# CMB likelihoods at the frontier

Squeezing parameters out of data

Giacomo Galloni

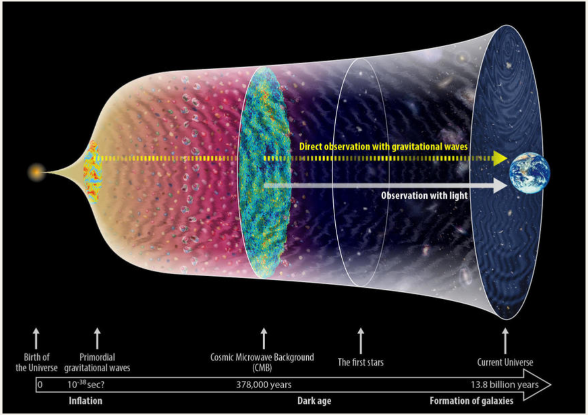
Challenges at the CMB Frontier · Chino, Japan · July 3, 2026

Department of Physics and Earth Science
University of Ferrara · INFN

# Inflation — and what it leaves behind

THE QUESTION

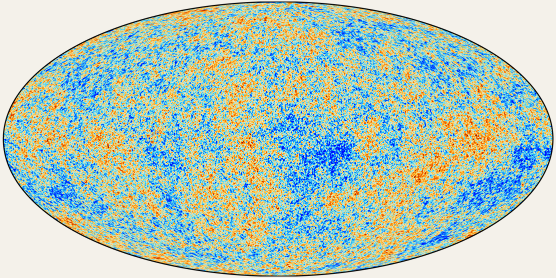
- Inflation is our best framework for the very early universe, yet still untested in detail.
- It predicts two kinds of primordial perturbations:
  - **scalar** (density) — the seeds of structure;
  - **tensor** (gravitational waves) — the smoking gun.
- Both leave an imprint on the cosmic microwave background.



# The CMB — our window on the early universe

THE QUESTION

- A snapshot of the universe at recombination ($z \approx 1100$).
- Tiny temperature fluctuations $\Delta T/T \sim 10^{-5}$ encode the seeds of all later structure.
- Subsequent reionization and primordial GWs leave their imprints in polarization.

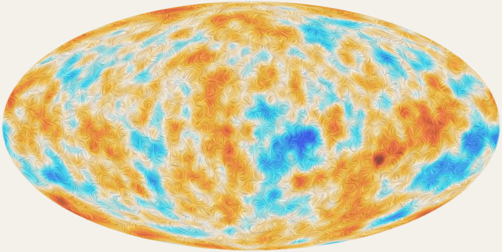


Planck Collaboration — arXiv:1807.06205

# CMB polarization — E-modes & B-modes

THE QUESTION

- Linear polarization decomposes into **E-modes** (parity-even) and **B-modes** (parity-odd).
- Scalar perturbations source E-modes; primordial GWs uniquely source large-scale B-modes.
- Both the reionization signal and primordial B-modes live at large angular scales (low $\ell$).



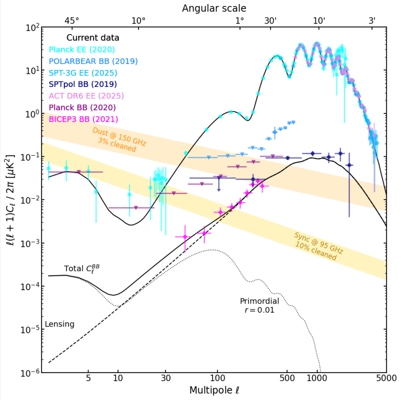
# The driving question

WHAT THIS TALK IS REALLY ABOUT

## The question

*How do we extract information on the parameters of interest from the data?*

- It hides real complexity: *which* data $D$, how to compress it, how to write a likelihood, how to dodge systematics.
- This era's challenge — a complete CMB description: small scales (not in this talk) *and* large scales.
- Large scales are hard: no central-limit theorem for the $C_\ell$.
- $\tau$ (EE) and $r$ (BB) are our proxies here.




*E/B-mode measurements to date — our grasp of the large scales is still partial.*

# The pipeline — sky to parameters via $\mathcal{L}(D \mid \theta)$

THE BRIDGE

Time-domain (raw detector) → Maps ($T, Q, U$) → Spectra ($C_\ell$) → Parameters ($\tau, r, \ldots$)

Each arrow compresses the data — but *how* differs:

- TOD $\rightarrow$ maps: through a likelihood (map-making);
- maps $\rightarrow$ spectra: brute linear compression (spherical-harmonic / QML estimators);
- spectra $\rightarrow$ parameters: a likelihood again — often with crude approximations of the experiment's response.

At each step the question is the same: *how much information do we keep, and what do we assume?*

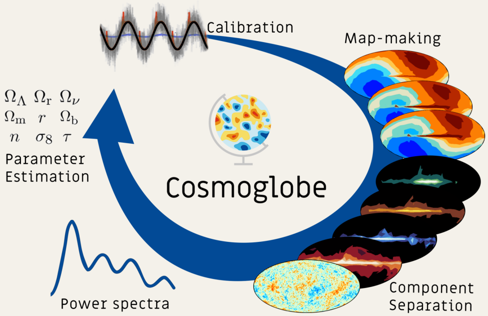
# $D$ = time-ordered data

THE RAW STREAM

The detectors deliver a time-ordered stream of raw data, tightly coupled to whatever happens in the sky as the instrument scans.

- It is the most informative data — every correlation is still in there.
- Working with it directly is the most ambitious likelihood — and computationally enormous; full end-to-end analyses are demonstrated only at limited scale.

Watts et al. — arXiv:2304.05484



# Map-making — compressing time into a map

TOD → MAPS · THE STEP WE WERE ASSUMING

Model the calibrated stream as $\mathbf{d} = \mathbf{P}\mathbf{m} + \mathbf{n}$, with $\mathbf{P}$ the pointing matrix ($T$, $Q$, $U$ projection) and $\mathbf{N} \equiv \langle \mathbf{n}\mathbf{n}^\top \rangle$ the time-domain noise covariance. Map-making compresses the time dimension: project the years-long stream onto a map.

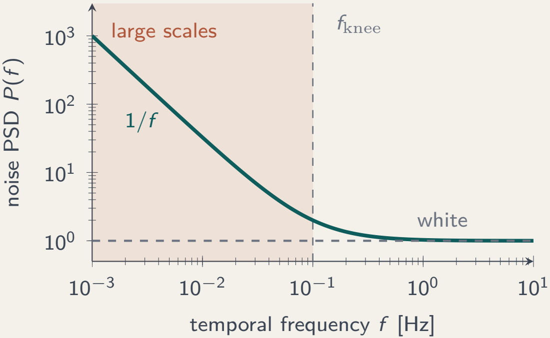
- Simplest, binning: average the samples falling in each pixel, $\hat{\mathbf{m}} = (\mathbf{P}^\top \mathbf{P})^{-1} \mathbf{P}^\top \mathbf{d}$.
- The natural place to remove some systematics: cosmic-ray glitches (deglitched on the timeline) and pointing / half-wave-plate wobble effects. → S. Stellati, Wed.
- The output is a map — the data product the rest of the pipeline consumes.

