A Novel Approach to Optimize Clustering of Parametric Map-Based Component Separation for Upcoming CMB Polarization Satellites

Wassim Kabalan[1], Wuhyun Sohn[1], Benjamin Beringue[1], Artem Basyrov[1], Pierre Chanial[1], Alexandre Boucaud[1], Josquin Errard[1]

[1]Université Paris Cité, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France

Tensor-to-Scalar Ratio & Foreground Removal

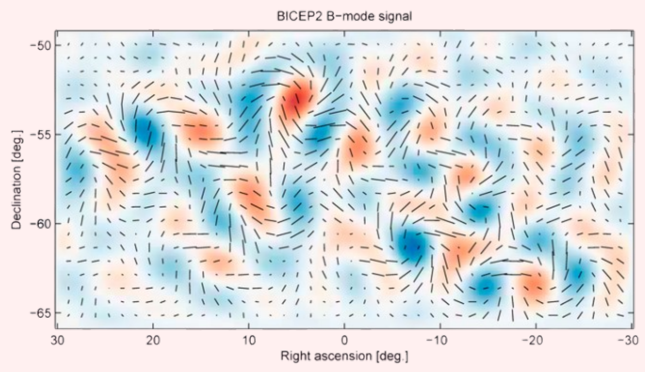


Figure 1: CMB B-mode polarization signal from primordial gravitational waves

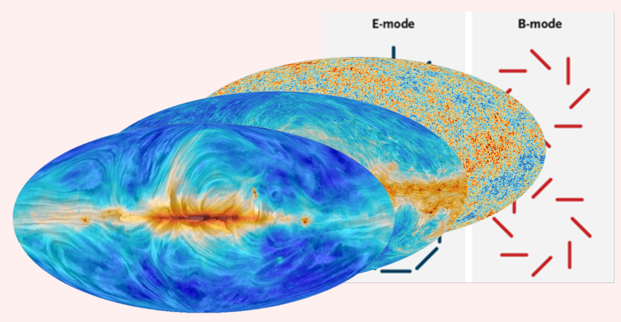


Figure 2: Galactic foreground contamination

The Challenge:

- Measuring tensor-to-scalar ratio r requires detecting faint B-mode polarization
- Galactic foregrounds dominate CMB signal by orders of magnitude
- Accurate foreground removal is critical for $r < 0.001$ detection

Parametric Component Separation

Data Model:

$$\mathbf{d} = \mathbf{A}(\boldsymbol{\beta})\,\mathbf{s} + \mathbf{n} \tag{1}$$

where $\mathbf{d}$ is observed data, $\mathbf{s}$ are sky components, $\mathbf{A}(\boldsymbol{\beta})$ encodes spectral dependencies, and $\mathbf{n}$ is noise.

Generalized Least Squares Solution:

$$\hat{\mathbf{s}} = \left(\mathbf{A}^\top \mathbf{N}^{-1} \mathbf{A}\right)^{-1} \mathbf{A}^\top \mathbf{N}^{-1} \mathbf{d} \tag{2}$$

Spectral Likelihood:

$$\ln \mathcal{L}_{\rm spec}(\boldsymbol{\beta}) \propto (\mathbf{A}^\top \mathbf{N}^{-1} \mathbf{d})^\top (\mathbf{A}^\top \mathbf{N}^{-1} \mathbf{A})^{-1} (\mathbf{A}^\top \mathbf{N}^{-1} \mathbf{d}) \tag{3}$$

Key Innovation: Data-driven optimization of spatial clustering rather than fixed patching schemes.

FURAX Implementation:

```
cmb = CMBOperator(nu, ...)
dust = DustOperator(nu, temp=params["T_d"],
    beta=params["beta_d"], ...)
synchrotron = SynchrotronOperator(nu,
    beta=params["beta_pl"], ...)
A = MixingMatrixOperator(cmb, dust, synchrotron)
AND = (A.T @ N.I)(d)
s = (A.T @ N.I @ A).I(AND)
likelihood = AND @ s
```

Spatial Variability

Foreground Spectral Parameters Vary Across Sky:

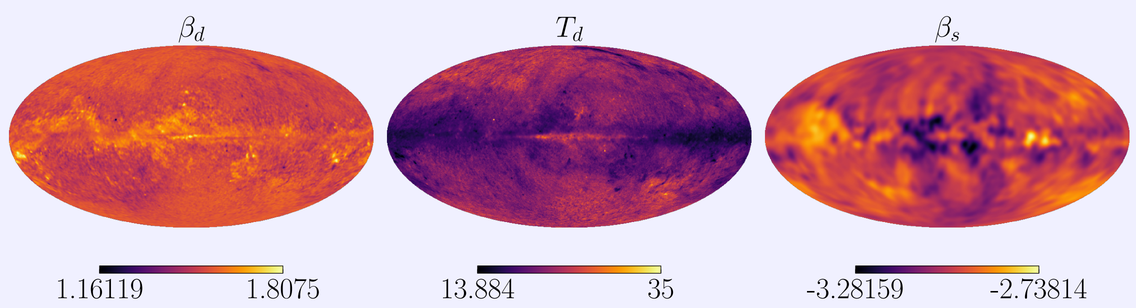


Figure 3: Modified blackbody spectral index map showing spatial variability in dust emission properties across the sky

