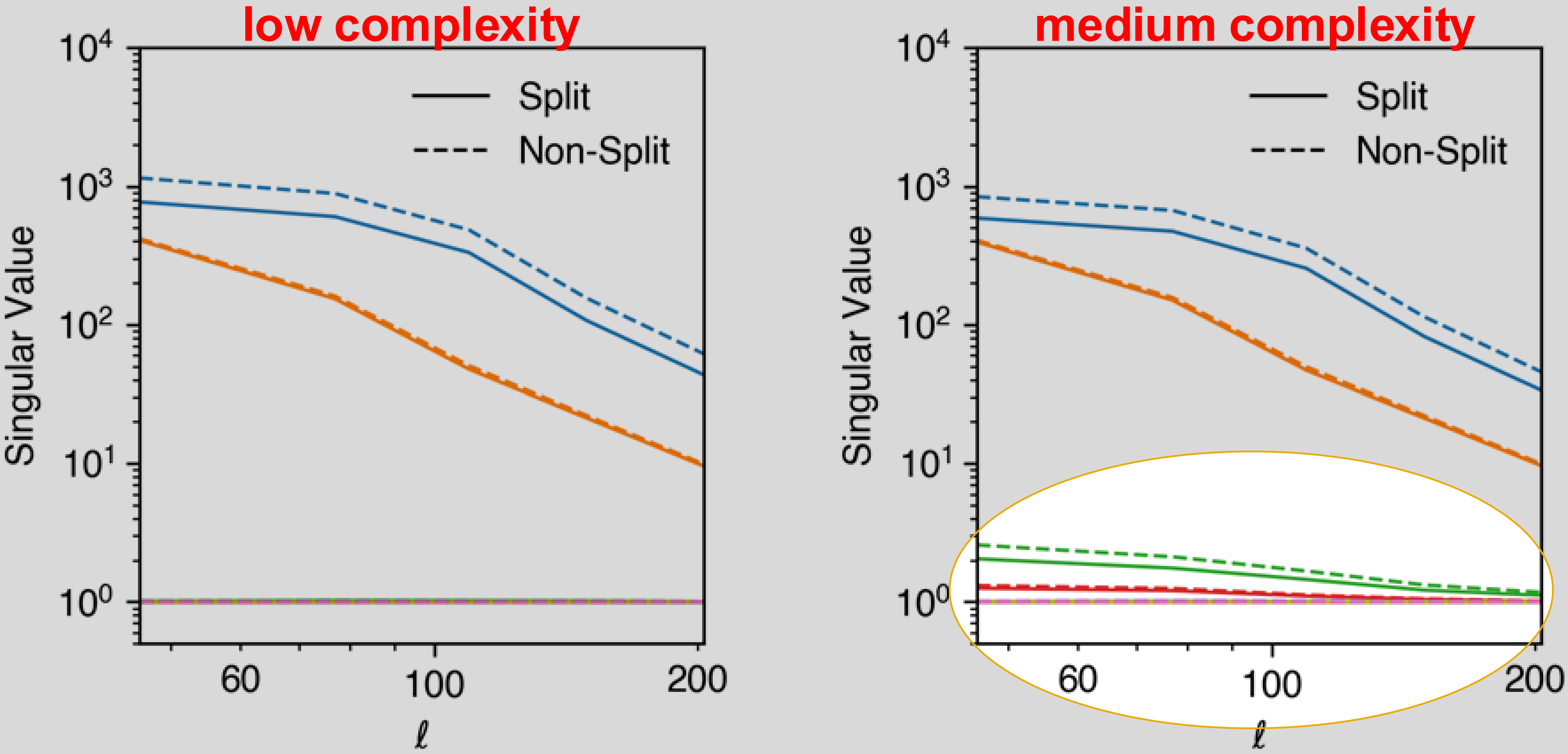
Foreground Bias!