Biquard 2024 — arXiv:2406.15415 · Planck HFI — arXiv:1502.01586.

# Why map-making is hard — 1/f and the atmosphere

## THE NOISE IS NOT WHITE

- Detector noise has a $1/f$ tail: power rises below a *knee* frequency $f_{\mathrm{knee}}$.
- Low temporal frequency $\leftrightarrow$ large angular scales $\leftrightarrow$ low $\ell$ — the $\tau$/B-mode regime.
- Ground-based: the atmosphere adds a huge, detector-correlated $1/f$-like signal — the dominant contaminant.
- *Naive binning* averages as if white $\Rightarrow$ stripes and lost large-scale power.



# From binning to the optimal map

GENERALISED LEAST SQUARES

Binning is optimal *only* for white noise. With $1/f$ correlations we must do better:

- Destriping: model the $1/f$ part as a few baseline offsets, fit and subtract them, then bin — cheap, and recovers most of the optimal answer.
- Map-making as a likelihood: maximise the Gaussian TOD likelihood $-2\log\mathcal{L} = (\mathbf{d}-\mathbf{P}\mathbf{m})^\top\mathbf{N}^{-1}(\mathbf{d}-\mathbf{P}\mathbf{m})$. Its generalised-least-squares (minimum-variance) maximum and pixel covariance are

$$\hat{\mathbf{m}} = (\mathbf{P}^\top\mathbf{N}^{-1}\mathbf{P})^{-1}\mathbf{P}^\top\mathbf{N}^{-1}\mathbf{d}, \quad \mathbf{C}_N = (\mathbf{P}^\top\mathbf{N}^{-1}\mathbf{P})^{-1}.$$

- It is the optimal compression — and reduces to binning when the noise is white ($\mathbf{N} \propto \mathbf{I}$).

Biquard 2024 — arXiv:2406.15415.

# Challenges of map-making

WHAT MAKES IT HARD

- **Correlated noise, where it hurts.** Instrumental $1/f$ + atmosphere (ground) pile up at low $\ell$ — the $\tau$/B-mode regime; detector decorrelation worsens the lowest multipoles.
- **Large-scale power loss.** Time-domain filtering / ground removal suppress the biggest scales; unbiased recovery needs a transfer function (filter-and-bin) or full GLS.
- **Data volume.** $\sim 10^4$–$10^5$ detectors $\times$ years, and growing. Optimal map-making is memory/IO-bound, pushing to massively-parallel HPC and GPU/exascale.
- **Systematics into $\mathbf{C}_N$.** Gain, pointing, bandpass, half-wave plates, $T \to P$ leakage must be propagated into the noise model.
- **The covariance wall.** The dense pixel–pixel $\mathbf{C}_N$ is infeasible at native resolution — the same memory + noise-covariance limit as the exact likelihood.

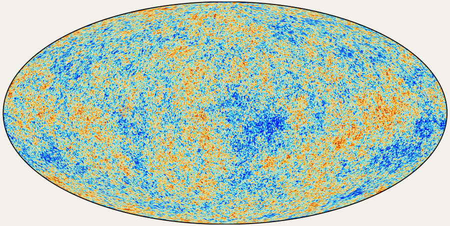
Biquard 2024 · CLASS 2025 (arXiv:2501.11904) · Planck HFI (arXiv:1502.01586).

# $D$ = pixels — the exact likelihood

PIXEL-BASED · THE GOLD STANDARD

$$-2\log\mathcal{L} \;=\; \mathbf{m}^{\top}\,\mathrm{COV}^{-1}\,\mathbf{m} \;+\; \log|\mathrm{COV}|$$

$$\mathbf{m} = \{T, Q, U\}, \quad \mathrm{COV} = \mathbf{S}(C_\ell) + \mathbf{C}_N$$



*Planck temperature map $T(\hat{n})$.*

- Exact: a multivariate Gaussian in *pixel* space, even on the cut sky — here $\mathbf{C}_N$ is the dense pixel–pixel map-making covariance.
- Template systematics marginalized analytically: add $\sigma^2\mathbf{T}\mathbf{T}^{\top}$ to $\mathrm{COV}$, $\sigma^2 \to \infty$ (foreground, monopole/dipole).
- Cost cubic, $\mathcal{O}(N_{\mathrm{pix}}^3)$ — long held intractable beyond very low resolution.

Gerbino et al. — arXiv:1909.09375 · template marg.: Planck 2018 V — arXiv:1907.12875.

# PICSLike — the exact likelihood, made easy

## REMOVING THE HISTORICAL ENTRY BARRIER

- PICSLike is a fast, user-friendly engine for the exact pixel-based likelihood — a few lines of config, not a bespoke pipeline.
- It dissolves two historical barriers: the entry cost of writing one, and the difficulty at large scales — now fully tractable at realistic resolution.
- The bottleneck is no longer CPU but memory and the noise-covariance model.
- Fully modular: cleanly generalizes to template fitting, new noise models, extra components, and more.

### LiteBIRD-like regime

- $N_{\rm side} = 16$, $f_{\rm sky} \approx 0.5$: interactive ($\sim$ seconds).
- $N_{\rm side} = 64$, $f_{\rm sky} \approx 0.6$, joint $Q, U$: $\sim$47 min on one node.

*Exact, end-to-end, at real survey sizes.*

Galloni & Pagano 2026 — arXiv:2605.21149 (CosmoForge I).

# CosmoForge — one framework

## ONE WORKSPACE: COSMOCORE · QUBE · PICSLIKE

**d** | mask | **N** | $C_\ell^{\rm fid}$ | $B_\ell, p_\ell$

↓

**CosmoCore** — *ComputationBasis interface*

*HarmonicBasis (SMW)*

$\mathbf{V},\ \Lambda(C_\ell^{\rm fid})$

$\mathbf{M}=\mathbf{V}\mathbf{N}^{-1}\mathbf{V}^\top,\ \ \mathbf{K}=\Lambda^{-1}+\mathbf{M}$

$\mathbf{C}^{-1}=\mathbf{N}^{-1}-\mathbf{N}^{-1}\mathbf{V}^\top\mathbf{K}^{-1}\mathbf{V}\mathbf{N}^{-1}$

$\ln|\mathbf{C}|=\ln|\mathbf{N}|+\ln|\Lambda|+\ln|\mathbf{K}|$

*PixelBasis (direct)*

build $\mathbf{S}$ in pixel space:

