Accounting for the beams in the parametric component separation

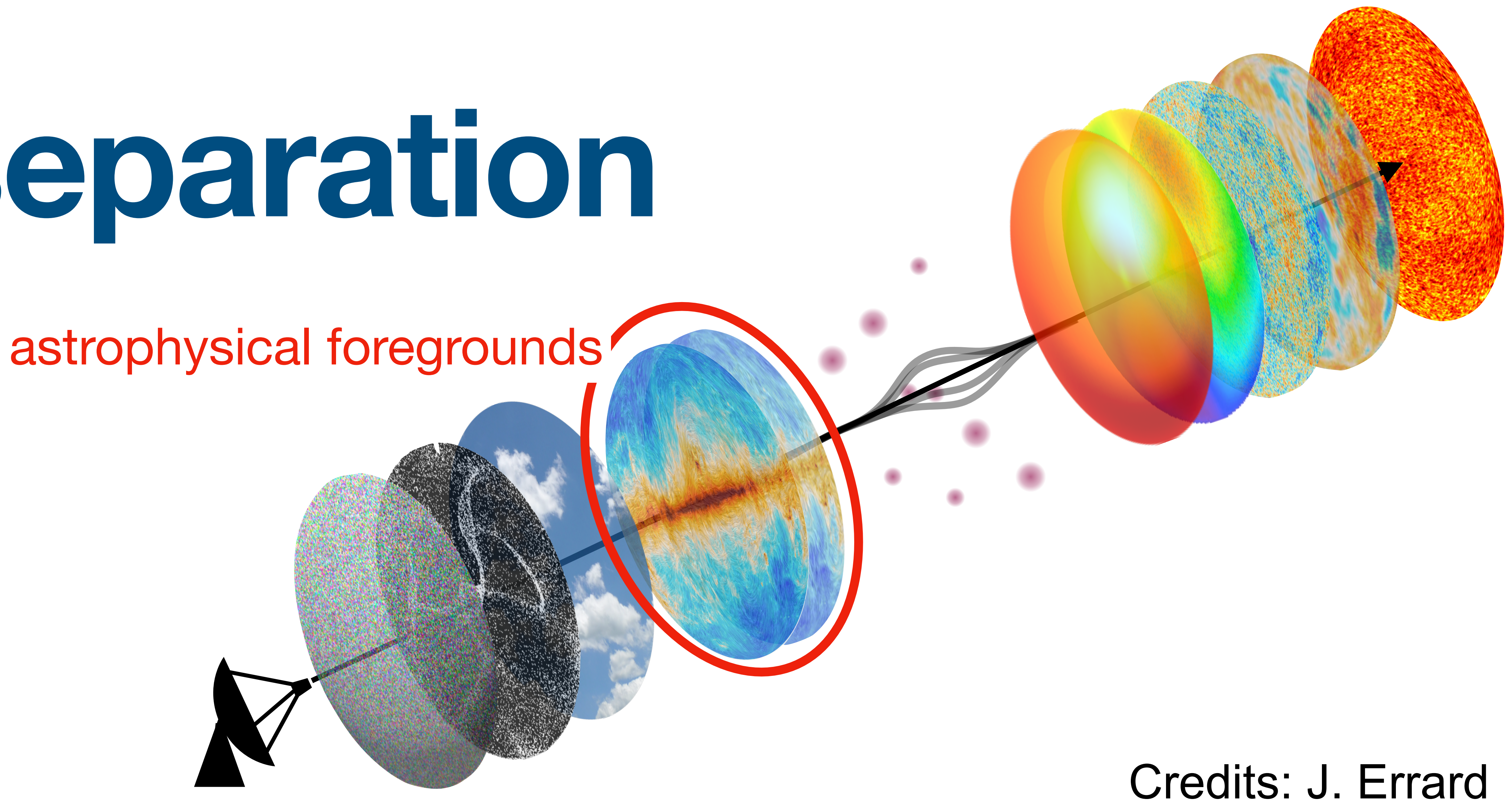
CMB-France 04/12/2023

Arianna Rizzieri, APC (Paris)
rizzieri@apc.in2p3.fr

PhD student with **Radek Stompor** and **Josquin Errard**

Component separation

The problem



Credits: J. Errard

The model

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

where: $\mathbf{d}$ = data vector of the measured signal for all the n_f frequencies and n_s Stokes parameters

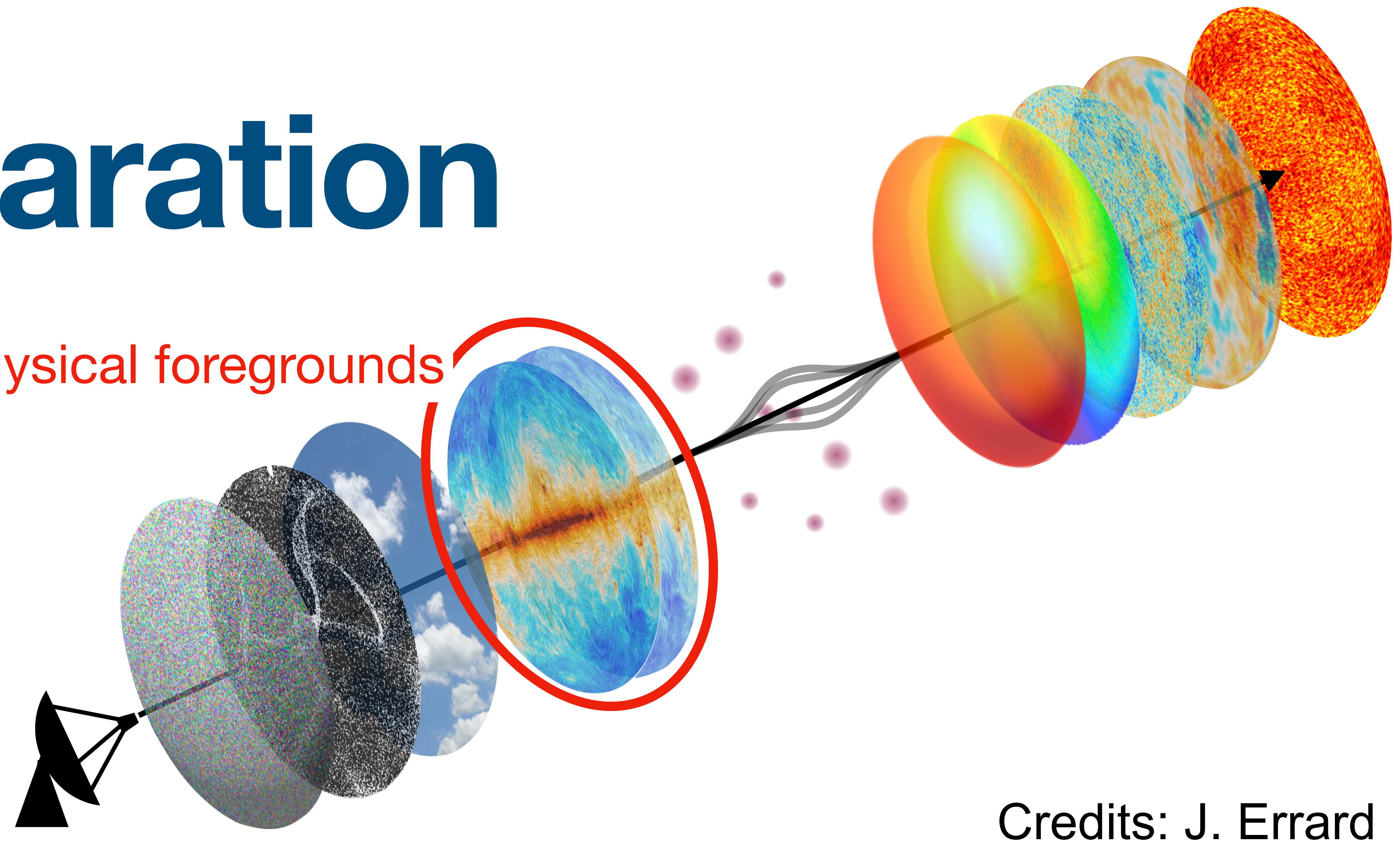
$\mathbf{A}$ = *component Mixing Matrix* unknown

$\mathbf{s}$ = true value signals for each component unknown

$\mathbf{n}$ = instrumental noise (assumed Gaussian distributed)

Component separation

The problem



Credits: J. Errard

The model

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

where: $\mathbf{d}$ = data vector of the measured signal for all the n_f frequencies and n_s Stokes parameters