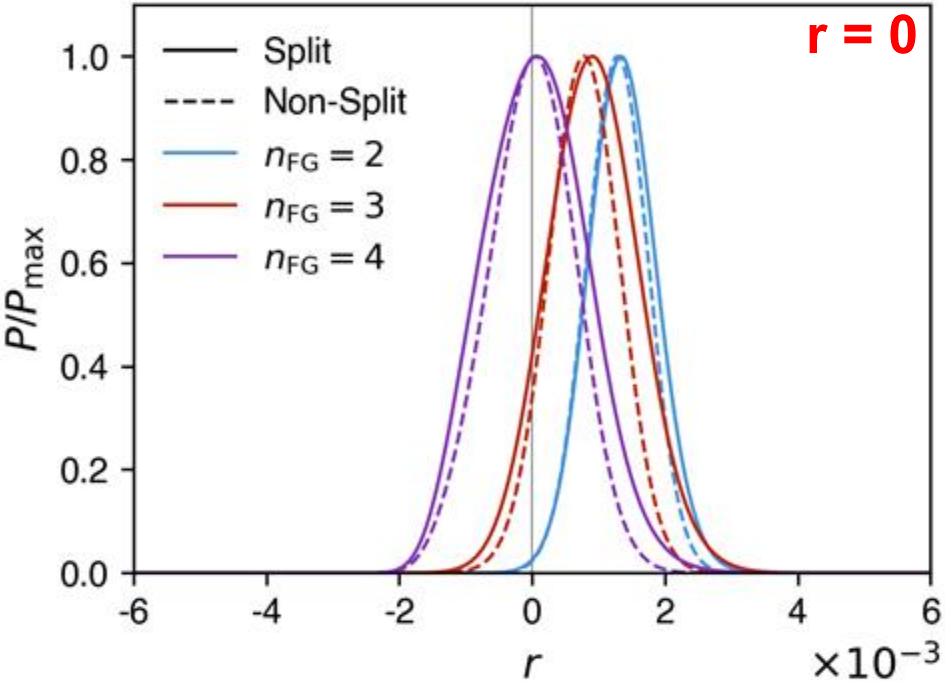
Split: 1.3 ± 0.4   4.3 ± 0.6

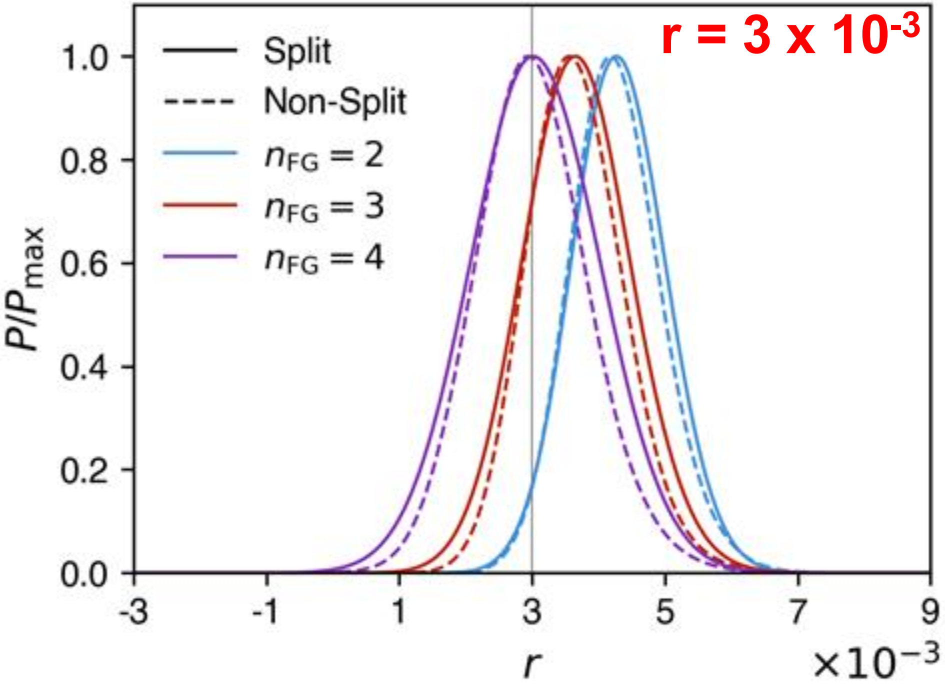
Non-Split: 1.3 ± 0.3   4.2 ± 0.6

Noise Whitened Foreground SVD - Low & Medium Complexity



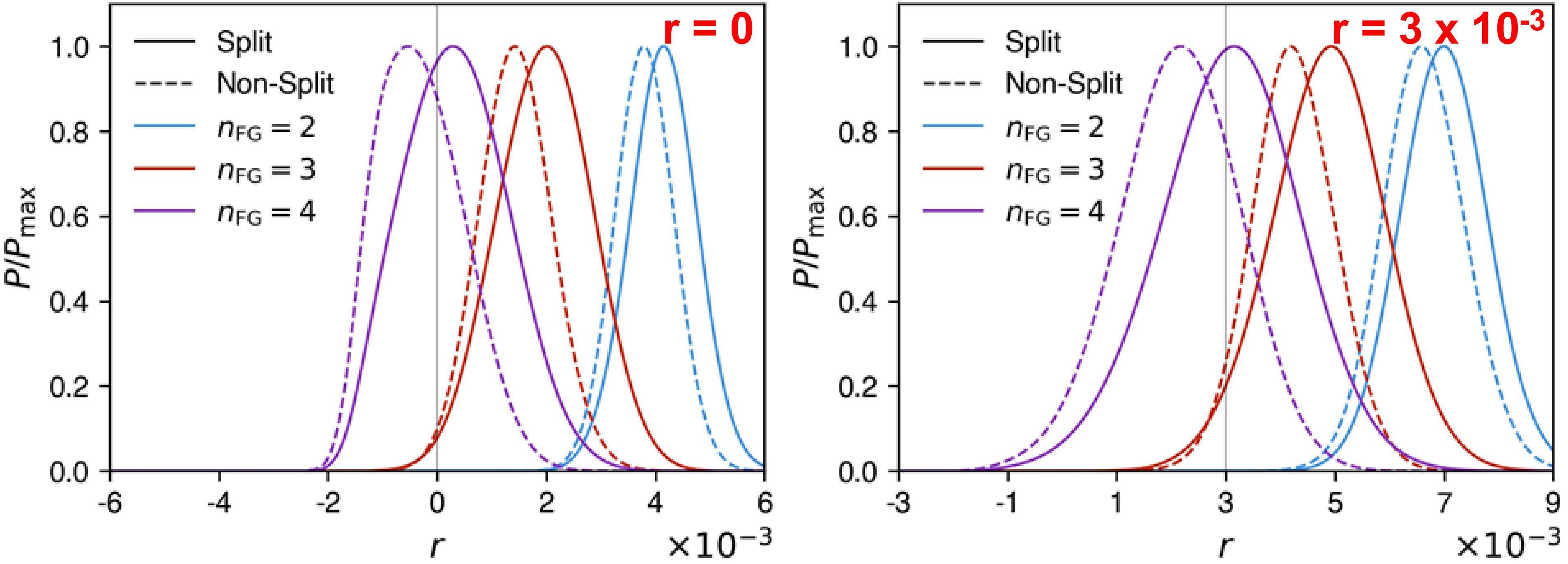
r-Posterior - Medium Complexity





	$r = 0$, $n_{FG} = 2$	$r = 0$, $n_{FG} = 3$	$r = 0$, $n_{FG} = 4$	$r = 3 \times 10^{-3}$, $n_{FG} = 2$	$r = 3 \times 10^{-3}$, $n_{FG} = 3$	$r = 3 \times 10^{-3}$, $n_{FG} = 4$
Split:	1.3 ± 0.4	0.9 ± 0.6	0.1 ± 0.7	4.3 ± 0.6	3.7 ± 0.8	3.0 ± 0.9
Non-Split:	1.3 ± 0.3	0.8 ± 0.4	0.0 ± 0.6	4.2 ± 0.6	3.6 ± 0.7	2.9 ± 0.8

r-Posterior - High Complexity



Split:	4.1 ± 0.5	2.0 ± 0.8	0.3 ± 0.9	7.0 ± 0.7	4.9 ± 1.0	3.1 ± 1.2
Non-Split:	3.8 ± 0.4	1.4 ± 0.6	- 0.5 ± 0.7	6.6 ± 0.7	4.2 ± 0.7	2.2 ± 1.1