$\mathbf{S}_{pp'}=\sum_\ell C_\ell W_\ell P_\ell(\cos\theta_{pp'})$

form $\mathbf{C}=\mathbf{S}+\mathbf{N}$

$\mathbf{C}^{-1}$, $\ln|\mathbf{C}|$ from Cholesky factorisation

↓ $\mathbf{C}^{-1}$, $\ln|\mathbf{C}|$

**QUBE**

$q_\ell=\frac{1}{2}\mathbf{d}^\top\mathbf{C}^{-1}\mathbf{P}_\ell\mathbf{C}^{-1}\mathbf{d}-b_\ell$

$F_{\ell\ell'}=\frac{1}{2}\,\mathrm{Tr}[\mathbf{C}^{-1}\mathbf{P}_\ell\mathbf{C}^{-1}\mathbf{P}_{\ell'}]$

$\hat{C}_\ell=(\mathbf{F}^{-1}\mathbf{q})_\ell$ three normalisations

**PICSLike**

$\ln\mathcal{L}(\theta)=-\frac{1}{2}\left[\mathbf{d}^\top\mathbf{C}^{-1}(\theta)\mathbf{d}+\ln|\mathbf{C}(\theta)|\right]$

over a parameter grid $\theta$

COSMOFORGE QUBE PICSLike

`pip install cosmoforge` · v1.0.1 · QML estimator + exact pixel likelihood, one config.

# $D = C_\ell$ (full sky)

SPECTRUM-BASED · WHY WE COMPRESS

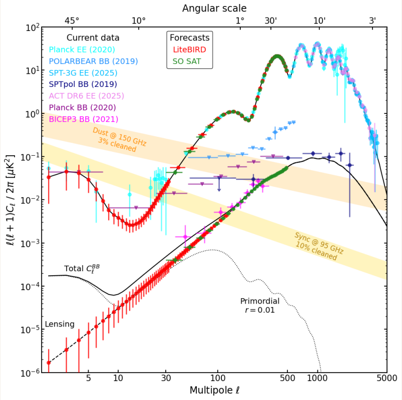
$$a_{\ell m} = \int Y^*_{\ell m}(\theta,\varphi)\,\frac{\Delta T}{T_0}\,\mathrm{d}\cos\theta\,\mathrm{d}\varphi$$

$$\langle a^*_{\ell m}\, a_{\ell' m'}\rangle = \delta_{\ell\ell'}\,\delta_{mm'}\,C_\ell$$

Wishart form (full sky):

$$-2\log\mathcal{L} = \sum_\ell (2\ell+1)\left[\frac{\widehat{C}_\ell}{C_\ell} - \log\frac{\widehat{C}_\ell}{C_\ell} - 1\right]$$

- Compresses maps to angular spectra — cheap, and exact on the full sky; scales where pixel-based cannot.
- Some systematics are natural here: the beam is (nearly) diagonal in $\ell$, and $E \to B$ leakage from the mask is handled by purification / pure-$B$ estimators.




*EE / BB spectra: current data and forecasts.*

# The catch — galactic foregrounds force a mask

COMPLICATION

Spherical harmonics are orthonormal and complete on the *full* sphere:

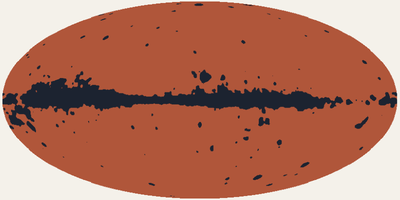
$$\int_{S^2} Y^*_{\ell m}\, Y_{\ell' m'}\, \mathrm{d}\Omega \;=\; \delta_{\ell\ell'}\, \delta_{mm'}\,.$$

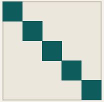
A mask restricts the integral to a subset of the sky:

$$\int_{S^2\setminus\mathrm{mask}} Y^*_{\ell m}\, Y_{\ell' m'}\, \mathrm{d}\Omega \;=\; W^{\ell\ell'}_{mm'} \;\neq\; \delta_{\ell\ell'}\delta_{mm'}\,.$$

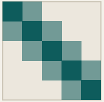
Different multipoles couple ($M_{\ell\ell'}$ broadens around the diagonal) and the Wishart form fails:

$$-2\log\mathcal{L} \;\neq\; \sum_{\ell}(2\ell+1)\left[\frac{\widehat{C}_\ell}{C_\ell} - \log\frac{\widehat{C}_\ell}{C_\ell} - 1\right].$$





full sky: $\delta_{\ell\ell'}$



cut sky: banded

# Two first attempts

APPROXIMATIONS 1 & 2 OF 4

### Fiducial Gaussian

$$-2\log\mathcal{L} = \left(\widehat{\boldsymbol{C}}_{\rm obs} - \boldsymbol{C}_{\rm th}\right) \mathrm{COV}_{\rm fid}^{-1} \left(\widehat{\boldsymbol{C}}_{\rm obs} - \boldsymbol{C}_{\rm th}\right)^{\top}$$

| | |
|---|---|
| $\sim$ **full-sky exact** | $\checkmark$ $\ell\ell'$ **couplings** |
| $\times$ **non-Gaussian $C_\ell$** | $-$ **negative-$C_\ell$ safe** |

- Couplings via a *fixed* covariance.
- Assumes Gaussian $C_\ell$ — fails at low $\ell$.

### Exact-with-$f_{\rm sky}$

$$-2\log\mathcal{L} = f_{\rm sky}\sum_{\ell}(2\ell+1)\left[\frac{C_\ell^{\rm obs}}{C_\ell^{\rm th}} - \log\frac{C_\ell^{\rm obs}}{C_\ell^{\rm th}} - 1\right]$$

| | |
|---|---|
| $\checkmark$ **full-sky exact** | $\times$ $\ell\ell'$ **couplings** |
| $\checkmark$ **non-Gaussian $C_\ell$** | $-$ **negative-$C_\ell$ safe** |

- Keeps the non-Gaussian shape; rescales the d.o.f. by the sky fraction $f_{\rm sky}$.
- Disregards multipole couplings entirely.

*Each fixes one failure of the other — neither captures couplings and non-Gaussianity at once.*

# Approximation 3 / 4 — Hamimeche–Lewis (HL)

APPROXIMATION 3 OF 4

✓ full-sky exact | ✓ $\ell\ell'$ couplings | ✓ non-Gaussian $C_\ell$ | × negative-$C_\ell$ safe

$$-2\log\mathcal{L} = \overrightarrow{\left[g\left(\frac{\boldsymbol{C}_{\rm obs}}{\boldsymbol{C}_{\rm th}}\right)\ \boldsymbol{C}_{\rm fid}\right]}\mathrm{COV}^{-1}_{\rm fid}\overrightarrow{\left[g\left(\frac{\boldsymbol{C}_{\rm obs}}{\boldsymbol{C}_{\rm th}}\right)\ \boldsymbol{C}_{\rm fid}\right]}^{\top}$$

$$g(x) = \mathrm{sign}(x-1)\sqrt{2(x-\log x-1)}$$

- $g(\,\cdot\,)$ symmetrises the Wishart shape around 1.
- Captures multipole couplings *and* non-Gaussianity.
- Negative $C_\ell$ values break the square root.

