Lessons Learned from Planck for cosmology from large angle polarization

Loris Colombo

“From Planck to the Future of CMB”
Ferrara, May 24th, 2022

Overview

- Planck represents the state-of-the-art for space CMB measurements.
- It mapped the microwave sky from L2, capturing ~29 Months of science data with the High Frequency Instruments, and ~49 Months of science data with the Low Frequency Instrument.
- Planck data analysis lasted 10+ years, with the main data and science results made available to the community in 4 public releases.
- Planck data will remain relevant for cosmology for years to come.

Planck PR1

- Planck first data release did not include any polarization information. Systematics and foreground contamination limited sensitivity to CMB polarization especially on large angular scales.

  "These factors are currently restricting Planck's ability to meet its most ambitious goals, e.g., to measure or set stringent upper limits on cosmological B-mode amplitudes. Although this situation is being improved at the present time, the possibility remains that these effects will be the final limitation for cosmology using the polarized Planck data." (Planck Collaboration I 2014)

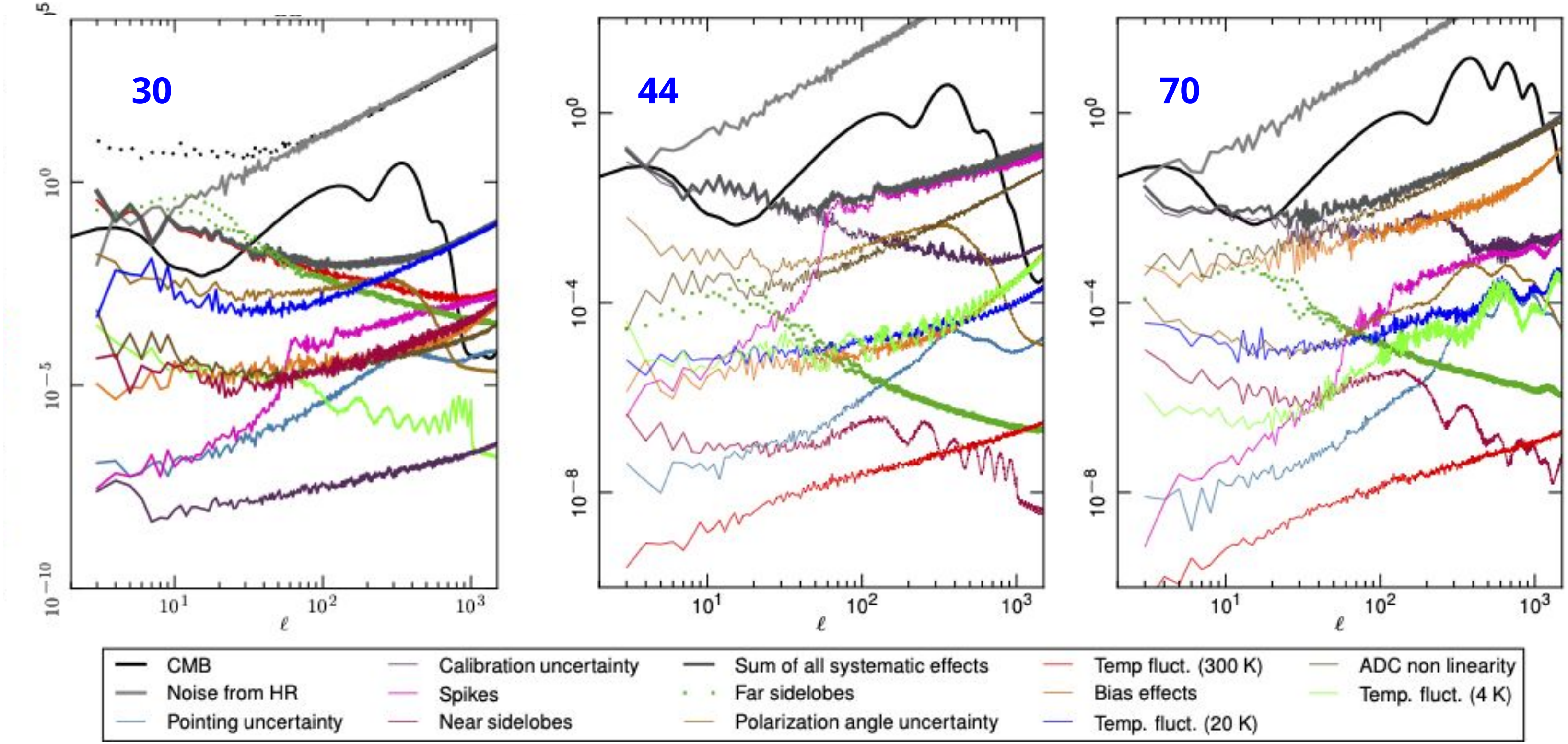
- The low-ℓ likelihood followed WMAP setup, but replacing the WMAP-based Commander temperature map with Planck-based Commander map.

- However, using Planck 353GHz channel to clean polarized dust contamination in WMAP data lead to a lower estimate, $\tau = 0.075 \pm 0.013$ vs $\tau = 0.089 \pm 0.013$ (Planck Collaboration XV 2014).

