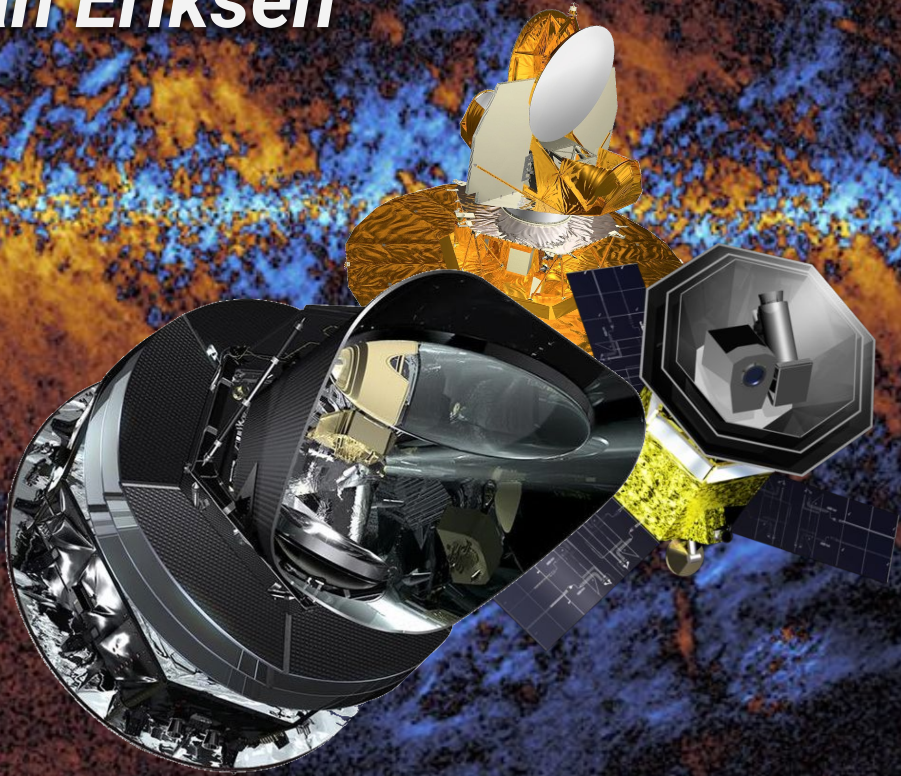
Systematics and degeneracies in CMB observations and the importance of global analysis