Hamimeche & Lewis — arXiv:0801.0554

# Approximation 4 / 4 — Offset HL

APPROXIMATION 4 OF 4

✓ full-sky exact | ✓ $\ell\ell'$ couplings | ✓ non-Gaussian $C_\ell$ | ✓ negative-$C_\ell$ safe

$$-2\log\mathcal{L} = \left[\overrightarrow{g\left(\tfrac{\boldsymbol{C}_{\rm obs}+\boldsymbol{O}_\ell}{\boldsymbol{C}_{\rm th}+\boldsymbol{O}_\ell}\right)\ \left(\boldsymbol{C}_{\rm fid}+\boldsymbol{O}_\ell\right)}\right]\mathrm{COV}^{-1}_{\rm fid}\left[\overrightarrow{g\left(\tfrac{\boldsymbol{C}_{\rm obs}+\boldsymbol{O}_\ell}{\boldsymbol{C}_{\rm th}+\boldsymbol{O}_\ell}\right)\ \left(\boldsymbol{C}_{\rm fid}+\boldsymbol{O}_\ell\right)}\right]^\top$$

$$g(x) = \mathrm{sign}(x)\,\mathrm{sign}(|x|-1)\sqrt{2\left(|x|-\log|x|-1\right)}$$

- Same machinery as HL with an additive offset $\boldsymbol{O}_\ell$.
- Offsets shift away from the negative-$C_\ell$ problem.
- Workhorse for low-$\ell$ B-mode polarisation.

A. Mangilli et al. — arXiv:1503.01347

# From one field to many

THE QUESTION THAT GIVES BIRTH TO MHL

- The spectrum likelihoods so far — Fiducial Gaussian, exact-$f_{\rm sky}$, HL, offset HL — were single-field: one map, one spectrum per $\ell$.
- But real analyses — and the future — combine many fields and experiments: channels, splits, whole surveys of the same sky.
- The challenge: consistency checks *between* experiments and *within* one (as we'll do for BICEP). When noise biases can't be modelled perfectly, cross-spectra may be the only robust way forward.

**Two questions**

1. How do we combine multiple fields into *one* coherent likelihood?
2. And how do we use *only the cross-spectra* — the part we trust — rigorously?

# Multi-field setup — and the noise-bias problem

## THE STARTING POINT

We proceed with HL — the best-performing approximation at large scales. With $n_{\mathrm{ch}}$ fields (channels or splits) the spectra at each $\ell$ form a symmetric matrix (e.g. $n_{\mathrm{ch}} = 3$):

$$\mathbf{C}_\ell = \begin{pmatrix} C_\ell^{A\times A} + N_\ell^A & C_\ell^{A\times B} & C_\ell^{A\times C} \\ C_\ell^{A\times B} & C_\ell^{B\times B} + N_\ell^B & C_\ell^{B\times C} \\ C_\ell^{A\times C} & C_\ell^{B\times C} & C_\ell^{C\times C} + N_\ell^C \end{pmatrix}$$

- Auto-spectra (diagonal) contain $\langle|\mathrm{noise}|^2\rangle$: we subtract an estimate $\widehat{N}_\ell^X$. If $\widehat{N}_\ell^X \neq N_\ell^{X,\mathrm{true}}$, the data vector is biased $\Rightarrow$ parameters shift.
- Cross-spectra carry no noise bias in expectation — the robust part of the data vector $X_\ell = (X_\ell^{AA}, X_\ell^{AB}, \ldots, X_\ell^{CC})$.

# The fix — marginalized HL (mHL)

CROSS-ONLY BY MARGINALIZATION [NEW]

Keep only the trusted crosses. Empirically $X_\ell \sim \mathcal{N}(\mu, M_{\ell\ell'})$, so apply the standard Gaussian marginalization rule:

$$Y \sim \mathcal{N}(\mu, \Sigma) \implies Y_{\text{cross}} \sim \mathcal{N}(\mu_{\text{cross}}, \, \Sigma_{\text{cross,cross}})$$

Build $X_\ell^{\text{cross}} = (X_\ell^{A\times B}, X_\ell^{A\times C}, X_\ell^{B\times C})$ and drop the rows/cols of $M_{\ell\ell'}$ for the auto-elements:

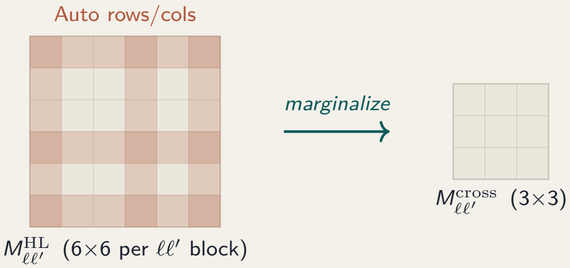
$$-2\log\mathcal{L} = \sum_{\ell\ell'} X_\ell^{\text{cross}\,T} \, [M^{\text{cross}}]^{-1}_{\ell\ell'} \, X_{\ell'}^{\text{cross}}$$

*mHL answers both questions — combine all fields, lean only on the trusted crosses.*

Galloni et al. 2025 — arXiv:2505.24829

# The geometric picture — strip auto rows/cols

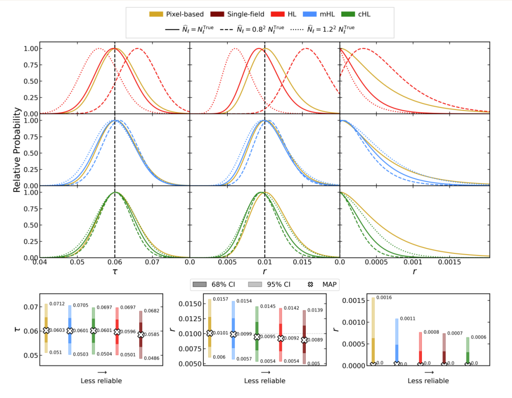
WHAT MHL DOES TO THE COVARIANCE




*Drop the rows and columns associated with $X^{X\times X}$. No new offsets, no per-cross transformation — just a sub-block of the original covariance.*

# Validation — and the takeaway

## CUT-SKY $f_{\rm sky} = 40\%$ VS THE EXACT PIXEL-BASED TRUTH




*Setup: 3 channels, $\ell_{\max} = 32$, $\pm 20\%$ noise-bias error; ground truth = exact pixel-based.*

- All HL approximations are over-confident vs PB — mHL the least.
- Exact within HL, no new offsets; robust to noise misestimation.
- Multi-field mHL beats single-field HL on the combined map.
- Public code: `CHARM-Like`.

*Now: put it to work on real data.*

# ELiCA — $\tau$ from all large-scale polarisation

MHL ON REAL DATA · IN BRIEF

- ELiCA combines Planck 100, 143 GHz and 70 GHz + WMAP as three fields in one mHL likelihood.
- It uses all the large-scale polarisation data available today, combined for the first time, to update the optical depth $\tau$.