Planck PR2

- A combination of improved understanding of instrumental properties, refinement of PR1 pipeline, and the addition of polarization specific processing, allowed for a first release of Planck polarization data.
- The main residual systematics in LFI large angle polarization were calibration errors and ADC non-linearities, which were shown to be below instrumental noise level at 70GHz.
- ADC non-linearities were also the main residual systematics affecting HFI polarization, but the time constraints prevented their reduction to a level below the instrumental noise on the largest scales for the CMB channels.
- A decision was taken for a partial release of polarization data:
    - Full multipole range at 30, 44, 70 (minus Survey 2 and 4) and 353GHz.
    - High-pass filtered ($\ell > 10$) maps at 100, 143, 217GHz.
- This allowed to release a Planck-only polarization likelihood covering all multipoles, by combining LFI large scales with HFI measurements at higher multipoles.

PR2 LFI systematics

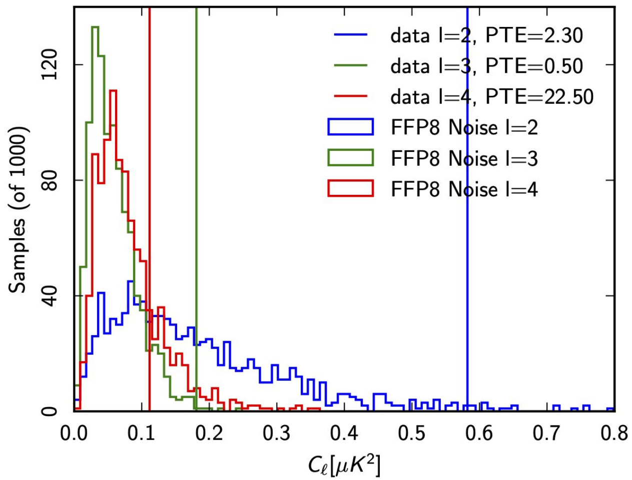


Planck Collaboration III 2016

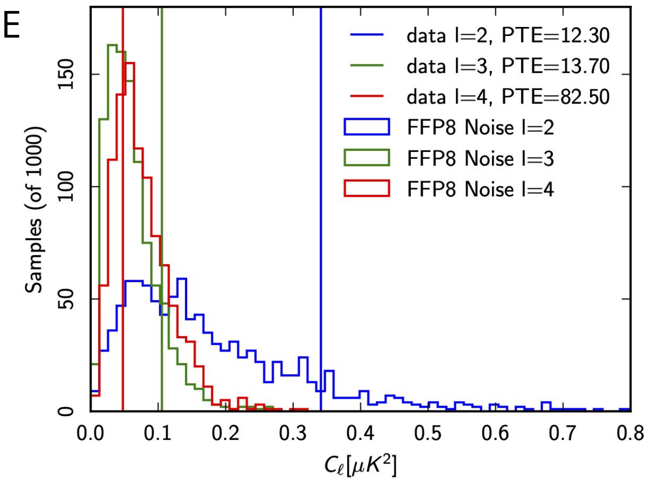
PR2 70GHz survey null tests

- Several null tests targeting S2 and S4 showed anomalies at the ~ 3σ level (as shown by a comparison with noise-only FFP8 simulations), and they were conservatively excluded from the cosmological analysis.

BB



EE



Planck Collaboration II 2016

PR2 Low-resolution dataset

- The low-ℓ P dataset was based on a template cleaned 70GHz map:

$$m_{Q,U} = \frac{1}{1-\alpha-\beta}(m_{70} - \alpha m_{30} - \beta m_{353})$$

- Scaling coefficients were estimated by minimizing the effective χ^2.

- The cleaned map was combined with a downgraded T Commander map, and the corresponding likelihood was assumed to be a multivariate Gaussian:

$$\mathcal{L}(\theta|\mathbf{m}) \propto P(\mathbf{D}=\mathbf{m}|\theta) \propto \frac{1}{\sqrt{|\mathbf{M}(\theta)|}} \exp\left(-\frac{1}{2}\mathbf{m}^t\mathbf{M}^{-1}(\theta)\mathbf{m}\right)$$

- with an effective (P) NCVM

$$N = \frac{1}{(1-\alpha-\beta)^2}(N_{70} + \sigma_\alpha^2 m_{30} m_{30}^T + \sigma_\beta^2 m_{353} m_{353}^T)$$

PR2 Low-resolution dataset

- The low-ℓ P dataset was based on a template cleaned 70GHz map:

$$m_{Q,U} = \frac{1}{1-\alpha-\beta}(m_{70} - \alpha m_{30} - \beta m_{353})$$

- Scaling coefficients were estimated by minimizing the effective χ^2.
- The cleaned map was combined with a downgraded T Commander map, and the corresponding likelihood was assumed to be a multivariate Gaussian:

$$\mathcal{L}(\theta|\mathbf{m}) \propto P(\mathbf{D}=\mathbf{m}|\theta) \propto \frac{1}{\sqrt{|\mathbf{M}(\theta)|}} \exp\left(-\frac{1}{2}\mathbf{m}^t\mathbf{M}^{-1}(\theta)\mathbf{m}\right)$$

- with an effective (P) NCVM

$$N = \frac{1}{(1-\alpha-\beta)^2}(N_{70} + \sigma_\alpha^2 m_{30} m_{30}^T + \sigma_\beta^2 m_{353} m_{353}^T)$$