Parametric component separation

$\mathbf{A}(\beta_i)$ = *component Mixing Matrix* **parametrised by a set of unknown parameters** β_i

$\mathbf{s}$ = true value signals for each component **unknown**

$\mathbf{n}$ = instrumental noise (assumed Gaussian distributed)

Parametric maximum likelihood based component separation

The solution - maximum likelihood principle

Full data likelihood:

$$-2\ln \mathscr{L}_{data}(\mathbf{s}, \beta_i) = \text{const} + (\mathbf{d} - \mathbf{A}\mathbf{s})^T \mathbf{N}^{-1} (\mathbf{d} - \mathbf{A}\mathbf{s})$$

Two possibilities:

Eriksen et al. 2006

- Characterising $\mathscr{L}_{data}$ numerically: mapping the likelihood by MCMC sampling De la Hoz et al. 2020
  (sampling at the same time signal and spectral parameters)
- Maximising $\mathscr{L}_{data}$:
  ➡ 2 step approach: Stompor et al. 2009

Step 1 recover the spectral parameter estimates by maximising

$$-2\ln \mathscr{L}_{spec}(\beta_i) = \text{const} - (\mathbf{A}^T\mathbf{N}^{-1}\mathbf{d})^T(\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{N}^{-1}\mathbf{d})$$

Step 2 recover the sky components

$$\mathbf{s} = (\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{N}^{-1}\mathbf{d}$$

Beams in component separation

The problem: ***Why do we need to account for the beams in the comp sep?***

$\mathbf{d} = \mathbf{A}\mathbf{s} + \mathbf{n}$ where $\mathbf{d} = [d_{\nu_0 \text{ fwhm}_0}, \ldots, d_{\nu_n \text{ fwhm}_n}]$

(different frequency channels have different beams)

Assumptions:

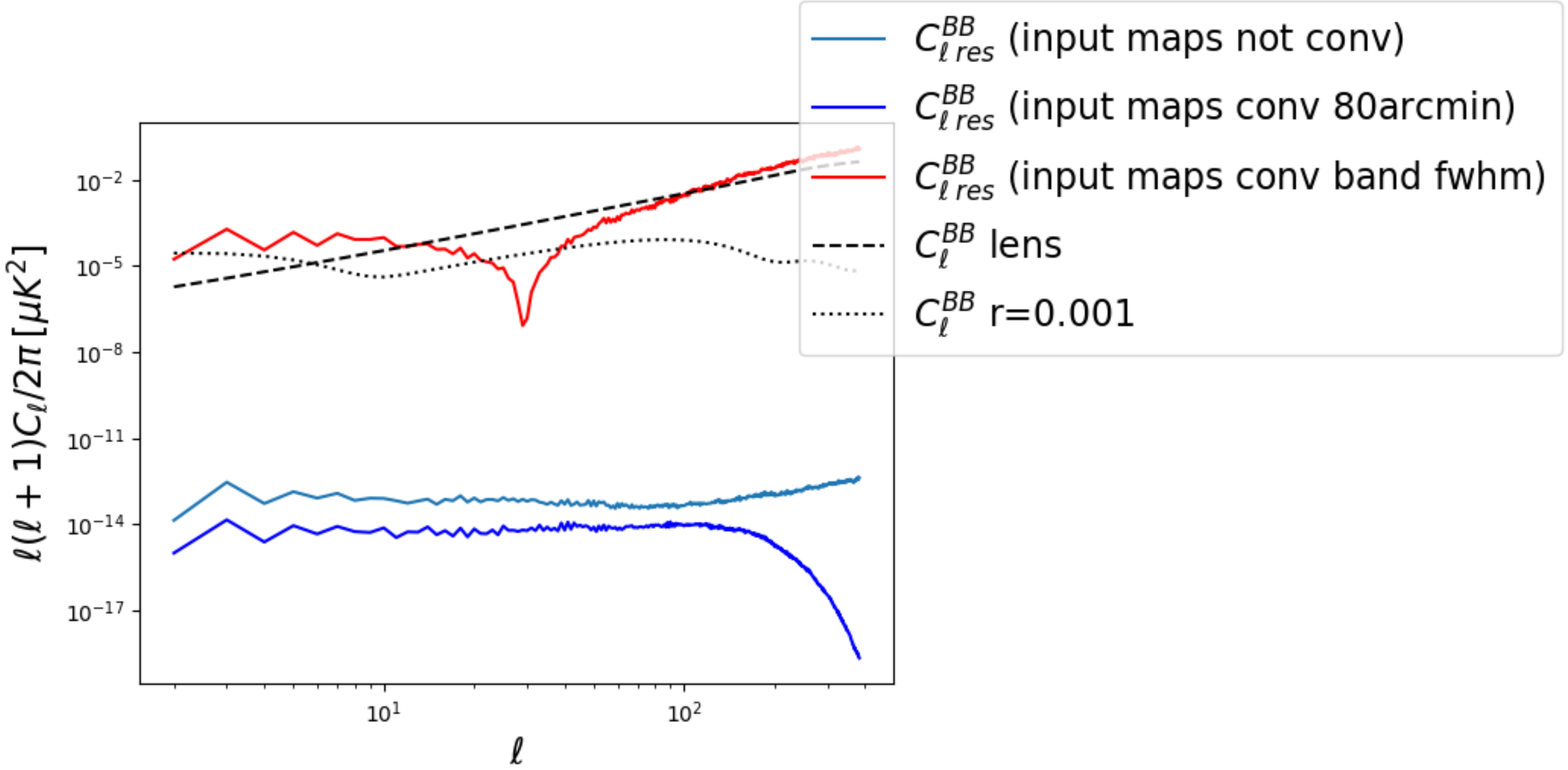
- only impact of the beam on the **component separation**
- only considered beam **main lobe**

$$-2 \ln \mathscr{L}_{spec}(\beta) = \text{const} - (\mathbf{A}^T\mathbf{N}^{-1}\mathbf{d})^T(\mathbf{A}^T\mathbf{N}^{-1}\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{N}^{-1}\mathbf{d})$$

the map-based spectral likelihood is biased

↓

the recovered CMB is biased (much higher CMB residuals)