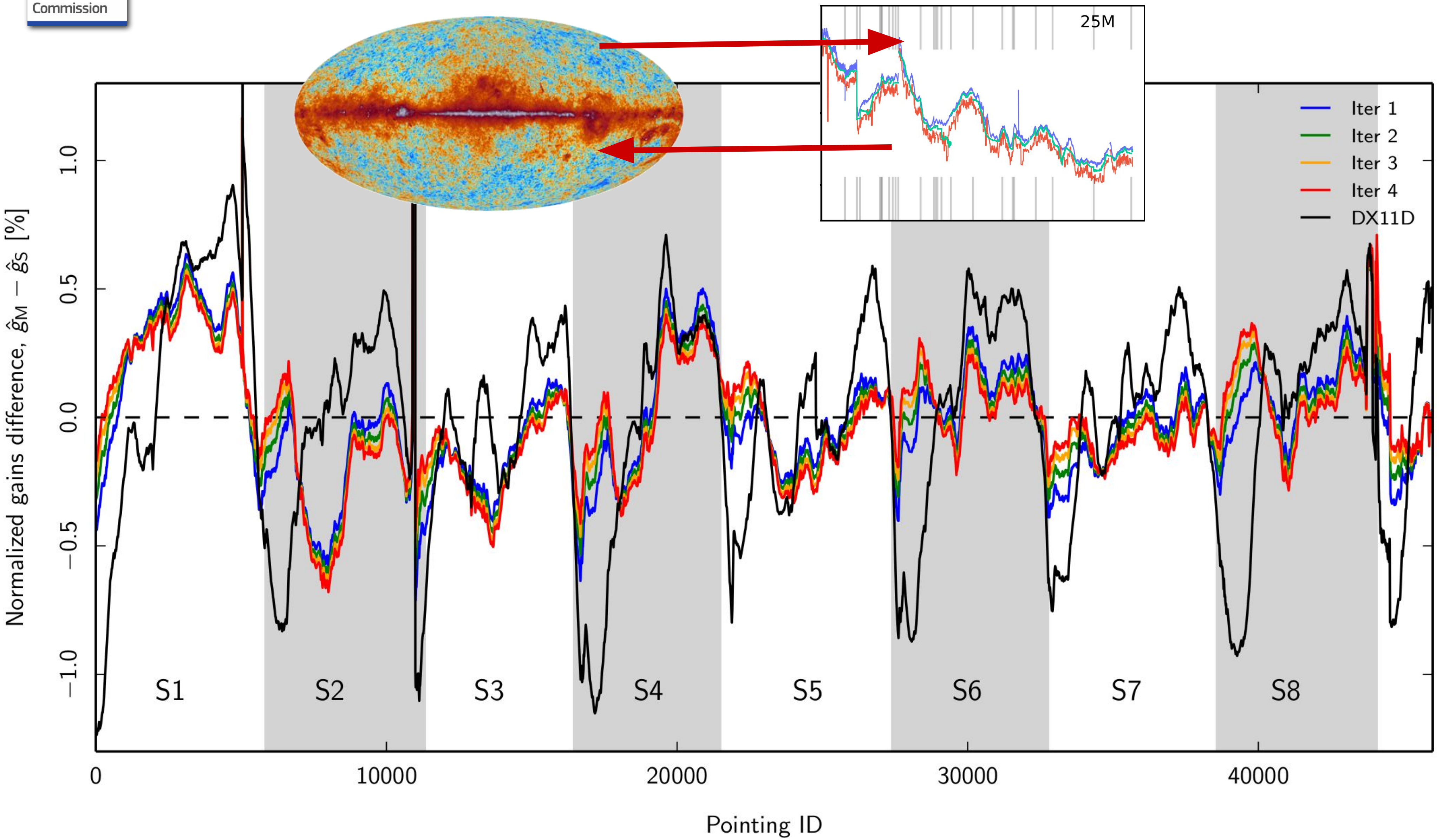
Hans Kristian Eriksen



"CMB systematics and calibration focus workshop", November 30, 2020

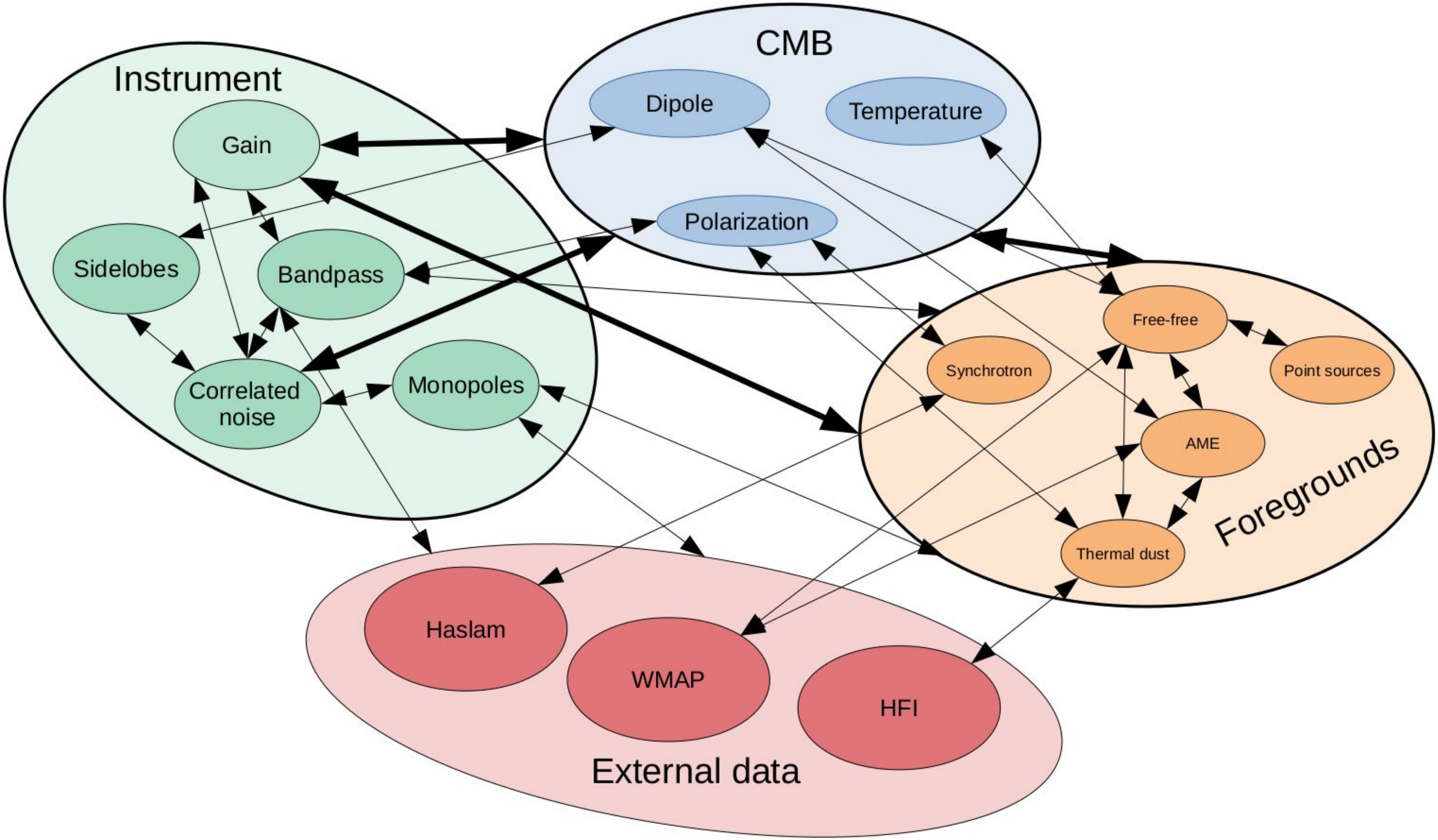
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776282