Instrument

PR2 Low-resolution dataset

- The low-ℓ P dataset was based on a template cleaned 70GHz map:

$$m_{Q,U} = \frac{1}{1-\alpha-\beta}(m_{70} - \alpha m_{30} - \beta m_{353})$$

- Scaling coefficients were estimated by minimizing the effective χ^2.
- The cleaned map was combined with a downgraded T Commander map, and the corresponding likelihood was assumed to be a multivariate Gaussian:

$$\mathcal{L}(\theta|\mathbf{m}) \propto P(\mathbf{D}=\mathbf{m}|\theta) \propto \frac{1}{\sqrt{|\mathbf{M}(\theta)|}} \exp\left(-\frac{1}{2}\mathbf{m}^t\mathbf{M}^{-1}(\theta)\mathbf{m}\right)$$

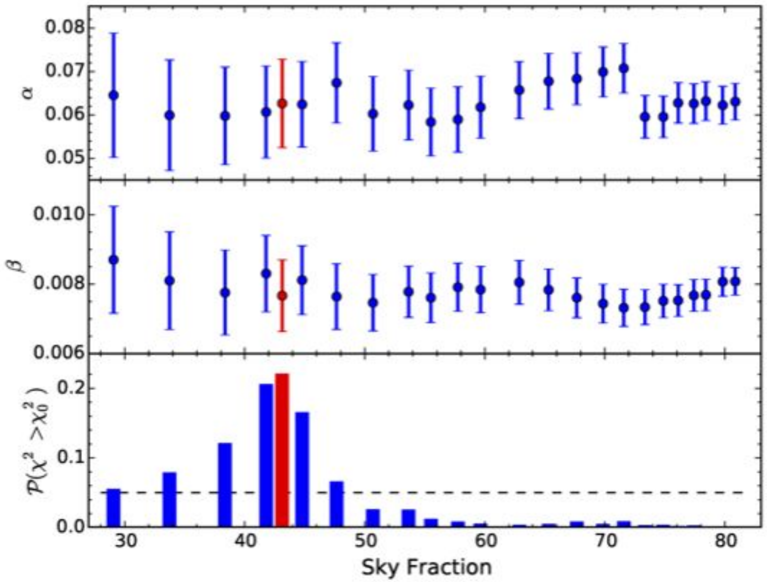
- with an effective (P) NCVM

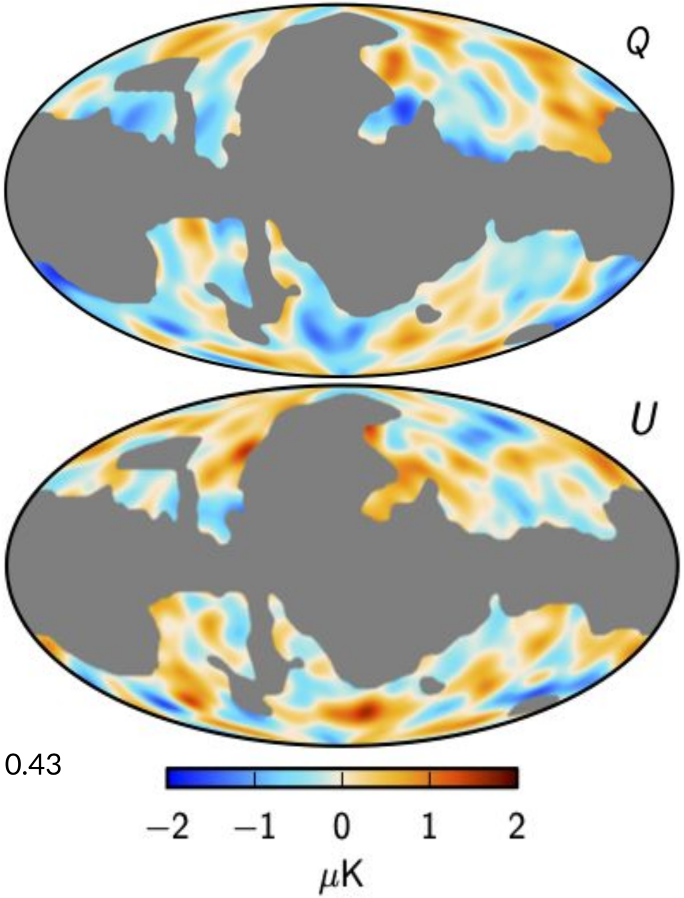
$$N = \frac{1}{(1-\alpha-\beta)^2}(N_{70} + \sigma_\alpha^2 m_{30} m_{30}^T + \sigma_\beta^2 m_{353} m_{353}^T)$$