Clustering Approach:

- Spherical K-means clustering groups pixels with similar spectral properties
- Balances statistical uncertainty with modeling flexibility

Optimized K-means Parameters:

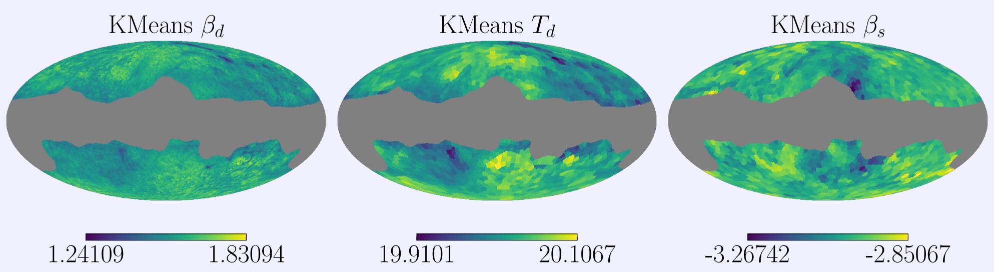
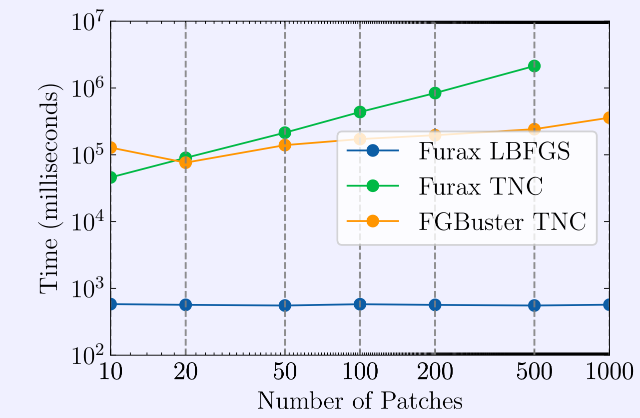


Figure 4: Optimized spectral parameter distributions from K-means clustering showing data-driven adaptation to foreground complexity

The FURAX Framework

- **JAX-Native**: Differentiable & GPU accelerated
- **Modular Design**: Composable algebraic operators
- **Memory Efficient**: Matrix-free linear operators
- **Scalable**: Multi-GPU execution

Grid Search: Configuration space $\mathcal{G} = \{K_{\beta_d}\} \times \{K_{T_d}\} \times \{K_{\beta_s}\}$ evaluated across 1.92M component separation runs.



.

Figure 5: Runtime comparison of FURAX Component Separation

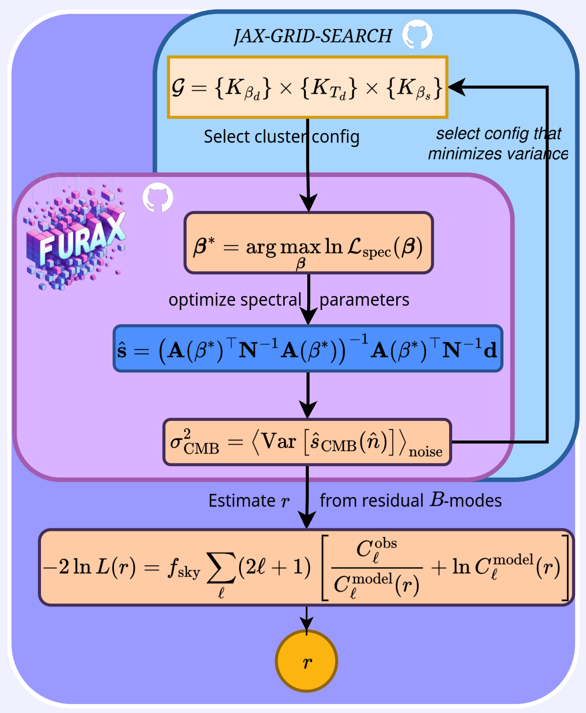


Figure 6: FURAX pipeline: Grid search over clustering configurations, spectral parameter optimization, and CMB variance minimization for model selection

Likelihood and Variance Analysis

Selection Criterion - CMB Variance Minimization:

$$\sigma^2_{\rm CMB} = \left\langle {\rm Var}_i \left[\hat{s}^{(i)}_{\rm CMB}\right] \right\rangle_{\rm pixels} \tag{4}$$