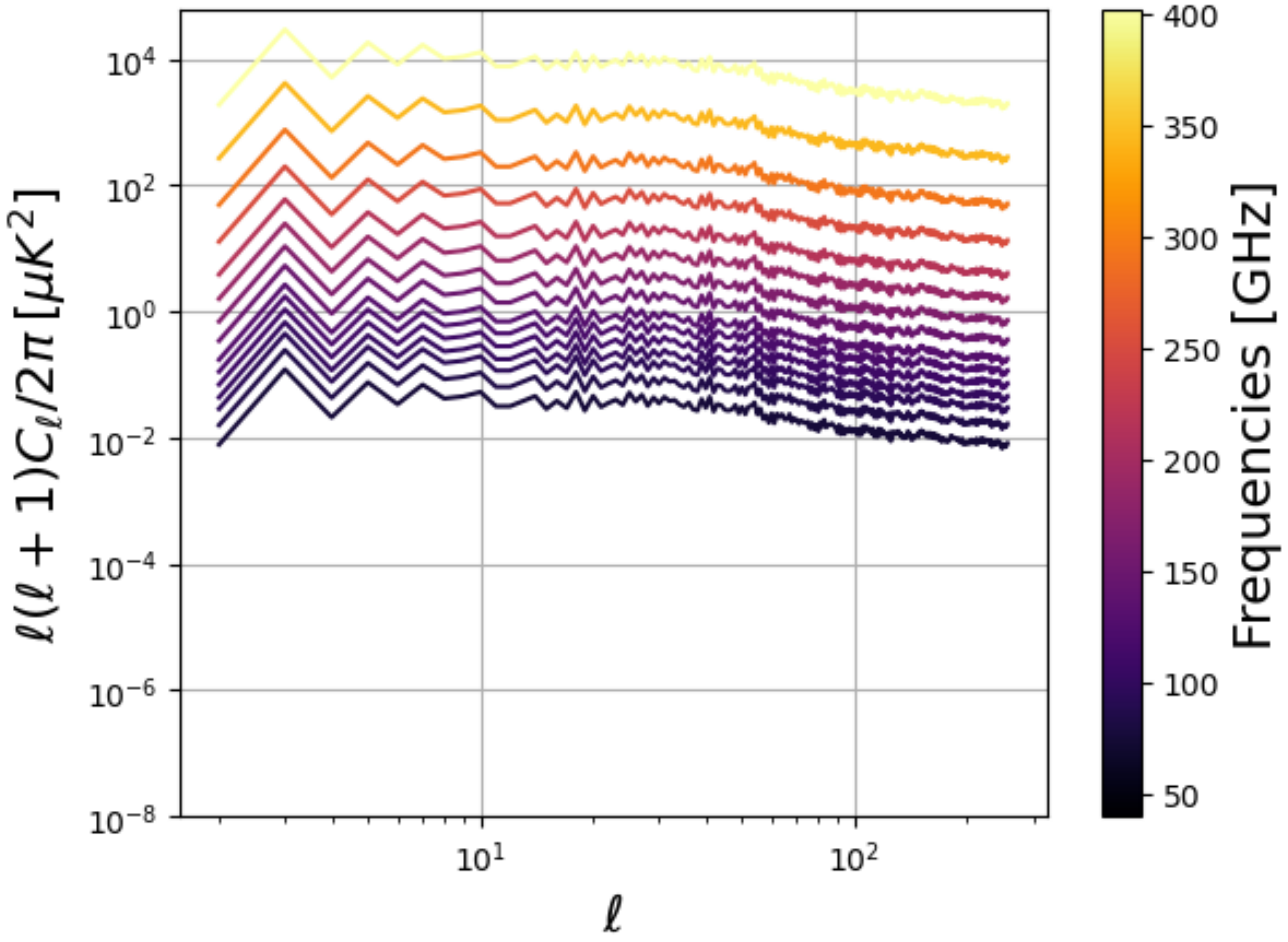
Beams in component separation

Extended data model

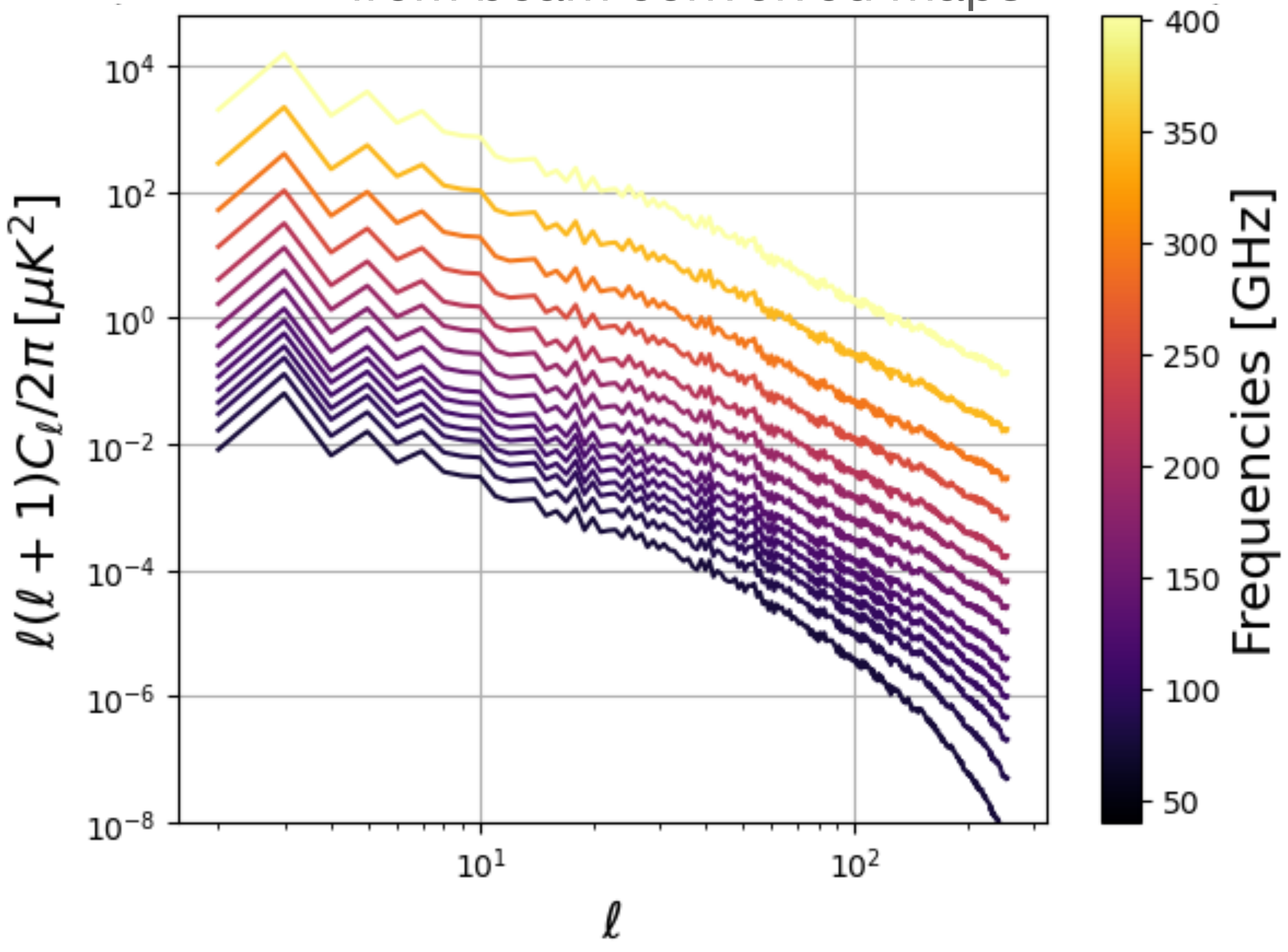
$$\mathbf{d} = \mathbf{B}\mathbf{A}\mathbf{s} + \mathbf{n}$$

where $\mathbf{B}$ true beams of the input frequency maps

Thermal dust emission power spectrum



Thermal dust emission power spectrum
from beam convolved maps



Beams in component separation

Extended data model

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

where $\mathbf{B}$ true beams of the input frequency maps

Can we perform the commutation of $\mathbf{B}$ and $\mathbf{A}$?

- ✓ if $\mathbf{B}$ is the same at each frequency band
- ✓ if there is no spatial variability in $\mathbf{A}$ (or spatial variability on scales bigger than the beam size)

Otherwise $\mathbf{B}$ and $\mathbf{A}$ do not commute, unless it is defined an effective Mixing Matrix $\tilde{\mathbf{A}}$ with more complex scaling laws depending on the beams: $\mathbf{B}\mathbf{A} = \tilde{\mathbf{A}}\hat{\mathbf{B}}$

Beams in parametric component separation

The problem: ***How to recover component maps from beam convolved frequency maps?***

Beams in parametric component separation

1st solution: *before* component separation

Beams in parametric component separation

1st solution: *before* component separation

Beams in parametric component separation

1st solution: *before* component separation

Modify input frequency maps and back to usual component separation:

Input frequency maps (each convolved at its own beam $\mathbf{B}_{\mathbf{true}}$):

$$\mathbf{d} = \mathbf{B}_{\mathbf{true}}(\mathbf{A}\mathbf{s}) + \mathbf{n}$$

Applying deconvolution and convolution on input frequency maps through:

$$\mathbf{B}_{\mathbf{eff}}\mathbf{d} = \mathbf{B}_{\mathbf{eff}}(\mathbf{B}_{\mathbf{true}}(\mathbf{A}\mathbf{s}) + \mathbf{n}) = \mathbf{B}_{\mathbf{final}}(\mathbf{A}\mathbf{s}) + \mathbf{B}_{\mathbf{eff}}\mathbf{n} = \mathbf{A}\hat{\mathbf{B}}_{\mathbf{final}}\mathbf{s} + \mathbf{B}_{\mathbf{eff}}\mathbf{n} = \mathbf{A}\mathbf{s}' + \mathbf{B}_{\mathbf{eff}}\mathbf{n}$$

($\mathbf{B}_{\mathbf{eff}} = \dfrac{\mathbf{B}_{\mathbf{final}}}{\mathbf{B}_{\mathbf{true}}}$ is the effective beam representing the deconvolution by $\mathbf{B}_{\mathbf{true}}$ and the convolution to a common beam $\mathbf{B}_{\mathbf{final}}$)

The usual component separation is performed on the $\mathbf{B}_{\mathbf{eff}}\mathbf{d}$ *frequency maps, to recover the component maps* $\mathbf{s}' = \hat{\mathbf{B}}_{\mathbf{final}}\mathbf{s}$ *beam-smoothed at the* final *resolution.*

Beams in parametric component separation

1st solution: *before* component separation

Modify input frequency maps and back to usual component separation:

Pros vs Cons

Pros

- No need to modify the component separation technique
- Fast

Cons

- $\mathbf{B}_{\mathbf{eff}}$ applied to $\mathbf{n}$, correlations introduced. The diagonal noise covariance matrix doesn't describe the data anymore. We can correct for it, but it remains an approximation.
- Common resolution has to be chosen $\geq$ than the worst input resolution
- Need to be able to smooth the maps in a reliable way (for example needed maps at sufficiently high resolution)