Instrument

Foregrounds

PR2 70GHz cleaned map



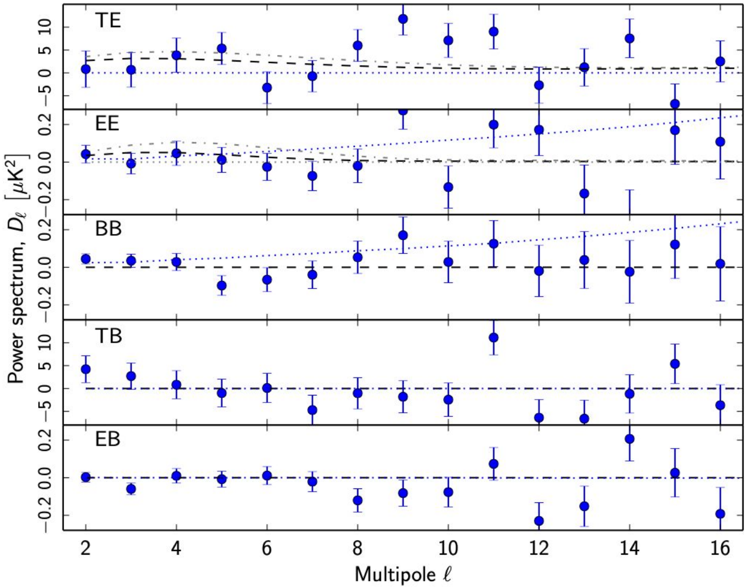


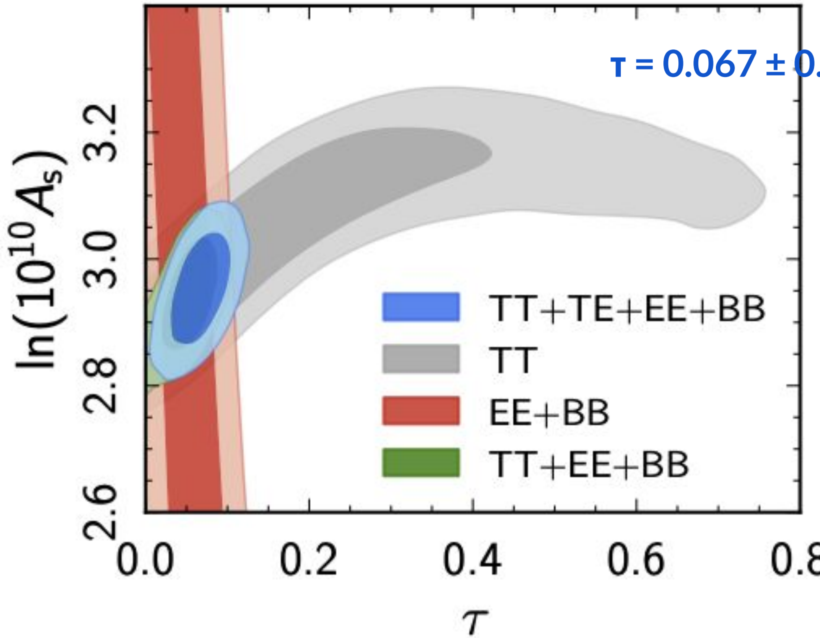
f_{sky} ~ 0.43

Planck Collaboration XIII 2016

PR2 Power Spectrum and Parameters

Planck Collaboration XIII 2016



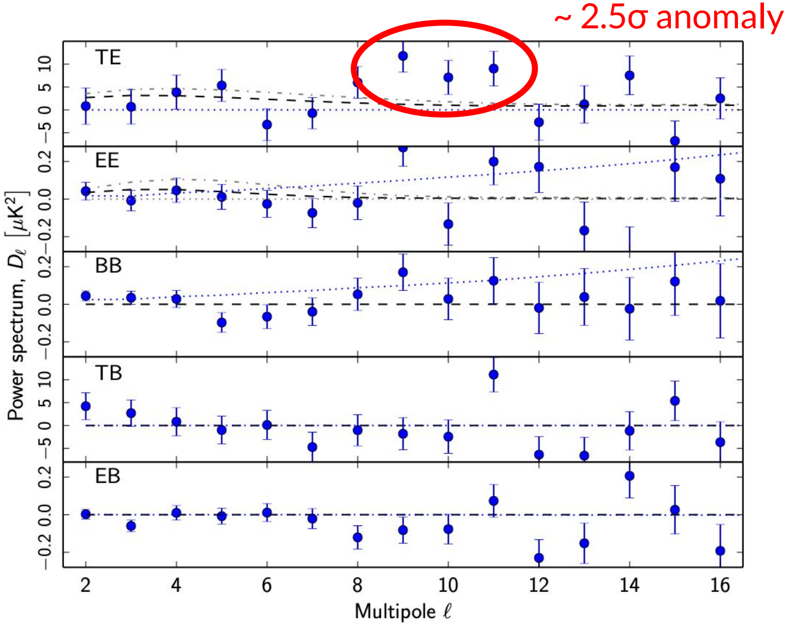


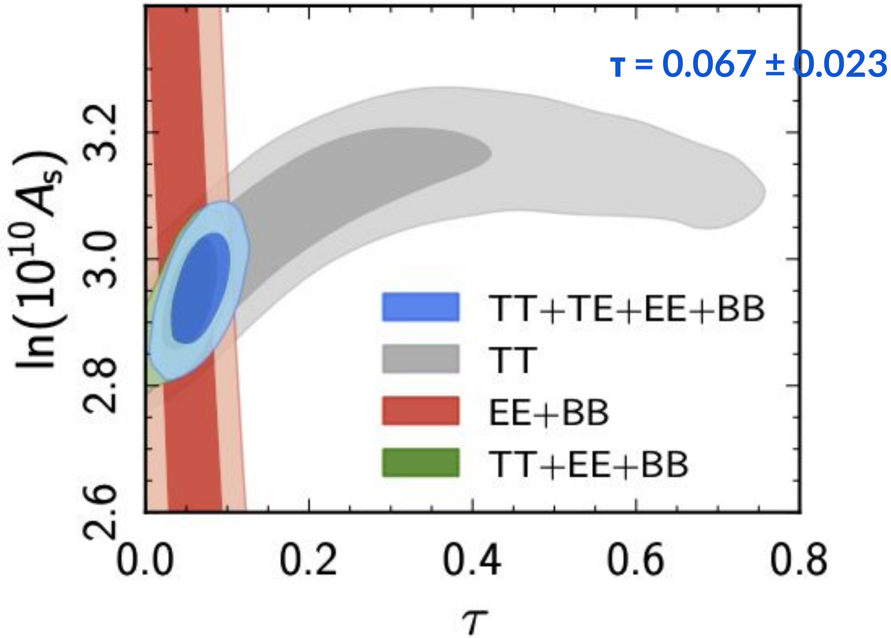
$\tau = 0.067 \pm 0.023$

Likelihood robustness was tested against data (e.g. null-tests) and simulation (e.g. harmonic χ^2) based checks

PR2 Power Spectrum and Parameters

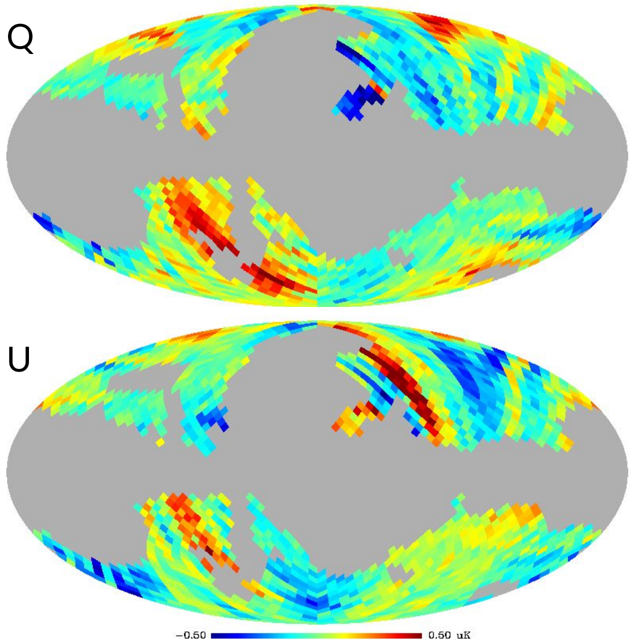
Planck Collaboration XIII 2016



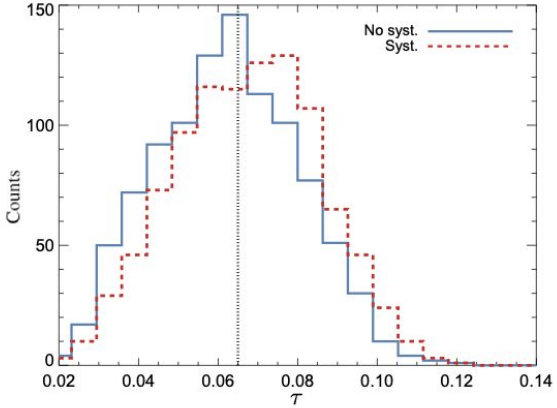


Likelihood robustness was tested against data (e.g. null-tests) and simulation (e.g. harmonic χ^2) based checks

LFI systematics error budget



- Simulation-based systematics template for cleaned 70GHz. Includes rescaled contribution from 30 and 353.