Genesini et al. 2026 — arXiv:2603.22454.

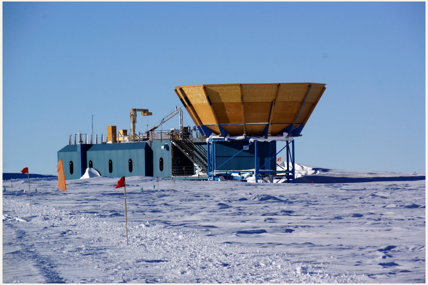
**A whole talk of its own**

Led by my student **V. Genesini** — the results are in her dedicated talk. Here, just proof that mHL works on real data.

# Diagnosing BICEP/Keck 2018 with mHL

THE SETTING

- Constraints on the tensor-to-scalar ratio $r$ from low-$\ell$ BB polarisation.
- 11 frequency maps combined into one likelihood:
  - BICEP3/Keck at 95, 150, 220 GHz + $\mathrm{B95_e}$
  - Planck at 30, 44, 143, 217, 353 GHz
  - WMAP at 23 and 33 GHz
- Eight sampled parameters: $r$ + 7 foreground (dust & synchrotron sector).
- Many auto-spectra with different noise properties $\Rightarrow$ natural setting for mHL.



*BICEP Array at the South Pole.*

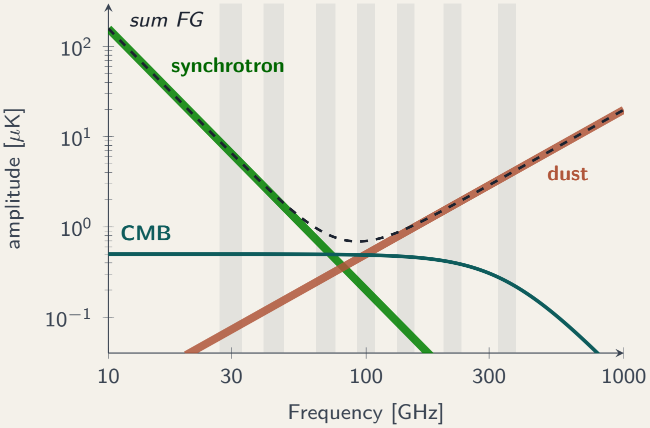
# Galactic foregrounds — dust and synchrotron

## WHY SO MANY CHANNELS AND PARAMETERS

The same foregrounds that forced the mask earlier now set the channel count. Two polarised galactic emissions dominate the large-scale B-mode sky:

- **Dust** — modified blackbody, rising at $\nu \gtrsim 100\,\text{GHz}$ ($\beta_d \sim 1.6$).
- **Synchrotron** — $e^-$ in galactic $\boldsymbol{B}$, falling at $\nu \lesssim 50\,\text{GHz}$ ($\beta_s \sim -3.1$).

Both smooth on the sky ($D_\ell^{\text{fg}} \propto \ell^\alpha$, $\alpha \sim -0.5$): only frequency dependence separates them from CMB — the job of component separation (see A. Carones, N. Krachmalnicoff).



# The test — mHL as a consistency tool

MHL IS MORE THAN A ROBUST $r$

mHL rests on one fact: after the $g$-transform the data vector $X$ is (empirically) Gaussian. So we apply the standard multivariate marginalization rule to keep only the cross-spectra.

- That already gives a consistency check: if the result *moves* when we marginalize a channel/auto, something may be wrong with the autos.
- Pushing the same Gaussianity further, the conditional likelihood is known too — so the full BICEP likelihood factorizes exactly, no approximation:

$$\mathcal{L}_{\mathrm{full}}(\theta) = \underbrace{\mathcal{L}_{\mathrm{marg}}(\theta)}_{\text{cross only (mHL)}} \cdot \underbrace{\mathcal{L}_{\mathrm{cond}}(\theta)}_{\text{auto} \mid \text{cross}}$$

- This unlocks a much deeper test: cross (marginal) vs auto | cross (conditional) accordance, parameter by parameter.

Galloni et al. 2025 — arXiv:2505.24829.

# Quantifying agreement — suspiciousness

## A PRIOR-INDEPENDENT TENSION MEASURE

The Bayes ratio $R$ measures concordance of two posteriors (here cross vs auto) but depends on the prior. The information $I$ — how much the data shrink the prior — depends on it the *same way*, so the difference cancels it:

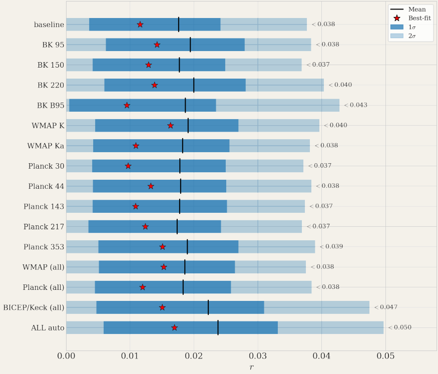
$$\ln R = \ln \frac{Z_{12}}{Z_1 Z_2}, \quad \ln I = D_{\mathrm{KL},1} + D_{\mathrm{KL},2} - D_{\mathrm{KL},12}, \quad \boxed{\ln S = \ln R - \ln I}$$

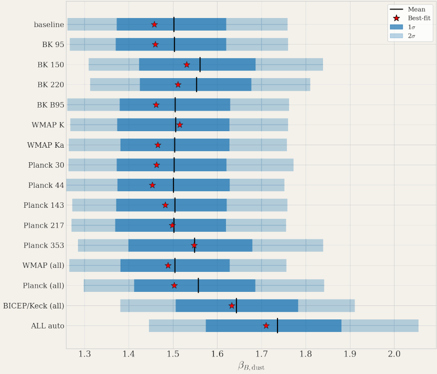
- $\ln S$ (suspiciousness) is therefore prior-independent.
- $d - 2\ln S \sim \chi^2_d$ ($d$ = effective shared constrained parameters): it scores both a discrepancy ($\ln S > 0$) and an over-concordance ($\ln S < 0$); $\ln S = 0$ is perfect concordance.

Marshall et al. 2006 · Handley & Lemos 2019.