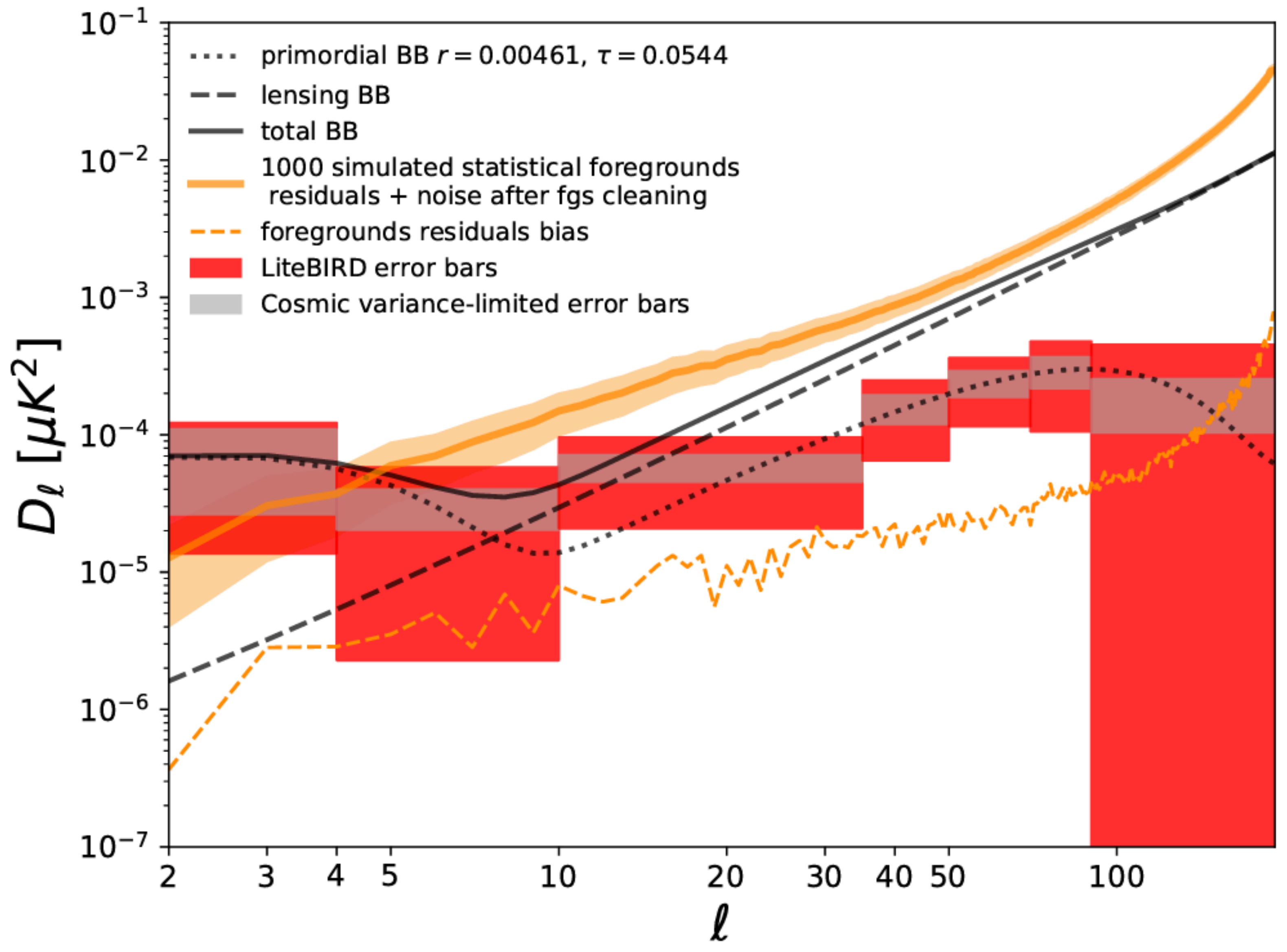
Beams in parametric component separation

1st solution: *before* component separation

Modify input frequency maps and back to usual component separation:

Example of this approach already used in the literature

Probing Cosmic Inflation with the LiteBIRD Cosmic Microwave Background Polarization Survey (LiteBIRD Collaboration, 2023)



The Simons Observatory: pipeline comparison and validation for large-scale B-modes (Wolz et al, 2023)
[fgbuster = pipeline C below]

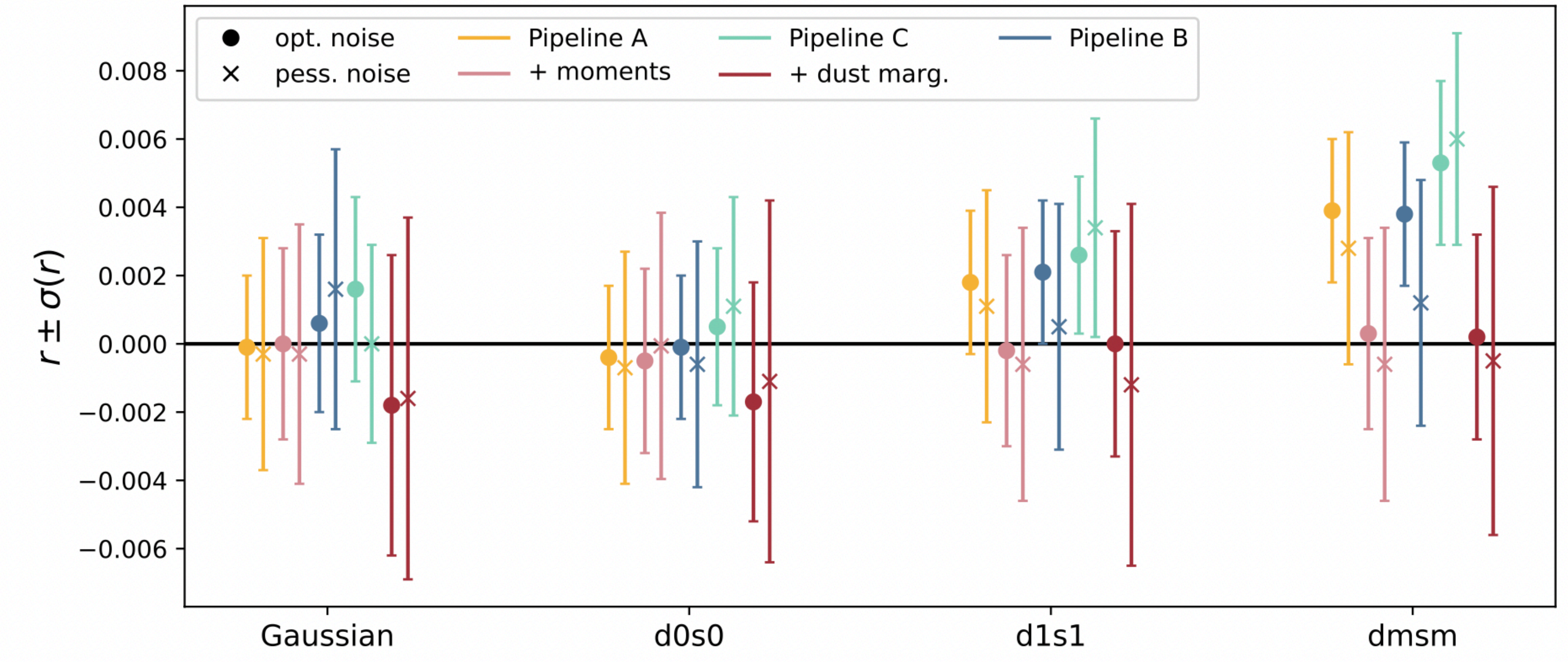


Fig. 7. Compilation of the mean r with its (16, 84)% credible interval as derived from 500 simulations, applying the three nominal pipelines (plus extensions) to four foreground scenarios of increasing complexity. We assume a fiducial cosmology with $r = 0$ and $A_{\rm lens} = 1$, inhomogeneous noise with goal sensitivity and optimistic $1/f$ noise component (dot markers), and inhomogeneous noise with baseline sensitivity and pessimistic $1/f$ noise component (cross markers). Note that the NILC results for Gaussian foregrounds are based on a smaller sky mask, see Appendix B.