- Dominant effect is calibration uncertainty.
- Impact on τ estimated at ≲ 0.25σ.
- Similar test based on data-driven simulation templates, (e.g. det-set difference map), suggested lower bias ≲ 0.10σ



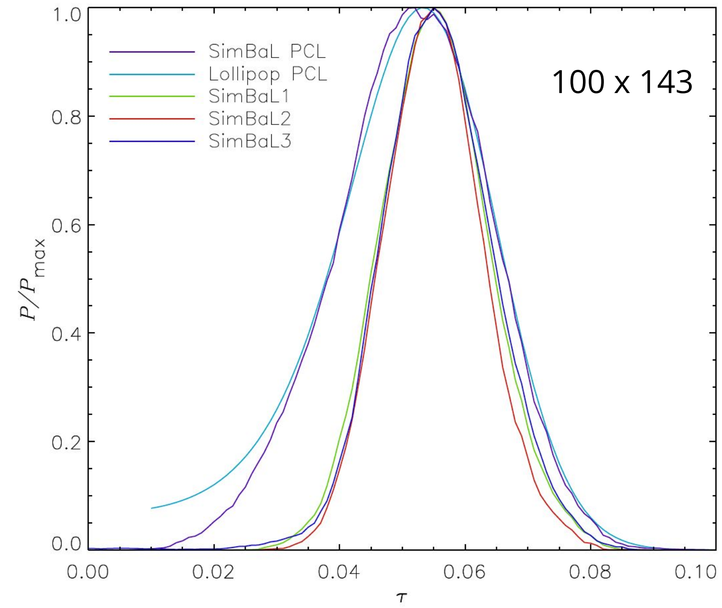
Planck Collaboration XLVI 2016

Intermediate results

- The main systematics affecting HFI large angle polarization was ADC non-linearities.
- In PR1 these were accounted for in terms of variable gain, which was enough for T analysis.
- A better model of ADC non-linearities was implemented for PR2, which allowed removal of the leading order effect, but leaving first order contributions affecting the largest polarization scales.
- SRoll took advantage of the refined ADC model and Planck foreground results to create templates of residual systematics in order to fit for and remove them during mapmaking, allowing to use HFI data also for low-ℓ cosmology.