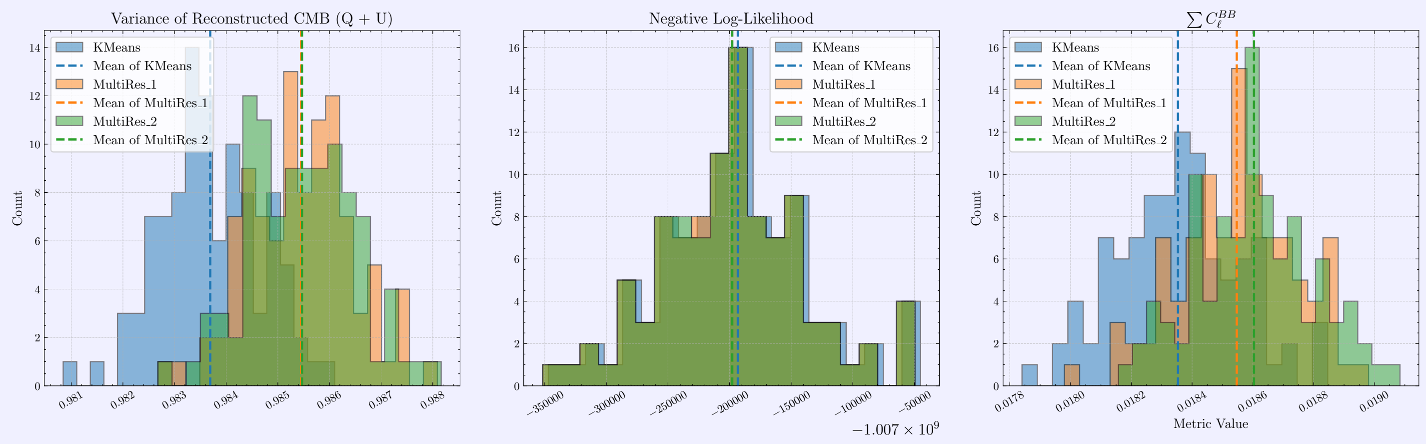


Figure 7: Distribution of CMB variance, spectral likelihood, and B-mode power for different spatial modeling approaches. K-means clustering achieves optimal balance.

Key Insight: Variance minimization acts as proxy for residual foreground contamination, leading to more robust cosmological constraints.

Results

Tensor-to-Scalar Ratio Constraints:

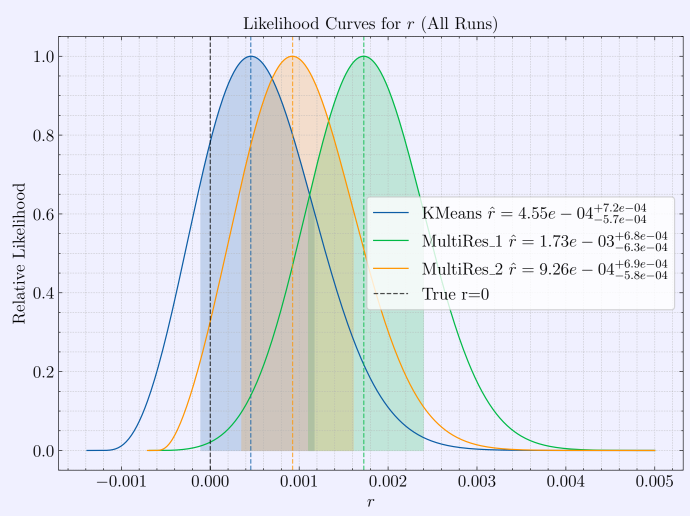


Figure 8: r likelihood distributions: K-means clustering (blue) yields $\hat{r} = 4.55 \times 10^{-4}$ with lowest bias and tightest constraints compared to multi-resolution approaches

Residual B-mode Spectra:

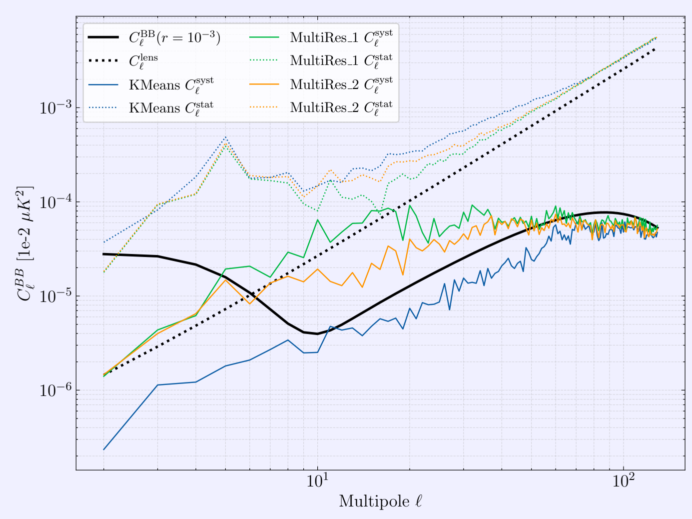


Figure 9: Residual B-mode power spectra: K-means clustering (blue) achieves significantly lower systematic residuals compared to multi-resolution approaches, falling below target sensitivity levels

References & Acknowledgements

Key References:

- FURAX framework & component separation methodology: Kabalan et al. (2024), arXiv:2024.xxxxx
- LiteBIRD collaboration forecasts: LiteBIRD Collaboration (2022), PTEP 2023
- JAX ecosystem for scientific computing: Bradbury et al. (2018), JAX library

Acknowledgements: This work was supported by the SciPol project (ERC Grant No. 101044073, PI: Josquin Errard) and performed using HPC resources at IDRIS Jean Zay supercomputer. Full code available at: `https://github.com/ASKabalan/furax-compsep-paper`