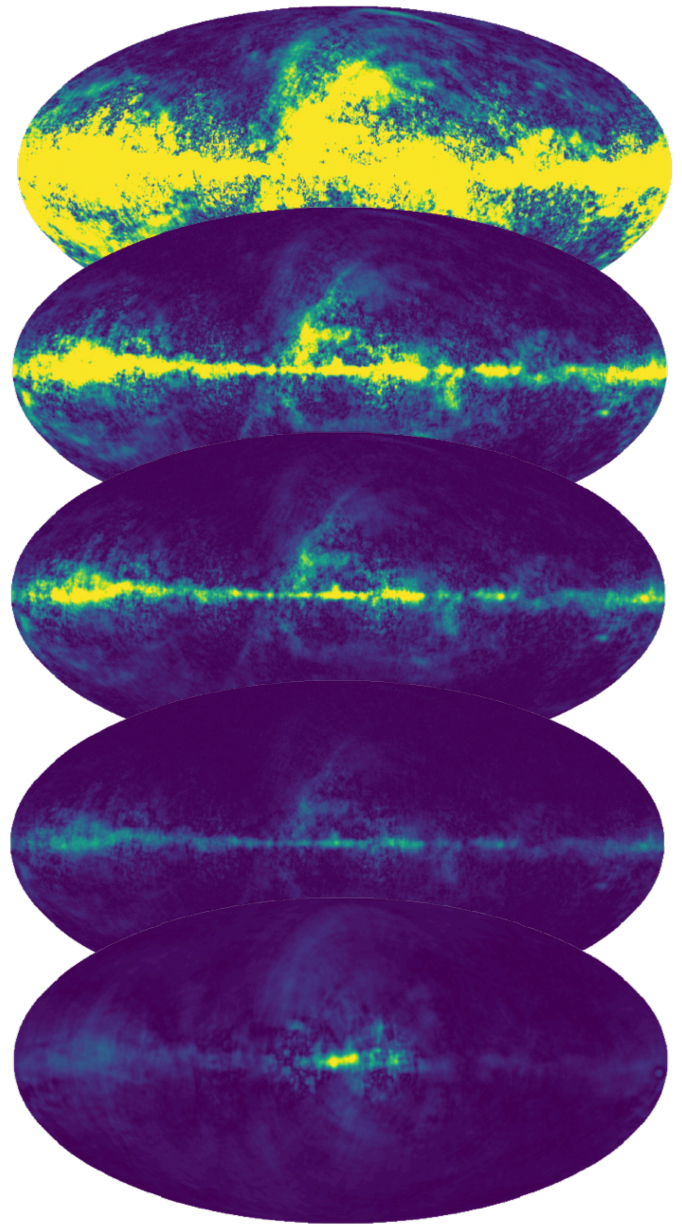
Beams in parametric component separation

2nd solution: *in the* component separation

Beam fwhm:

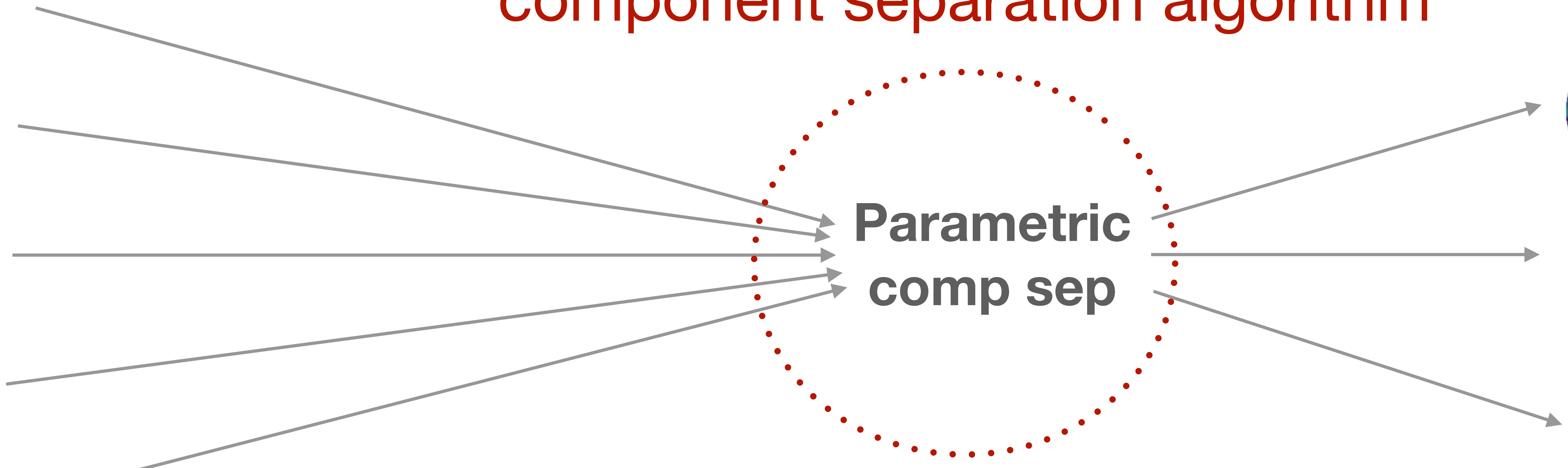
80 arcmin

20 arcmin



frequency maps
each smoothed at its beam value

2nd solution:
dealing with the beams ***in*** the
component separation algorithm

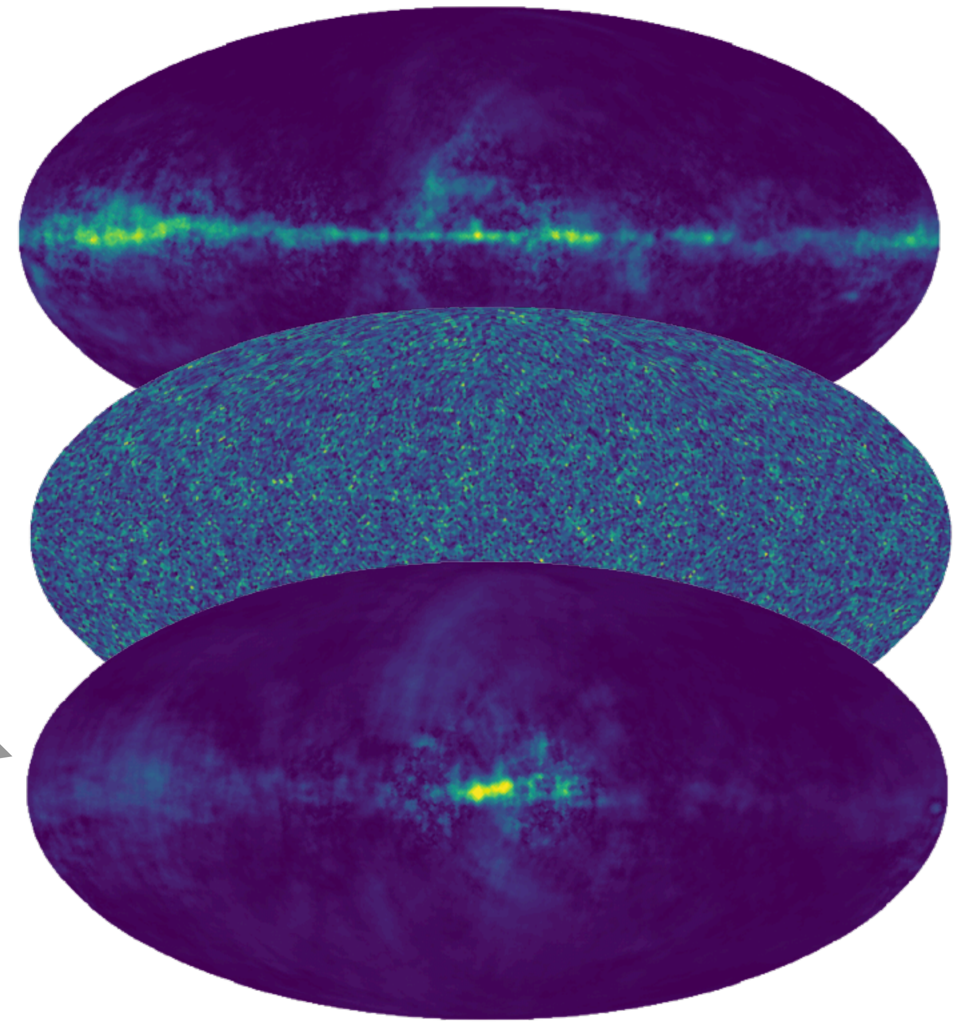


Beam fwhm:

?? arcmin

?? arcmin

?? arcmin



component maps
smoothed at a common beam

Beams in parametric component separation

2nd solution: *in the* component separation

domain	applying B	spatial variability of A	homogenous noise	ell correlated noise
harmonic	✓	X	X	✓
pixel	X	✓	✓	X
mixed	✓	✓	✓	✓

back and forth pixel/harmonic domain, to apply each object where it naturally lives

Beams in parametric component separation

2nd solution: *in the* component separation

Input frequency maps (each convolved at its own beam $\mathbf{B}_{\mathbf{true}}$): $\mathbf{d} = \mathbf{B}_{\mathbf{true}}(\mathbf{A}\mathbf{s}) + \mathbf{n}$

It can be rewritten as: $\mathbf{d} = \mathbf{B}_{\mathbf{true}}\mathbf{B}_{\mathbf{final}}{}^{-1}\mathbf{A}\hat{\mathbf{B}}_{\mathbf{final}}\mathbf{s} + \mathbf{n} = \mathbf{A}'\mathbf{s}' + \mathbf{n}$

where: $\mathbf{s}' = \hat{\mathbf{B}}_{\mathbf{final}}\mathbf{s}$ is the recovered components beam-smoothed at the final resolution

$\mathbf{A}' = \mathbf{B}_{\mathbf{true}}\mathbf{B}_{\mathbf{final}}{}^{-1}\mathbf{A}$ (redefinition of the Mixing Matrix)

Comp sep:

➡ 2 step approach

Step 1 recover the spectral parameter estimates by maximising

$$-2\ln \mathscr{L}_{spec}(\beta_i) = \text{const} - (\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{d})^T(\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{A}')^{-1}(\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{d})$$

Step 2 recover the sky components

$$\mathbf{s} = (\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{A}')^{-1}\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{d}$$

Beams in parametric component separation

2nd solution: *in the* component separation

Comp sep:

➡ 2 step approach

Step 1 recover the spectral parameter estimates by maximising

$$-2\ln\mathscr{L}_{spec}(\beta_i) = \text{const} - (\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{d})^T(\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{A}')^{-1}(\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{d})$$

Step 2 recover the sky components

$$\mathbf{s} = (\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{A}')^{-1}\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{d}$$

Solved with a **PCG solver**: $\mathbf{A}'^T\mathbf{N}^{-1}\mathbf{A}'\mathbf{x} = \mathbf{A}'^T\mathbf{N}^{-1}\mathbf{d}$

Step 1 $\mathscr{L}_{spect} = -(\mathbf{d}^T\mathbf{N}^{-1}\mathbf{A}')\mathbf{x}$ (PCG at each step of the maximisation)

Step 2 recovered components: $\mathbf{x}$

bonus:
can handle input frequency maps at different **nsides**

Beams in parametric component separation

2nd solution: *in the* component separation

Adding the beam operator in the spectral likelihood

Pros vs Cons

Pros

- More rigorous treatment
- The recovered component maps resolutions are not limited by the worst input map resolution

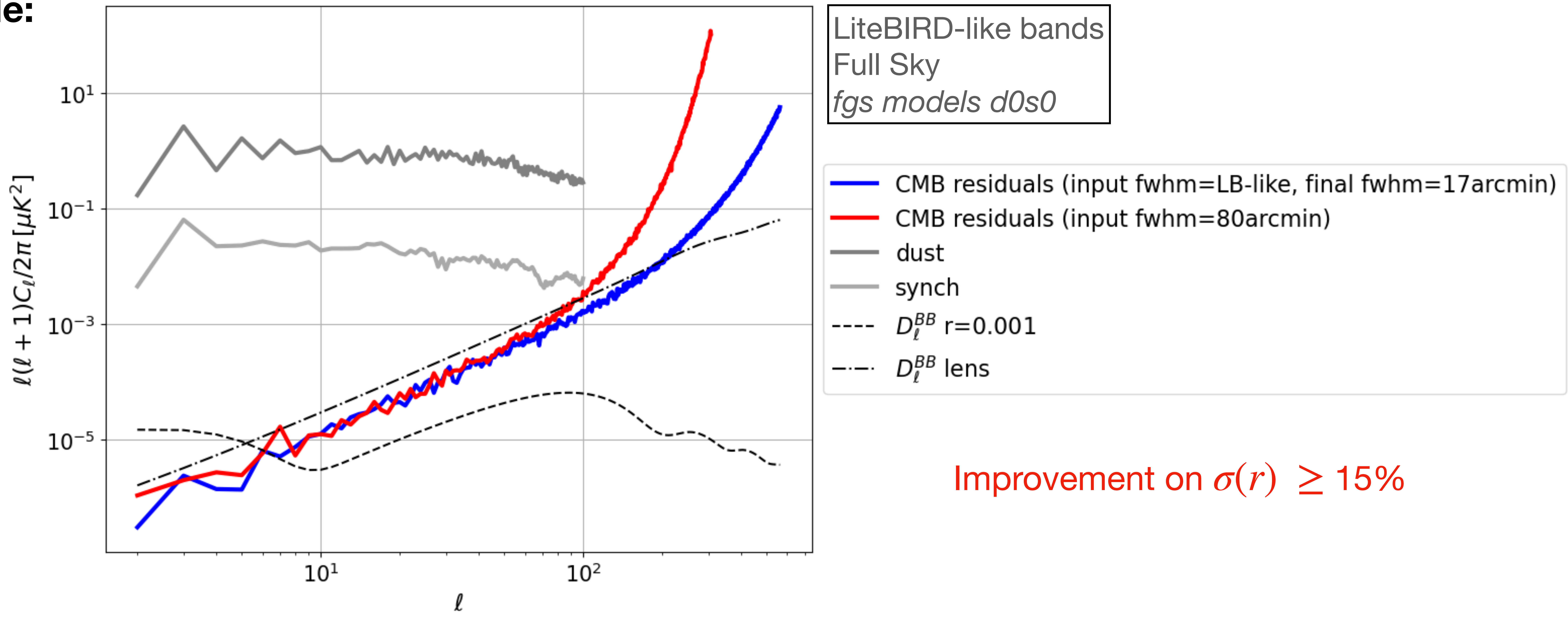
Cons

- Edge effects due to cut sky, where the convolution doesn't perform well
- Potentially more computationally involved

Beams in parametric component separation

2nd solution: *in the* component separation

Example:



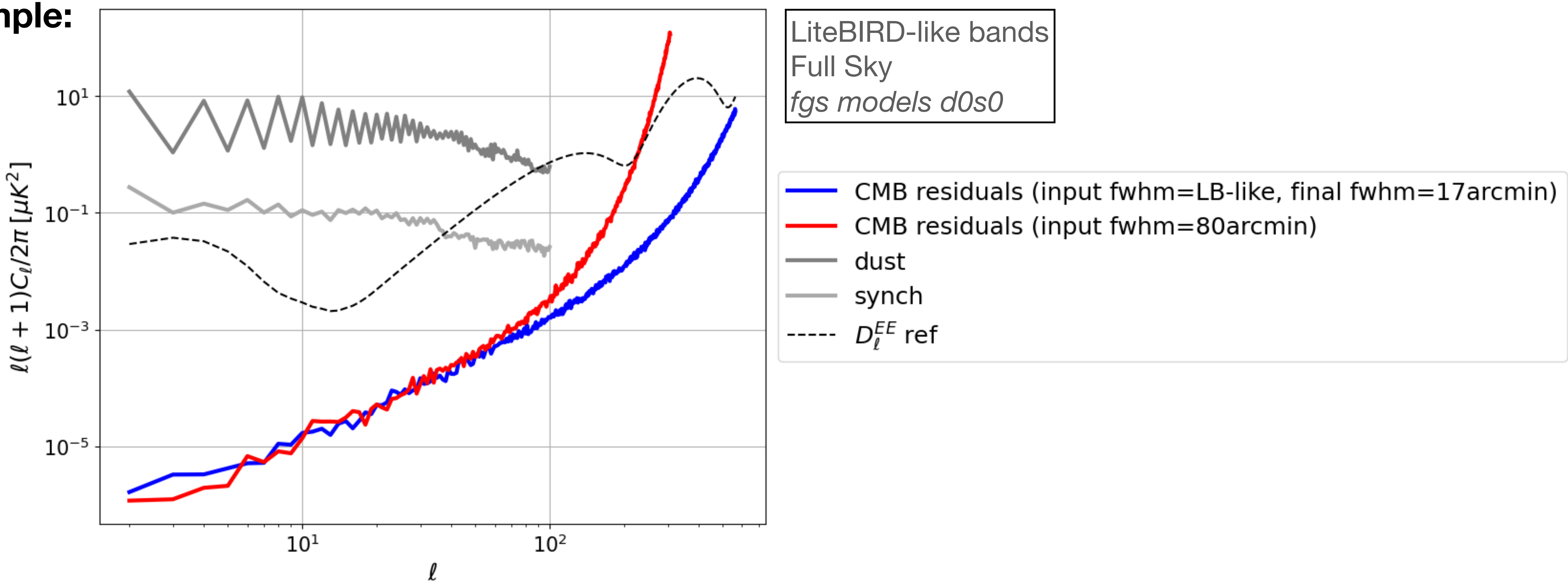
LiteBIRD-like bands
Full Sky
fgs models d0s0