Critical question: How well do we really know the gain?



Planck (2018), A&A, 641, A2

"Planck LFI dependency map"



"The swamp"

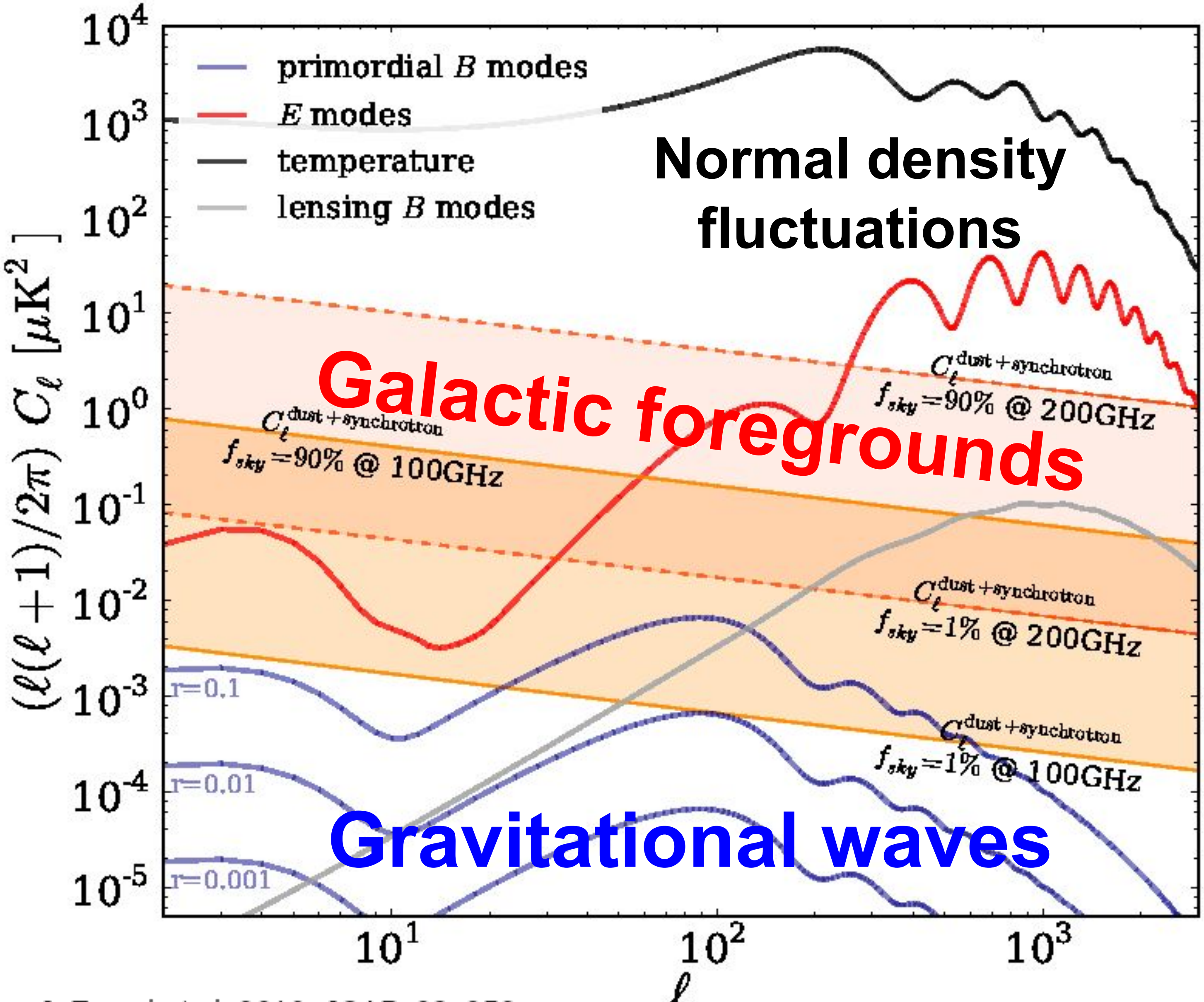