# mHL on BK18 — marginalized constraints

## APPLY *JUST* MHL — MARGINALIZE EACH CHANNEL · PRELIMINARY

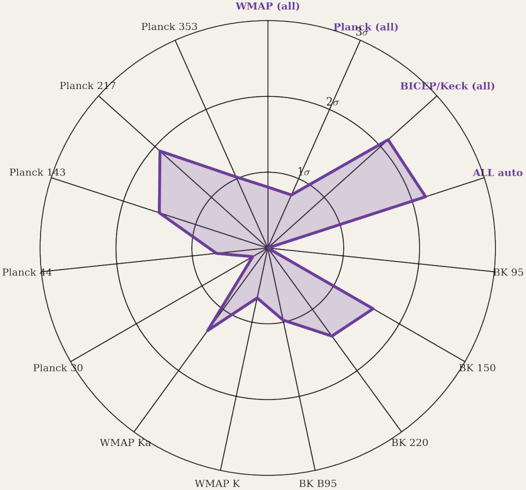



*Baseline $r < 0.038 \rightarrow$ all-auto marginalized $r < 0.050$: the 95% bound degrades by $\approx 32\%$, while $\beta_{\mathrm{dust}}$ stays put — mHL alone, no conditional yet.*

# Suspiciousness across channels

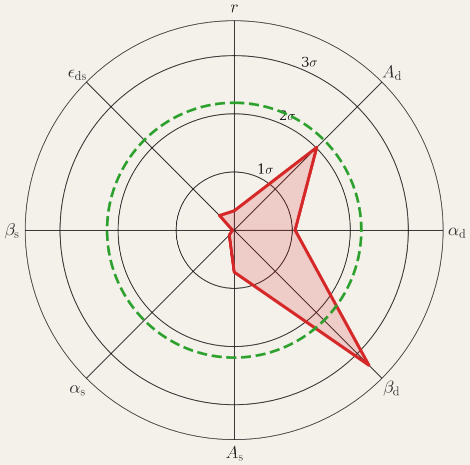
CROSS VS AUTO, PER CHANNEL AND COMBINATION · PRELIMINARY




- Suspiciousness between the cross (marginal) and auto | cross (conditional) posteriors, per channel.
- All-auto reaches $2.2\sigma$; all BICEP/Keck $2.1\sigma$; Planck / WMAP alone $\lesssim 0.8\sigma$.
- Single channels are weakly-constrained conditionals — not a shared-parameter signal.

# Where the discrepancy lives — parameter shifts

PER PARAMETER, MARGINAL VS CONDITIONAL · PRELIMINARY



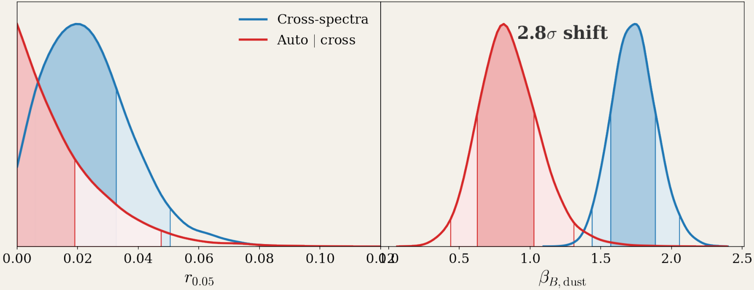

Shift between the cross (marginal) and auto | cross (conditional) means:

$$S_i = \frac{(\mu_i^\mathrm{cr} - \mu_i^\mathrm{cd})^2}{\sigma_i^{2,\mathrm{cr}} + \sigma_i^{2,\mathrm{cd}}}.$$

- "cd" is the autos given the crosses — *not* a standalone auto fit.
- $\beta_\mathrm{dust}$ and $A_\mathrm{dust}$ stand out — the whole dust sector; $r$ and the rest stay small.

## $r$ vs $\beta_{\rm dust}$

CROSS (MARGINAL) VS AUTO | CROSS (CONDITIONAL) · PRELIMINARY

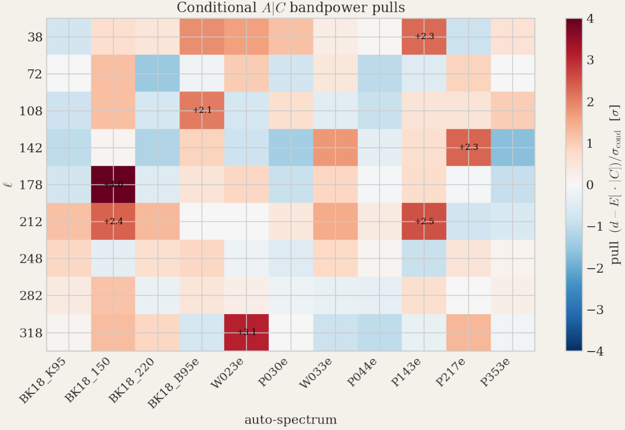



- $r_{0.05}$ stable: cross 0.024 vs auto 0.017 ($0.3\sigma$).
- $\beta_{\rm dust}$ shifts: cross $1.74 \pm 0.16$ vs auto $0.86 \pm 0.22 \Rightarrow 2.8\sigma$ (non-Gaussian, quoted — the naive Gaussian estimate gives $3.3\sigma$).

*The "auto" curve is the conditional (autos given the crosses), not a standalone auto fit.*

# Where in the data — bandpower pulls

AUTO | CROSS RESIDUALS, PER CHANNEL AND MULTIPOLE · PRELIMINARY




Standardized residual of each auto bandpower against what the cross predicts for it:

$$\mathrm{pull}_\ell = \frac{\hat{X}^{\mathrm{a}}_\ell - \bar{X}^{\mathrm{a}}_\ell}{\sigma_{\mathrm{a|c}}}.$$

- Dominated by one cell: BK 150 GHz at $\ell \approx 178$, $+4.04\sigma$.
- No other cell exceeds $\sim 3.1\sigma$ (a lone noisy WMAP 23 GHz bin).
- Not broadband — one channel, one multipole bin.

All BICEP results preliminary — Galloni & Ballardini, in prep.

# A diagnostic tool

FROM A LIKELIHOOD FIX TO A DIAGNOSTIC TOOL

- mHL's exact factorization is a powerful, model-faithful diagnostic of an experiment's self-consistency: here, $r$ robust, a dust-sector cross/auto discrepancy.
- Naturally promising for multi-experiment setups — the same construction cross-checks *different experiments*.
- All data points are *first-class citizens* of one data vector (like different fields), so marginalizing the high-$\ell$ part pinpoints the low-$\ell$ origin of the discrepancy.
- The origin is an open question — a residual effect in the deep BK auto-spectra; we do *not* present it as a detection.

All BICEP results preliminary — Galloni & Ballardini, in prep.

# Open challenges — and two we softened

IN ONE SLIDE

- Extracting parameters at large scales is hard — the likelihood is non-trivial at *every* step (map-making, the exact pixel likelihood, the $C_\ell$ approximations), and the open challenges are many.
- Today we softened two:
  - `PICSLike` — the exact pixel likelihood, now fast and tractable at large scales;
  - `mHL` — robust multi-field $C_\ell$ likelihoods, *and* an exact-factorization self-consistency test.
- On real data: $r$ robust from BICEP/Keck (with a preliminary dust-sector cross/auto discrepancy), and $\tau$ from all large-scale polarisation (V. Genesini).
- The same direction: making future, harder CMB measurements — down to B-modes at $r \sim 10^{-3}$ — trustworthy.