Improvement on $\sigma(r) \geq 15\%$

Beams in parametric component separation

2nd solution: *in the* component separation

Example:



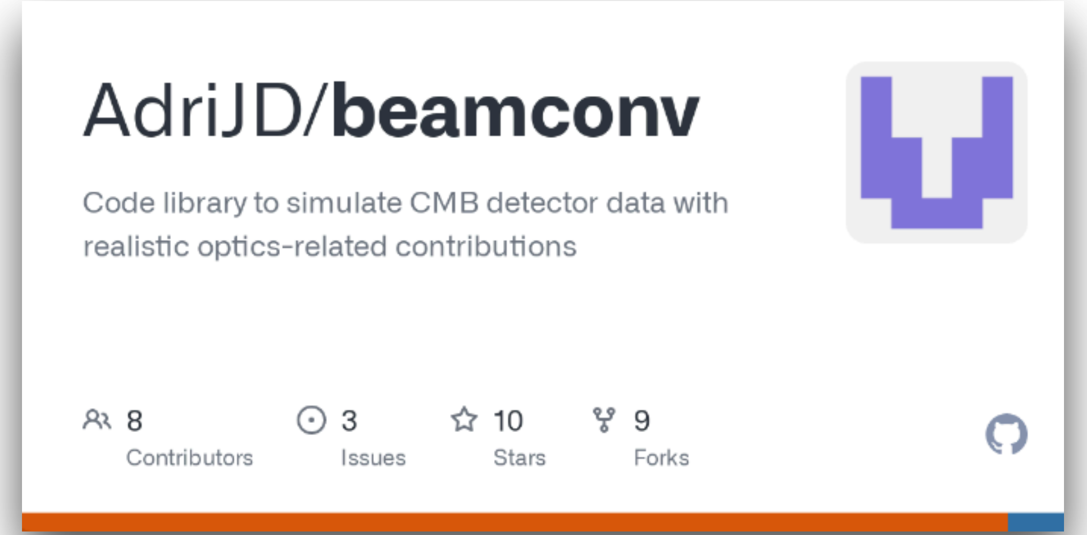
Conclusion and prospects

The input frequency maps are often convolved to different resolutions, we need a strategy to deal with this:

- 1st solution: dealing with the beams ***before*** component separation
- 2nd solution: we add the beam operator **in** the component separation method

↓

Apply the beam in the harmonic domain, but stay in pixel domain as much as possible (to be able to address the spatial variability of the foregrounds)



<u>*Next steps*</u>:

- study the effect of **more complex beam** in the component separation, exploiting more advanced softwares to perform the deconvolution ex.
- use implementation with PCG to deal with correlated noise → advantage of the mixed harmonic / pixel approach
- going to **TODs domain** (to include pixel dependent beams) → performing the component separation at the same time of the mapmaking

Thanks!

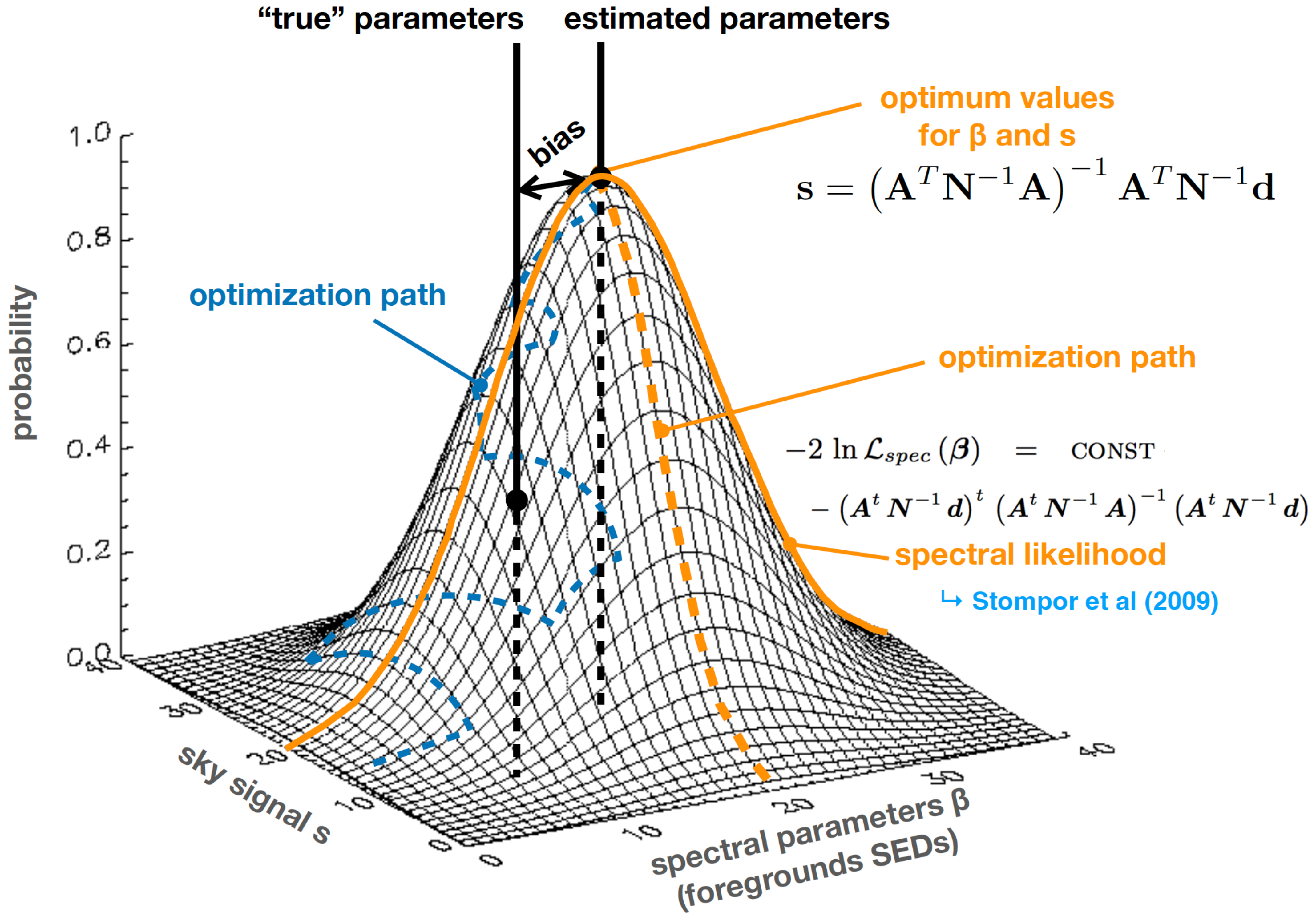
Questions?

Parametric maximum likelihood based component separation

The solution - two step procedure

Stompor et al. 2009

➡ 2 step procedure: we do not directly maximised the full data likelihood, but rather the *spectral likelihood* (computationally easier!)



Credits: J. Errard

Beams in parametric component separation

The problem: ***How to recover component maps from beam convolved frequency maps?***

$\mathbf{d} = \mathbf{BAs} + \mathbf{n}$ where $\mathbf{B}$ are the true beams (could be the same for all the frequencies or not)

Commutation of the beam operator and the Mixing Matrix:

➡ If $\mathbf{A}$ doesn't depend on the pixel and $\mathbf{B}$ is the same for all frequency channels: $\mathbf{BA} = \mathbf{A\hat{B}}$

The data model becomes: $\mathbf{d} = \mathbf{As}' + \mathbf{n}$ where $\mathbf{s}' = \mathbf{\hat{B}s}$ are beam-smoothed component maps.

(This can be extended to $\mathbf{A}$ with spatial variability as long as those are at scale bigger than the beam size.)

➡ If $\mathbf{A}$ does depend on the pixel: $\mathbf{BA} \neq \mathbf{A\hat{B}}$.