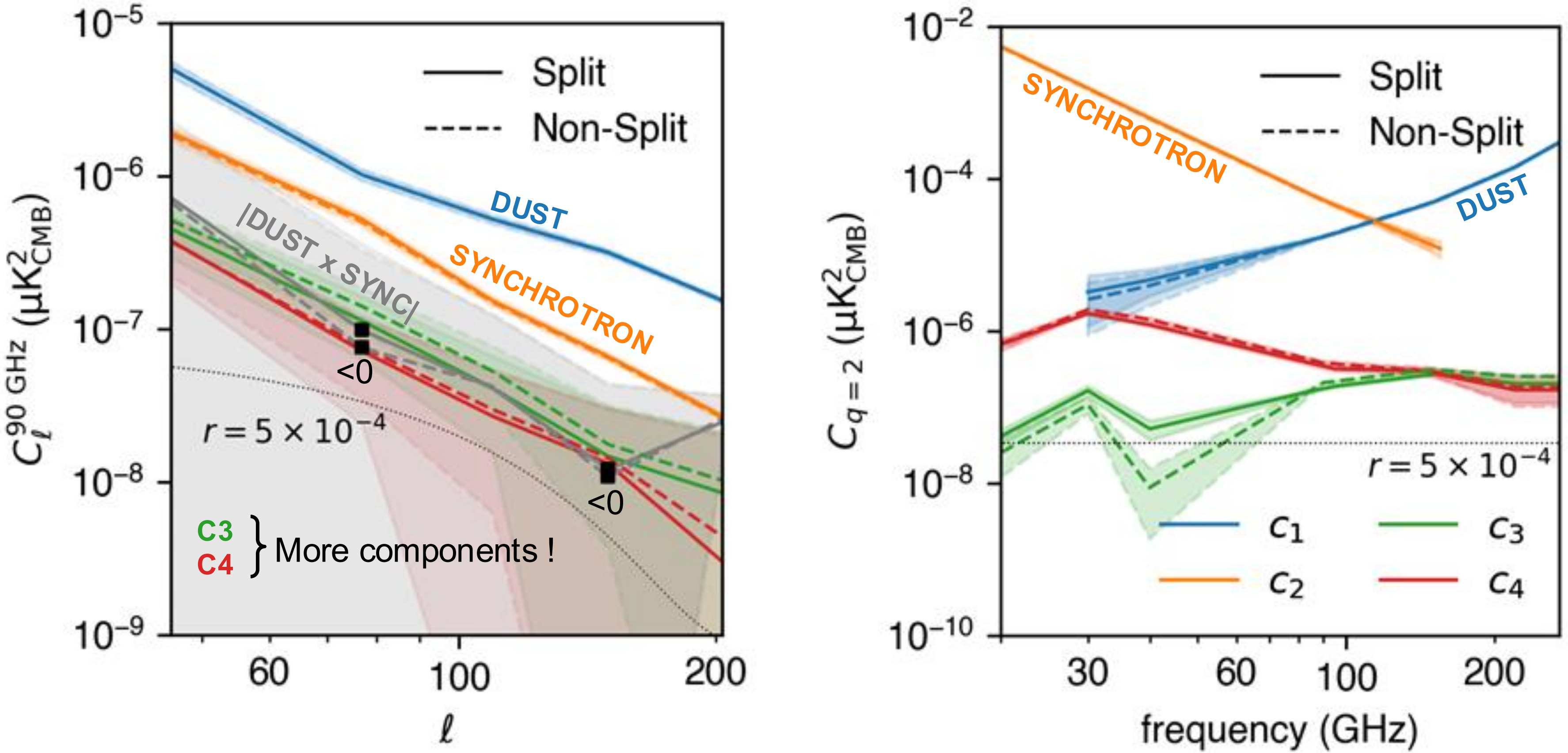
Chi Squared per DOF as Goodness of Fit

	n_{FG}	Low Complexity $r_{MAP} \pm \sigma_r$ (10^{-3})	Low Complexity χ^2/n_{dof}	Medium Complexity $r_{MAP} \pm \sigma_r$ (10^{-3})	Medium Complexity χ^2/n_{dof}	High Complexity $r_{MAP} \pm \sigma_r$ (10^{-3})	High Complexity χ^2/n_{dof}
Split	2	0.0 ± 0.3	1.27	1.3 ± 0.4	1.16	4.1 ± 0.5	1.37
	3	-	-	0.9 ± 0.6	1.17	2.0 ± 0.8	1.27
	4	-	-	0.1 ± 0.7	1.16	0.3 ± 0.9	1.19
Non-Split	2	0.0 ± 0.3	1.19	1.3 ± 0.3	1.24	3.8 ± 0.4	1.73
	3	-	-	0.8 ± 0.4	1.16	1.4 ± 0.6	1.33
	4	-	-	0.0 ± 0.6	1.10	-0.5 ± 0.7	1.26

$$\chi^2/n_{\rm dof} = \frac{-2\ln\mathcal{L}(\theta_{\rm MAP})}{\frac{1}{2}n_f(n_f+1)Q - n_{\rm params}}$$