Figure from J. Errard et al. 2016, JCAP, 03, 052

“The swamp”

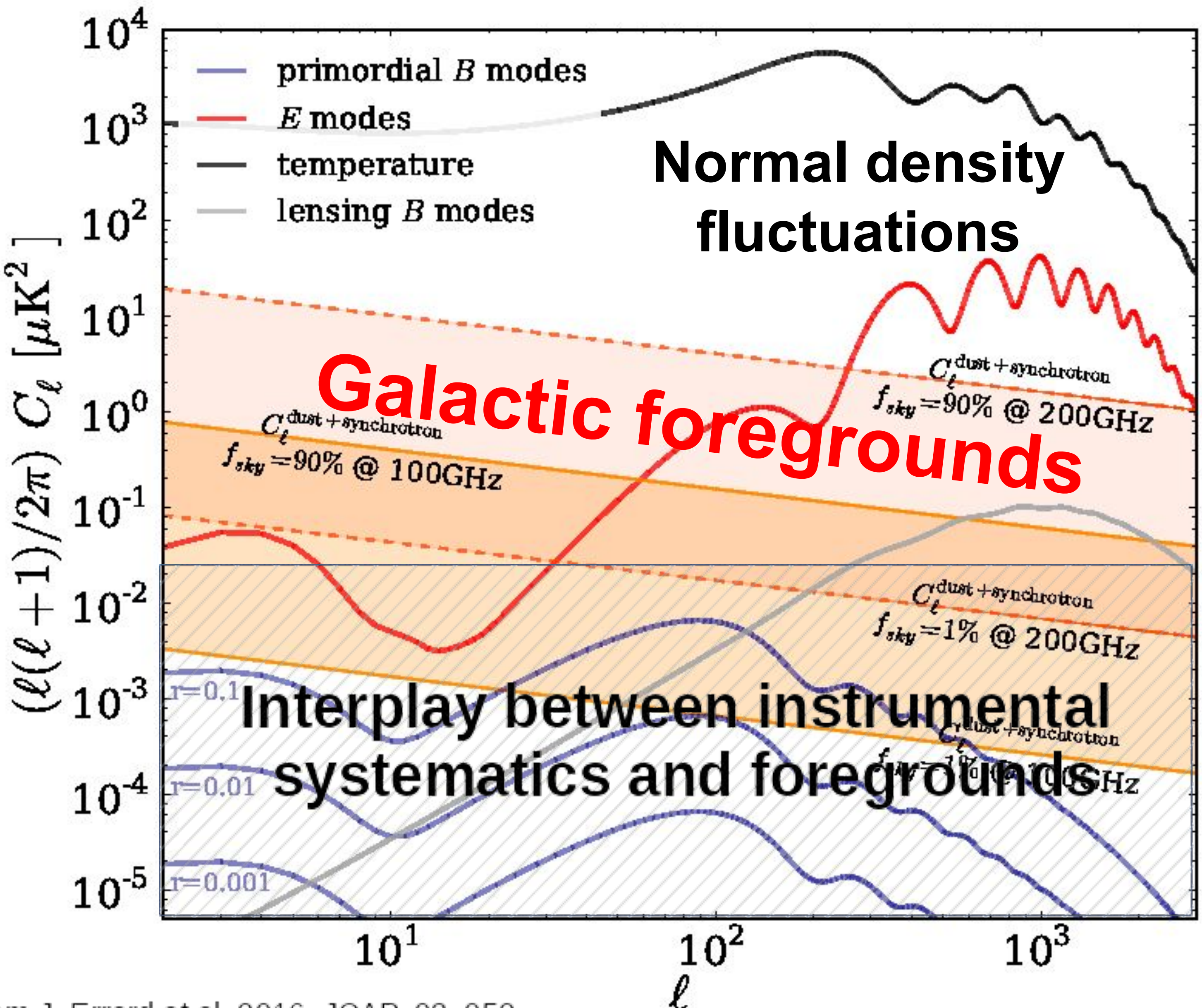
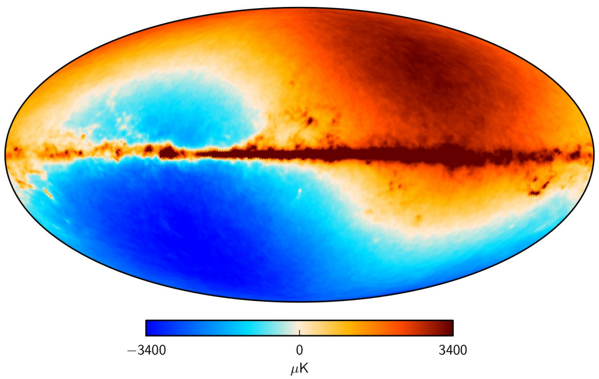