Simulation based likelihood

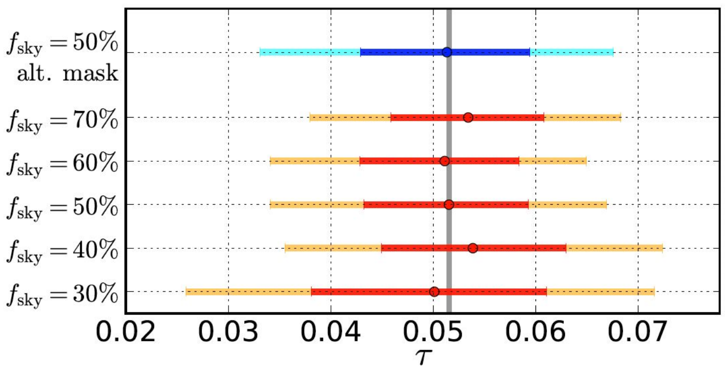
- Cosmological inference from SRoll maps was based on cross-spectra between foreground cleaned maps at different frequencies, reducing the impact of possible biases due to modeling errors in noise or systematics properties.
- SRoll maps were supported by a set of E2E noise+instrumental effects simulations, which allowed to empirically determine the statistical properties of cross-spectra distribution, and to propagate the relevant uncertainty to the cosmological parameters.
- τ estimates from HFIxHFI and HFIxLFI cross-spectra were in good agreement with PR2 LFI constraints.

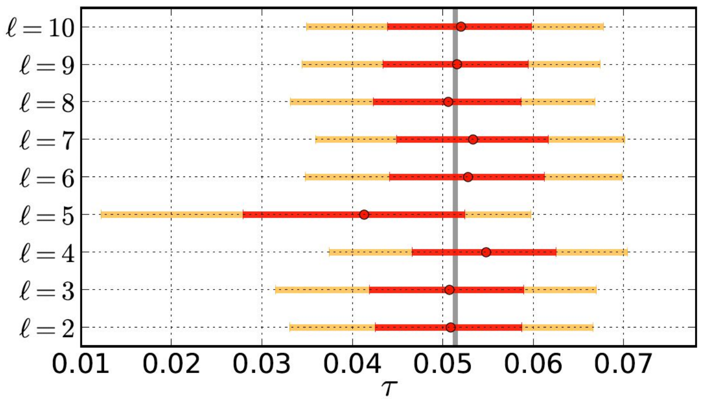


Planck Collaboration Int. XLVI 2016

Planck PR3

- SRoll approach was adopted as default HFI mapmaking, and the corresponding likelihood based on 100 and 143GHz foreground cleaned maps provided the baseline P low-ℓ approach, in combination with a Blackwell-Rao T likelihood, based on Commander map.
- Extensive sets of robustness test, checking for dependence from sky fraction, multipole range, simulation sets, data cuts, ...
- τ = 0.051 ± 0.009

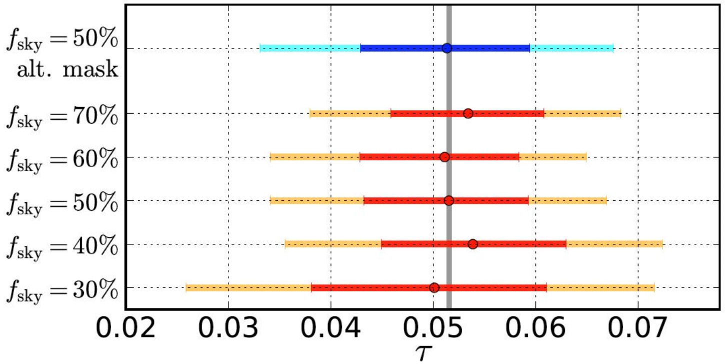


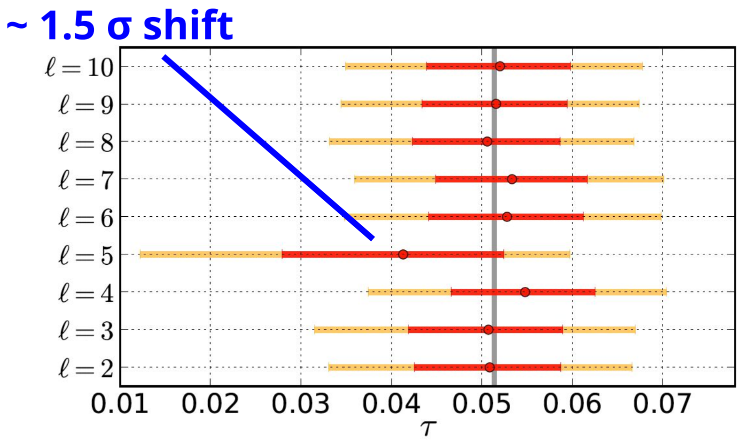


Planck Collaboration V 2020

Planck PR3

- SRoll approach was adopted as default HFI mapmaking, and the corresponding likelihood based on 100 and 143GHz foreground cleaned maps provided the baseline P low-ℓ approach, in combination with a Blackwell-Rao T likelihood, based on Commander map.
- Extensive sets of robustness test, checking for dependence from sky fraction, multipole range, simulation sets, data cuts, …
- $\tau = 0.051 \pm 0.009$