Chi Squared per DOF is ***not*** a good indicator of r-bias

Best Fit Power Spectra & Mixing Vectors - High Complexity



Summary on SMICA

- SMICA offers unbiased *r*-measurement (within 1σ) even for high complexity foregrounds
- The method does not rely on a parametric model of foregrounds (no assumption on emission laws, harmonic spectrum, correlation between synchrotron and dust, frequency decorrelation)

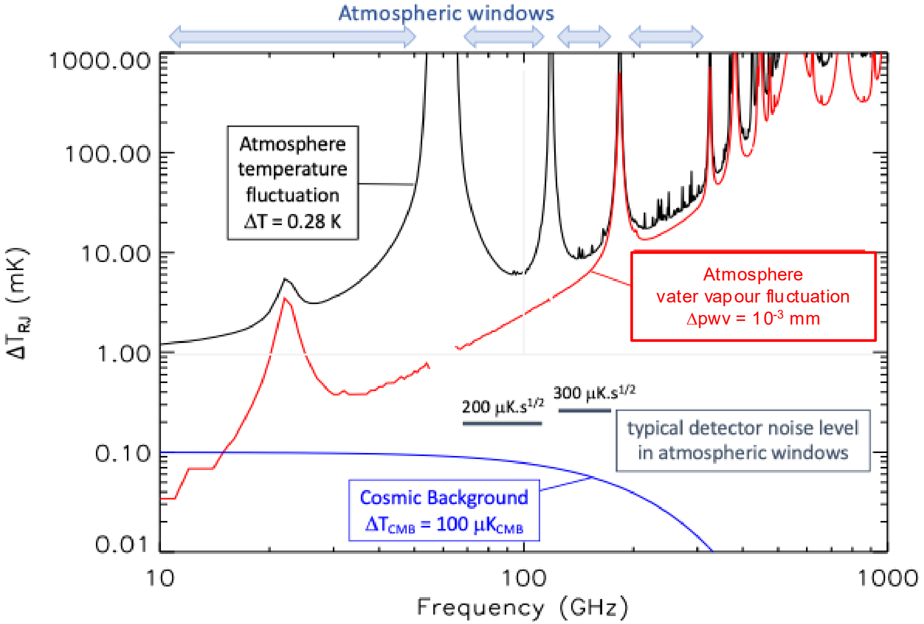
The data "speaks for itself" !!

- Non-split bands offer much needed sensitivity — better results than with split bands
- ⚠ BEWARE: Chi Squared per DOF is not a good indicator of *r*-bias

What NEXT ??

The next component separation challenge

- **ATMOSPHERE** — One of the biggest foregrounds ! (Time domain...)
- Can we reduce its impact by a combination of map-making and component separation methods ?



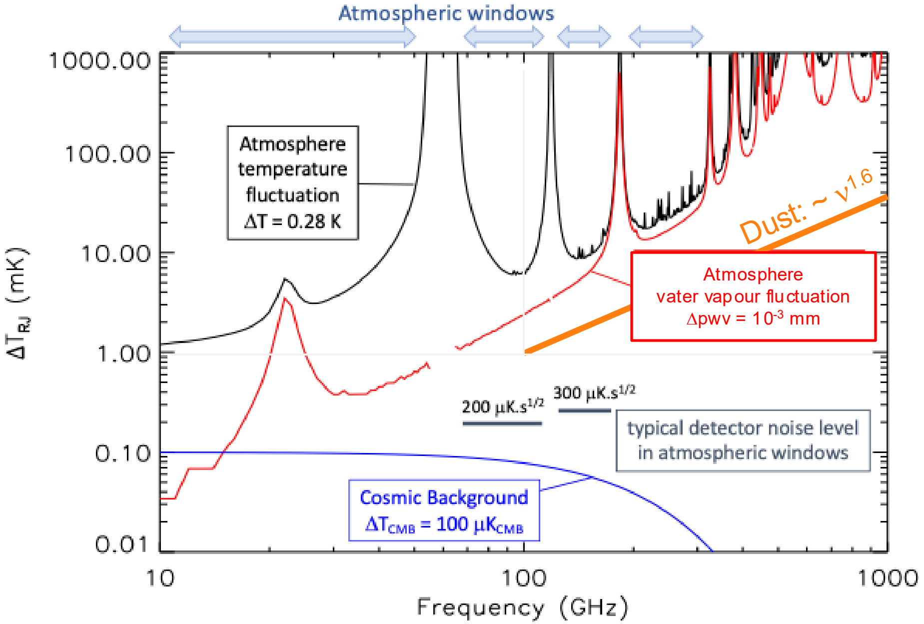
We can think of using extra frequency channels, but the atmosphere is multi-dimensional...