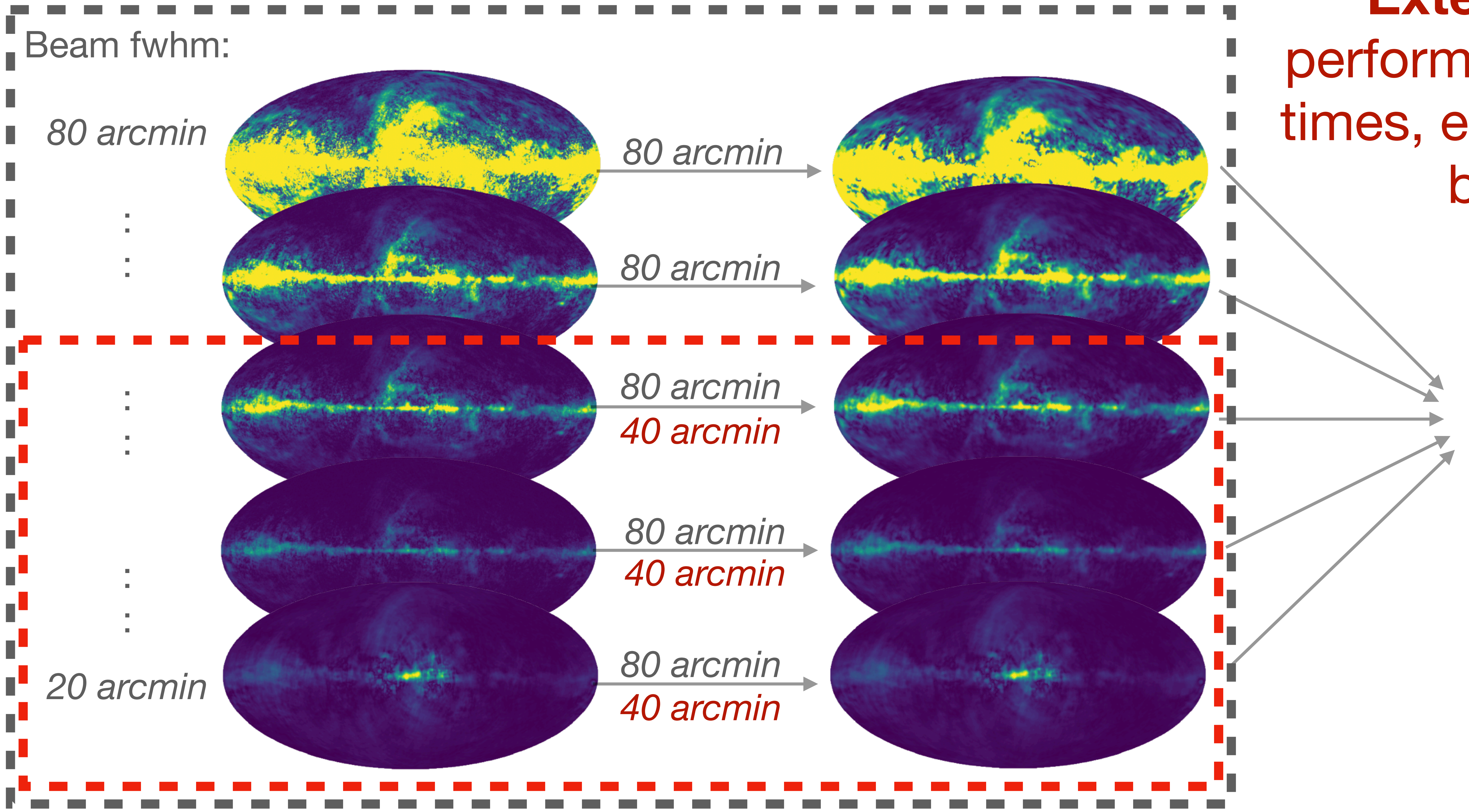
We can still have: $\mathbf{BA} = \mathbf{A}'\mathbf{\hat{B}}$, where $\mathbf{A}'$ is a modified ("effective") Mixing Matrix.

The data model becomes: $\mathbf{d} = \mathbf{A}'\mathbf{s}' + \mathbf{n}$

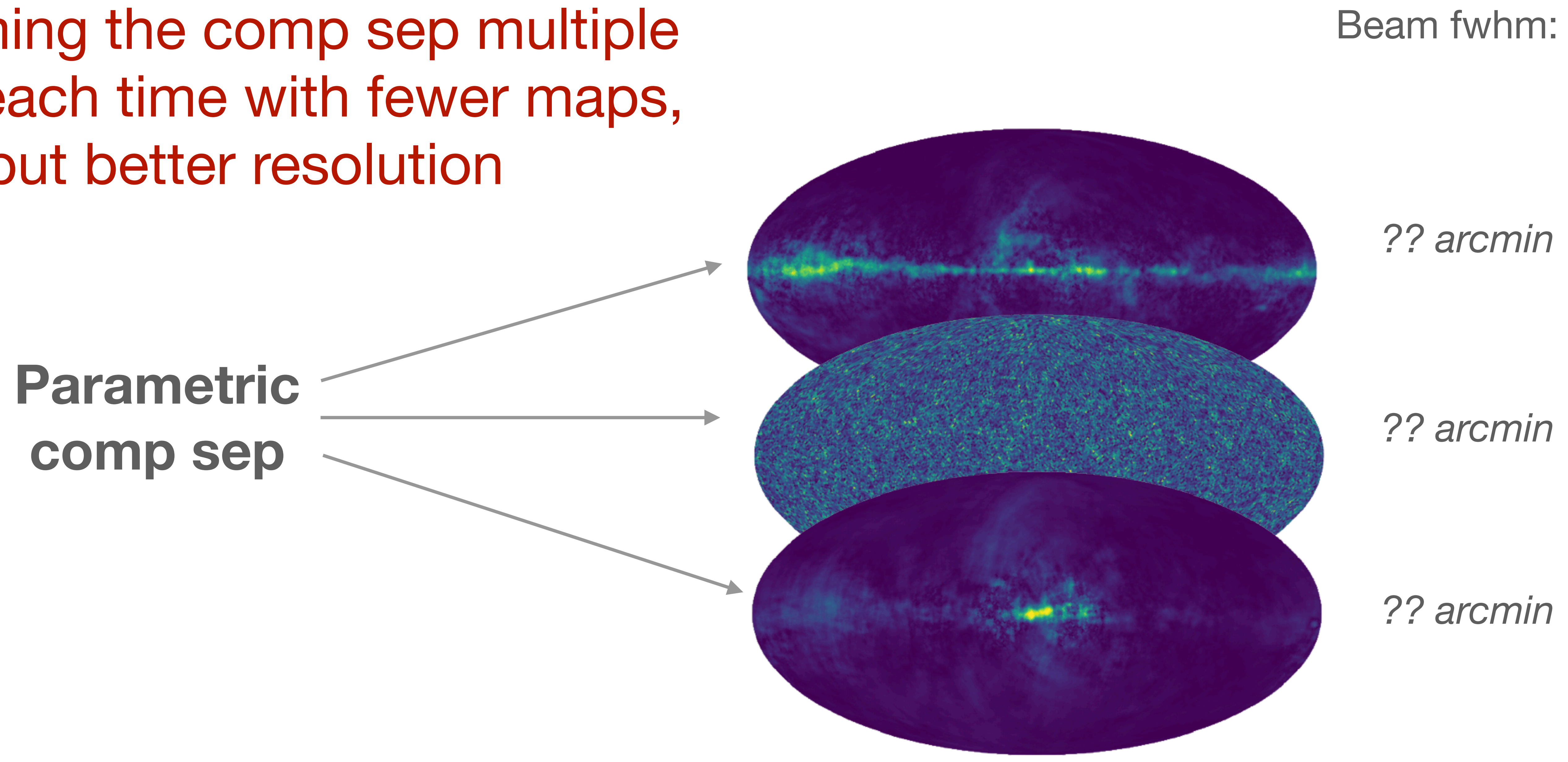
Beams in parametric component separation

Extension to 1st solution: *before* component separation

Extension to 1st solution:
performing the comp sep multiple times, each time with fewer maps, but better resolution



frequency maps
each smoothed at its beam value



component maps
smoothed at a common beam

Beams in parametric component separation

Extension to 1st solution: *before* component separation

Performing the comp sep multiple times, each time with less maps, but better resolution.

The resolution is limited by the resolution of the larger fwhm, therefore two regimes explored.

The final result is a combination of those.

	final (common) resolution	frequency band used	reconstructed components
regime 1	80 arcmin	all bands	CMB, dust, synch
regime 2	40 arcmin	excluded lower freq bands	CMB, dust