# Thank you

*giacomo.galloni@unife.it*

APPENDIX

# Backup slides

*Material for questions.*

# HL likelihood — the eigenvalue trick

BACKUP · HAMIMECHE & LEWIS 2008

Start from the exact full-sky Wishart likelihood at fixed $\ell$, with $\widehat{\boldsymbol{C}}_\ell \sim \mathcal{W}_n(\boldsymbol{C}_\ell, 2\ell+1)$. Diagonalize:

$$\boldsymbol{D}_\ell \;=\; \boldsymbol{C}_\ell^{-1/2}\,\widehat{\boldsymbol{C}}_\ell\,\boldsymbol{C}_\ell^{-1/2} \;=\; \boldsymbol{U}_\ell\,\boldsymbol{\Lambda}_\ell\,\boldsymbol{U}_\ell^T,$$

and define the change of variable

$$\boldsymbol{g}(\boldsymbol{D}_\ell) \;=\; \boldsymbol{U}_\ell\,\mathrm{diag}\big[g(\lambda_i)\big]\,\boldsymbol{U}_\ell^T, \qquad g(x) \;=\; \mathrm{sign}(x-1)\sqrt{2\,(x-\log x-1)}.$$

In this basis, the data vector is approximately Gaussian:

$$-2\log\mathcal{L}_{\mathrm{HL}} \;=\; \sum_{\ell\ell'}\big[\boldsymbol{g}\,\boldsymbol{C}_\ell^{\mathrm{fid}}\big]_\ell^T M_{\ell\ell'}^{-1}\big[\boldsymbol{g}\,\boldsymbol{C}_{\ell'}^{\mathrm{fid}}\big]_{\ell'},$$

with $M_{\ell\ell'}$ estimated from simulations. Cut sky / coupling: $M$ also picks up the off-diagonal $\ell\ell'$ correlations.

# mHL — Schur complement of the partitioned Gaussian

BACKUP · WHY MHL IS EXACT (WITHIN HL)

Partition the data vector into auto and cross blocks, $X_\ell = (X_\ell^{\mathrm{a}}, X_\ell^{\mathrm{c}})$, with $M = \begin{pmatrix} M_{\mathrm{aa}} & M_{\mathrm{ac}} \\ M_{\mathrm{ca}} & M_{\mathrm{cc}} \end{pmatrix}$.

**Marginal** (drop autos):

$$p(X^{\mathrm{c}}) \;\propto\; \exp\!\Big[-\tfrac{1}{2}(X^{\mathrm{c}})^T M_{\mathrm{cc}}^{-1} X^{\mathrm{c}}\Big] \quad\Longrightarrow\quad \mathcal{L}_{\mathrm{marg}} \text{ uses } M_{\mathrm{cc}}.$$

**Conditional** (autos given crosses):

$$p(X^{\mathrm{a}} \,|\, X^{\mathrm{c}}) \;\propto\; \exp\!\Big[-\tfrac{1}{2}(X^{\mathrm{a}} - \bar{X}^{\mathrm{a}})^T \Sigma_{\mathrm{a|c}}^{-1} (X^{\mathrm{a}} - \bar{X}^{\mathrm{a}})\Big],$$

with $\bar{X}^{\mathrm{a}} = M_{\mathrm{ac}} M_{\mathrm{cc}}^{-1} X^{\mathrm{c}}$ and $\Sigma_{\mathrm{a|c}} = M_{\mathrm{aa}} - M_{\mathrm{ac}} M_{\mathrm{cc}}^{-1} M_{\mathrm{ca}}$ (Schur complement).

**Factorization:** $\mathcal{L}_{\mathrm{full}}(\theta) = \mathcal{L}_{\mathrm{marg}}(\theta) \cdot \mathcal{L}_{\mathrm{cond}}(\theta)$ — exact within HL, enables the auto-vs-cross diagnostic of Part V.

# Sellentin–Heavens correction

BACKUP · FINITE-$N_{\rm sim}$ COVARIANCE

$M_{\ell\ell'}$ is estimated from $N_{\rm sim}$ realisations $\Rightarrow$ it is a noisy Wishart-distributed matrix. Marginalising over the true covariance gives a $t$-distribution-like form:

$$-2\log\mathcal{L} \;=\; N_{\rm sim}\,\log\!\left[1+\tfrac{1}{N_{\rm sim}-1}\,X^T\widehat{M}^{-1}X\right].$$

Reduces to the Gaussian when $N_{\rm sim}\to\infty$; prevents posterior over-confidence for small $N_{\rm sim}$.

In ELiCA: $N_{\rm sim}=500$, $n_{\rm data}\approx 28\times 6=168$ — correction is non-negligible.

Sellentin & Heavens 2015 — arXiv:1511.05969

# $\ln R$, $\ln I$, $\ln S$ — formulas

BACKUP · SUSPICIOUSNESS

Bayes ratio for two datasets $D_1$, $D_2$ on a shared parameter $\theta$:

$$R \;=\; \frac{Z_{12}}{Z_1\, Z_2} \;=\; \frac{P(D_1, D_2)}{P(D_1)\, P(D_2)},$$

with evidences $Z_i = \int \mathcal{L}_i(\theta)\, \pi(\theta)\, \mathrm{d}\theta$. Information from likelihoods:

$$D_{\mathrm{KL},i} \;=\; \int P_i(\theta) \log \frac{P_i(\theta)}{\pi(\theta)}\, \mathrm{d}\theta, \qquad \ln I \;=\; D_{\mathrm{KL},1} + D_{\mathrm{KL},2} - D_{\mathrm{KL},12}.$$

Suspiciousness (prior-independent):

$$\boxed{\ln S \;=\; \ln R - \ln I}$$

Both $\ln R$ and $\ln I$ are estimated from nested sampling along with the evidences.

Marshall et al. 2006 · Handley & Lemos 2019

# Calibration — $d - 2\ln S \sim \chi^2_d$

BACKUP · TURNING $\ln S$ INTO A TENSION SCALE

For Gaussian posteriors of dimension $d$, the suspiciousness is calibrated by

$$d - 2\ln S \;\sim\; \chi^2_d.$$

$d =$ effective number of constrained shared parameters (Bayesian model dim.):

$$d \;=\; 2\left(\langle(\ln\mathcal{L})^2\rangle_P - \langle\ln\mathcal{L}\rangle_P^2\right).$$

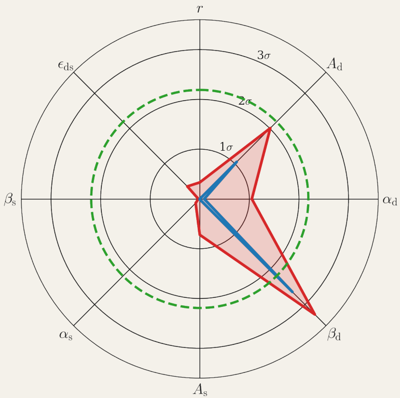
Tension probability:

$$p \;=\; \int_{d-2\ln S}^{\infty} \chi^2_d(x)\,\mathrm{d}x, \qquad n_\sigma = \sqrt{2}\,\mathrm{erfc}^{-1}(p).$$

Why the dimensionality: a high-dim. shift can have small $T$ per axis but a large total $\chi^2_d$ — averaging across $d$ dilutes signals (slide 46).