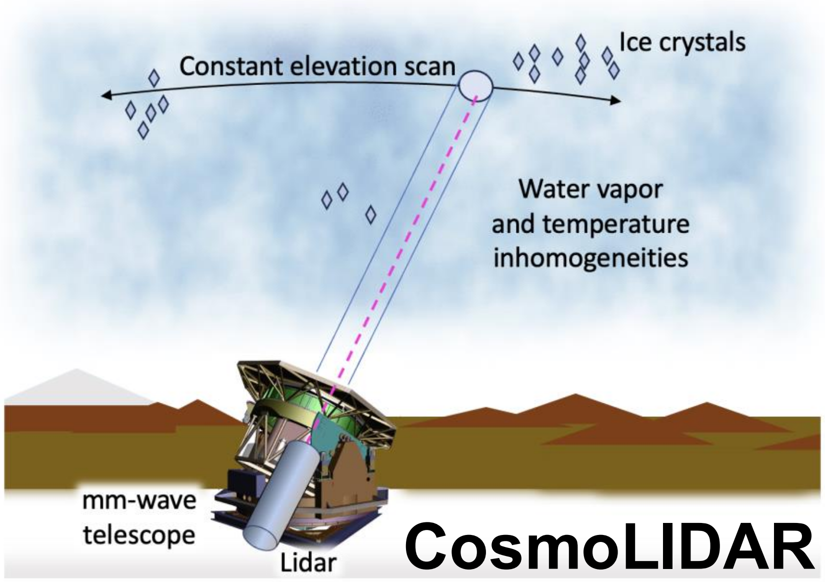
Already struggling with astrophysical components, we need **more channels** !

Why not **dedicated channels** ??

The next component separation challenge

- **ATMOSPHERE** — One of the biggest foregrounds ! (Time domain...)
- Can we reduce its impact by a combination of map-making and component separation methods ?
- With the help of ***DEDICATED observation channels*** ?