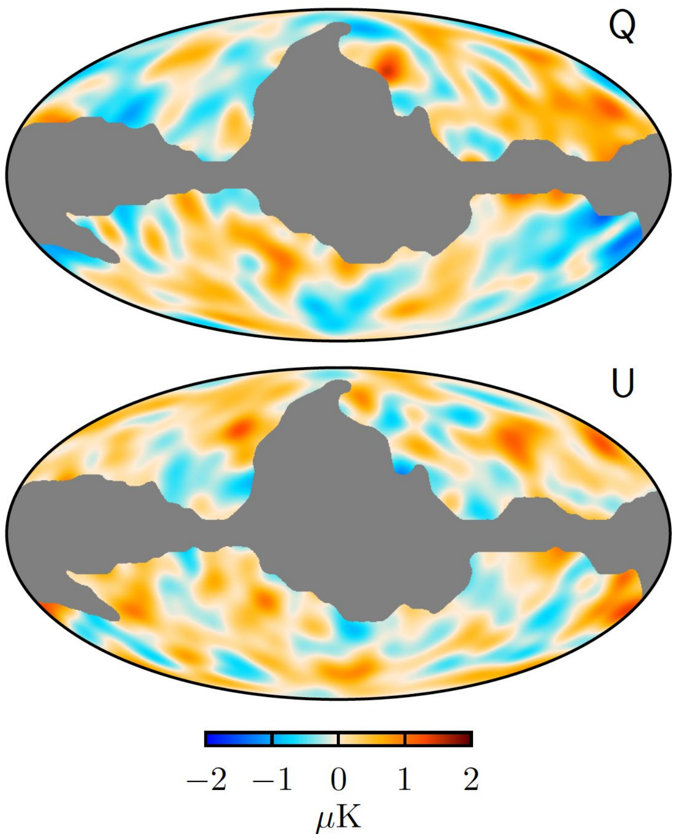
Planck Collaboration V 2020

PR3 LFI Calibration Changes

- In regions of the sky where the dipole is ~0, the dominant timeline contribution at 30 and 44 GHz are the Galactic foregrounds.
- The intrinsic fg polarization caused spurious anti-correlated fluctuations in the relative gain between two radiometers of the same horns.
- A circular problem: in order to measure the sky we need to understand the instrument, but this requires knowing the sky.
- LFI adopted an iterative pipeline, in which the polarized foreground solution from a previous step was fed as input to the next calibration and mapmaking iteration.
- The procedure converged at 30 and 44GHz, but not at 70GHz, where the polarized fg signal was lower.

LFI Gain leakage template

- A leakage template based on the difference between subsequent 30GHz iterations was fitted and removed from 70GHz data.
- Combined with other improvements, this allowed to increase sky fraction to ~ 0.63 and include all 8 surveys in the final pixel-based likelihood.
- $\tau = 0.063 \pm 0.020$



Post Legacy Analysis

- Three main efforts built on the first 3 releases to further improve Planck results:
  - Joint processing of HFI and LFI data, which became the base for Planck PR4 release (Reijo Keskitalo);
  - SRoll 2, focused on HFI data;
  - BeyondPlanck, a reanalisys of LFI data in a Bayesian framework (Mathew Galloway, Trygve Svalheim' poster).
- All these approaches adopt some level of "iterative" processing

BeyondPlanck

- BeyondPlanck is in a sense a natural continuation and the end point of the iterative LFI processing started in PR3.
- The starting point is a parametric model for the observed timeline:

$$d_{j,t} = g_{j,t}\mathsf{P}_{tp,j}\left[\mathsf{B}^{\mathrm{symm}}_{pp',j}\sum_{c}\mathsf{M}_{cj}(\beta_{p'},\Delta^{j}_{\mathrm{bp}})a^{c}_{p'} + \mathsf{B}^{\mathrm{asymm}}_{j,t}\left(\boldsymbol{s}^{\mathrm{orb}}_{j} + \boldsymbol{s}^{\mathrm{fsl}}_{t}\right)\right] + n^{\mathrm{corr}}_{j,t} + n^{\mathrm{w}}_{j,t}.$$

- With ω = {all parameters of interest}, map out the joint posterior with standard MCMC methods (Gibbs sampling).

$$P(\omega \mid \boldsymbol{d}) = \frac{P(\boldsymbol{d} \mid \omega)P(\omega)}{P(\boldsymbol{d})} \propto \mathcal{L}(\omega)P(\omega),$$

- The output of the pipeline is a set of samples for all the parameters, e.g. gain, correlated noise, foreground spectral properties, component maps,...