# The look-elsewhere effect (LEE)

BACKUP · CORRECTING THE PER-PARAMETER SHIFTS




Preliminary.

Scanning 8 parameters inflates the largest shift — the look-elsewhere effect; the trials factor folds it in.

- LEE-corrected here (vs uncorrected in the main flow): the ring moves, $\beta_{\rm dust}$ still stands out.
- Localizes to a single deep BK 150 GHz bandpower ($\ell \approx 178$); overall significance spans $\approx 2.2$–$2.8\sigma$ — reduced, not erased.

# Is it just that one bandpower?

BACKUP · ROBUSTNESS CHECKS · PRELIMINARY

- Drop the peak bandpower (BK 150 GHz, $\ell \approx 178$) and repeat the full fit:
  - Conditional $\beta_{\rm dust}$ moves back toward the cross, $0.86 \pm 0.22 \to 1.01 \pm 0.28$ (cross unchanged, $1.74 \pm 0.16$).
  - Shift drops $2.8\sigma \to 2.1\sigma$; omnibus suspiciousness $2.2\sigma \to 1.7\sigma$.
  - $r$ unaffected — the feature lives in the foreground sector, not the cosmology.
- So the single cell carries about half the effect; a coherent residual across the remaining BK auto-bandpowers carries the other half — localized, not pointlike.
- An untargeted omnibus $\chi^2$ over all 99 conditional residuals is unremarkable ($0.9\sigma$): the discrepancy is coherent and localized, not a broadband excess.

All BICEP results preliminary — Galloni & Ballardini, in prep.

# The non-Gaussian parameter-shift test

BACKUP · THE ACCURATE $\beta_{\rm dust}$ SHIFT

The Gaussian shift assumes both posteriors are Gaussian. The conditional $\beta_{\rm dust}$ is skewed and pressed against its lower bound (0.86 with a tail toward 0), so the Gaussian $\sigma$ overstates the separation. Form $\Delta\beta = \beta_{\rm cross} - \beta_{\rm auto}$ from the two independent posteriors; the tension is the chance there *is* a shift:

$$P_{\rm shift} = \int_{p(\Delta\beta)>p(0)} p(\Delta\beta)\,\mathrm{d}\Delta\beta, \qquad n_\sigma = \sqrt{2}\,\mathrm{erf}^{-1} P_{\rm shift}.$$

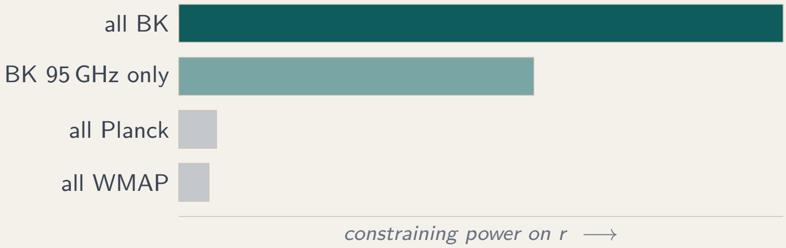
- *Gaussian* limit: $n_\sigma = |\Delta\mu| / \sqrt{\sigma_{\rm c}^2 + \sigma_{\rm a}^2} = 3.3\sigma$.
- *Non-Gaussian* (KDE from the samples, no Gaussian assumption): $P_{\rm shift} = 0.9952 \Rightarrow 2.82\sigma$ (stable 2.82/2.89 across resamplings) $\Rightarrow \approx 2.8\sigma$.
- All other 7 parameters $< 2\sigma$ by the same test — $\beta_{\rm dust}$-localized.

Raveri, Zacharegkas & Hu 2019 · `tensiometer`.

# Where does the $r$ information actually come from?

BACKUP · CHANNEL-BY-CHANNEL

Restrict the analysis to one channel block at a time:




*BK channels carry $r$; Planck/WMAP enter as foreground templates. Removing all BK gives $r = 0.68 \pm 0.34$.*

# Map-making — GLS and destriping

BACKUP · TOD $\to$ MAP, IN DETAIL

The Gaussian TOD likelihood for $\mathbf{d} = \mathbf{Pm} + \mathbf{n}$ is $-2\log\mathcal{L} = (\mathbf{d} - \mathbf{Pm})^\top \mathbf{N}^{-1}(\mathbf{d} - \mathbf{Pm})$. Setting $\partial/\partial\mathbf{m} = 0$ gives the normal equations

$$(\mathbf{P}^\top \mathbf{N}^{-1}\mathbf{P})\,\hat{\mathbf{m}} = \mathbf{P}^\top \mathbf{N}^{-1}\mathbf{d}.$$

- $\mathbf{N}^{-1}$ acts as a Fourier-domain noise filter (stationary $1/f$); the system is solved iteratively by preconditioned conjugate gradient.
- **Destriping** splits the noise $\mathbf{n} = \mathbf{Fa} + \mathbf{w}$ into baseline offsets $\mathbf{Fa}$ (the $1/f$ part) and white $\mathbf{w}$. ML for the amplitudes,

  $$\hat{\mathbf{a}} = \left(\mathbf{F}^\top \mathbf{Z}\mathbf{F} + \mathbf{C}_a^{-1}\right)^{-1}\mathbf{F}^\top \mathbf{Z}\mathbf{d}, \quad \mathbf{Z} = \mathbf{C}_w^{-1} - \mathbf{C}_w^{-1}\mathbf{P}(\mathbf{P}^\top \mathbf{C}_w^{-1}\mathbf{P})^{-1}\mathbf{P}^\top \mathbf{C}_w^{-1},$$

  then bin the cleaned stream. A baseline prior $\mathbf{C}_a$ regularizes the fit; the solution $\to$ full ML as the baseline length $\to$ one sample.

Keskitalo et al. 2005 (arXiv:astro-ph/0412517) · Biquard 2024.

# Map-making — transfer functions and $\mathbf{C}_N$

BACKUP · WHY LARGE SCALES AND MEMORY ARE HARD

- **Filter-and-bin.** A time-domain filter applied before binning biases the map; the bias is calibrated as an empirical *transfer function* from input-vs-filtered simulations (CLASS: a pixel-space transfer matrix). Cheap, but the large-scale correction is the delicate part.
- **The covariance wall.** The dense pixel–pixel $\mathbf{C}_N = (\mathbf{P}^\top \mathbf{N}^{-1} \mathbf{P})^{-1}$ is $\mathcal{O}(N_{\rm pix}^2)$ to store — infeasible at native resolution, and reconstructed only on low-resolution maps via Monte-Carlo / destriping-baseline estimation.
- **Scale.** Optimal map-making is memory/IO-bound, pushing to massively-parallel HPC and GPU/exascale.

Biquard 2024 — arXiv:2406.15415 · CLASS 2025 — arXiv:2501.11904.