The next component separation challenge

Shamik Ghosh
and CPB / LBNL-C3 / APC / CEA groups

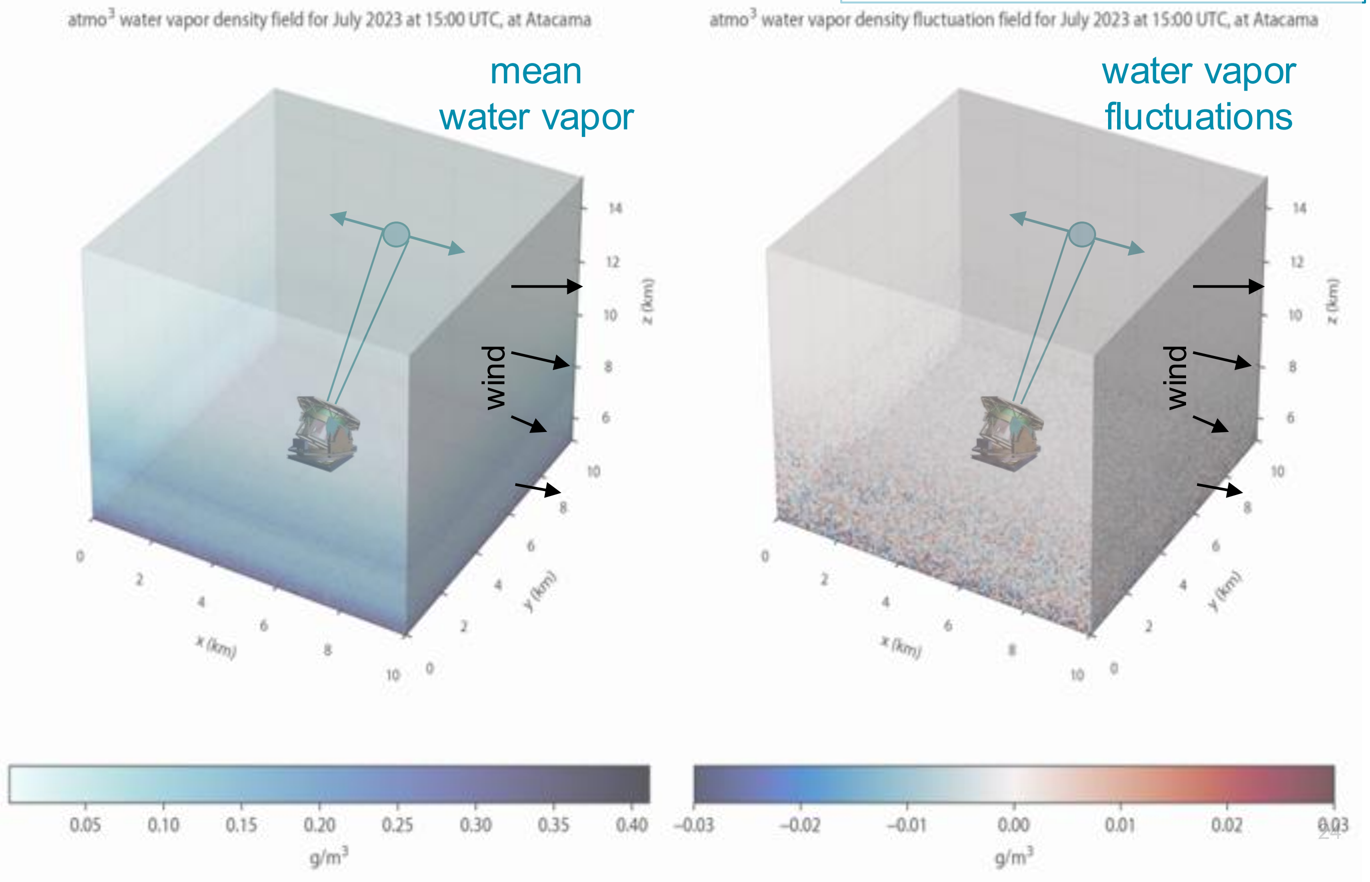
- Simulations ongoing:

  Atmo^3

- Forecasts next ...

- Could we field a LIDAR in Atacama and make some tests jointly with SO data ?