BeyondPlanck CMB processing

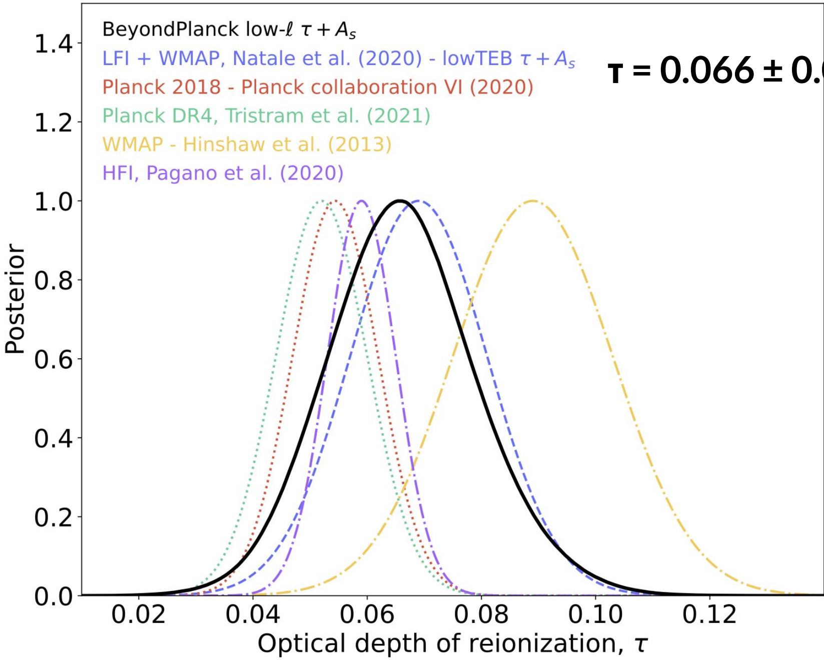
- For computational reasons, CMB extraction in BeyondPlanck is a two-step procedure.
- The role of the CMB component within the main BP chain is to get a robust foreground emission model while optimizing the computational resource use. These maps however have artifacts and residuals which renders them unsuitable for cosmology.
- The CMB maps used for the final science are a postprocessing of the maps from the main chain. There are two resampling steps:
    - Full resolution T-only maps.
    - Low-resolution I,Q,U maps.

Low resolution CMB products

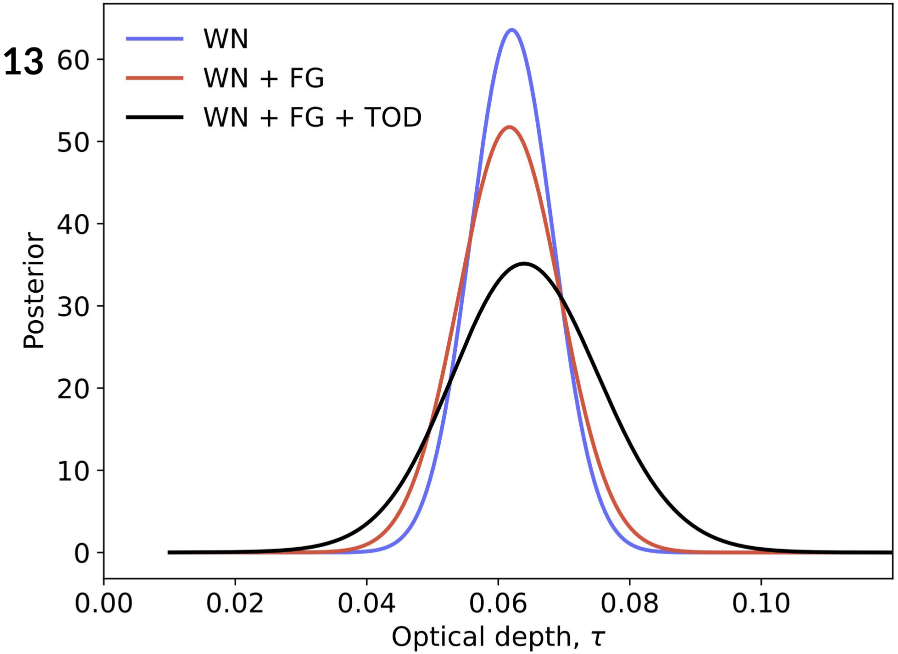
- For each main chain sample, BP resamples $\mathbf{a}^{\mathrm{CMB}}$ at multipoles $\ell \leq 64$, fixing higher multipoles and all instrumental and foreground parameters.
- Under some assumptions, this amounts to drawing samples from a Gaussian distribution with diagonal covariance (white noise), and can be done (comparatively) quickly.
  - For each main chain sample (4000) we draw 50 additional low-resolution samples.
- These samples define a mean CMB map and effective noise covariance matrix, which includes all uncertainty contributions from all steps of the pipeline.
- The corresponding likelihood has a multivariate Gaussian expression.

BeyondPlanck Collaboration 2022; Paradiso et al. 2022

BP low-l constraints



τ = 0.066 ± 0.013



- Good consistency with previous results.
- BP approach naturally allows to incorporate and fully propagate all sources of uncertainty to the final cosmological parameters.

Paradiso et al. 2022 https://arxiv.org/abs/2205.10104

Summary

- Planck highlighted the importance of a "holistic" approach to CMB data analysis. The interplay between the sky, the instrument, and the analysis pipeline will be a fundamental challenge for future generations of CMB experiments, and likely set the limit to our ability of constraining cosmology.
- Simulations will play a key role in understanding the instrument, assessing systematics and propagating them to final parameter estimates.