- Could be quite relevant for 350 GHz *KAIROS*



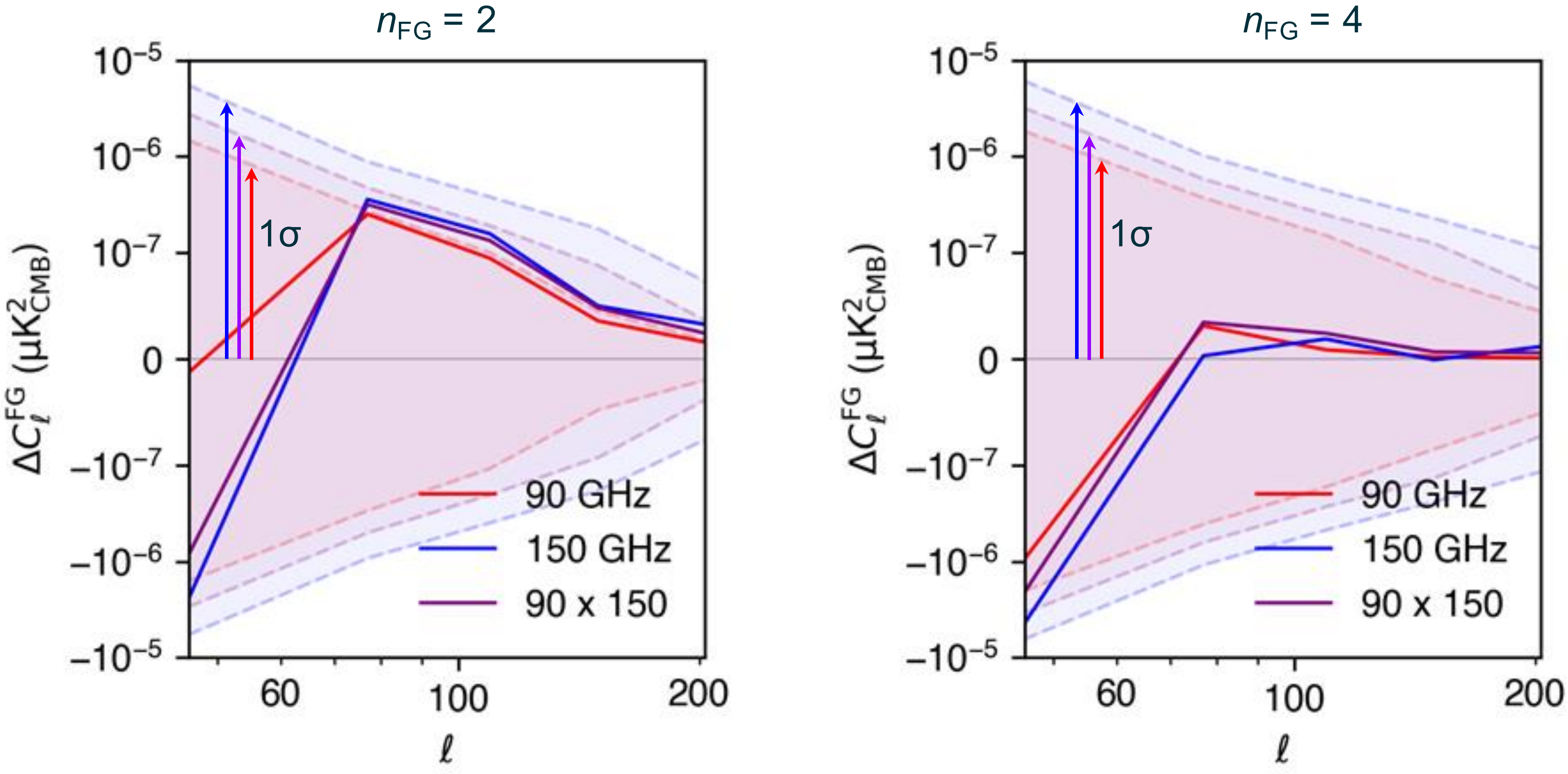
BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY
Office of Science

Thank You

CNRS • Berkeley
Centre Pierre Binétruy

Foreground Residuals - Non-Split High Complexity



Measurement Table (r = 0)

	n_{FG}	Low Complexity		Medium Complexity		High Complexity	
		$r_{\text{MAP}} \pm \sigma_r$ (10^{-3})	χ^2/n_{dof}	$r_{\text{MAP}} \pm \sigma_r$ (10^{-3})	χ^2/n_{dof}	$r_{\text{MAP}} \pm \sigma_r$ (10^{-3})	χ^2/n_{dof}
Split	2	0.0 ± 0.3	1.27	1.3 ± 0.4	1.16	4.1 ± 0.5	1.37
	3	-	-	0.9 ± 0.6	1.17	2.0 ± 0.8	1.27
	4	-	-	0.1 ± 0.7	1.16	0.3 ± 0.9	1.19
Non-Split	2	0.0 ± 0.3	1.19	1.3 ± 0.3	1.24	3.8 ± 0.4	1.73
	3	-	-	0.8 ± 0.4	1.16	1.4 ± 0.6	1.33
	4	-	-	0.0 ± 0.6	1.10	-0.5 ± 0.7	1.26

Measurement Table (r = 0.003)

	n_{FG}	Low Complexity		Medium Complexity		High Complexity	
		$r_{MAP} \pm \sigma_r$ (10^{-3})	χ^2/n_{dof}	$r_{MAP} \pm \sigma_r$ (10^{-3})	χ^2/n_{dof}	$r_{MAP} \pm \sigma_r$ (10^{-3})	χ^2/n_{dof}
Split	2	2.9 ± 0.5	1.33	4.3 ± 0.6	1.16	7.0 ± 0.7	1.28
	3	-	-	3.7 ± 0.8	1.20	4.9 ± 1.0	1.26
	4	-	-	3.0 ± 0.9	1.17	3.1 ± 1.2	1.21
Non-Split	2	2.9 ± 0.5	1.30	4.2 ± 0.6	1.32	6.6 ± 0.7	1.66
	3	-	-	3.6 ± 0.7	1.26	4.2 ± 0.7	1.34
	4	-	-	2.9 ± 0.8	1.21	2.2 ± 1.1	1.26