Figure from J. Errard et al. 2016, JCAP, 03, 052

CMB's "chicken and egg" problem

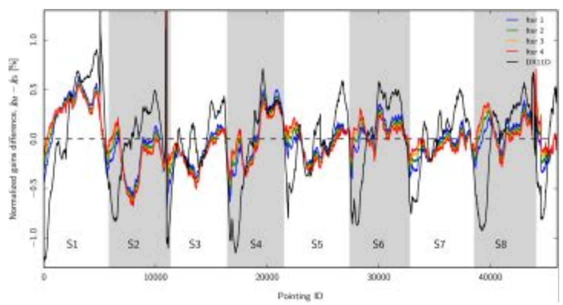
Need to know the instrument to measure the sky...

Data



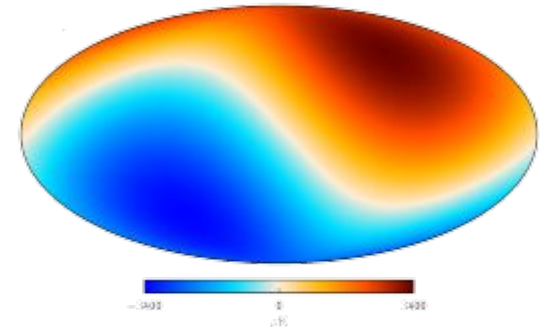
=

Instrument calibration

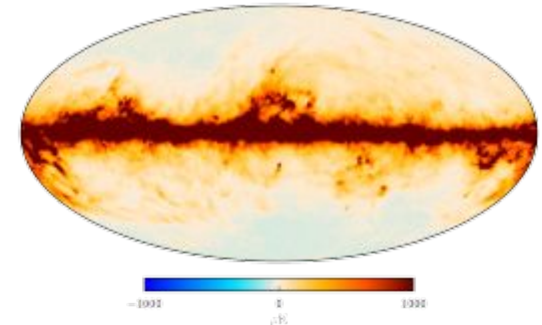


·

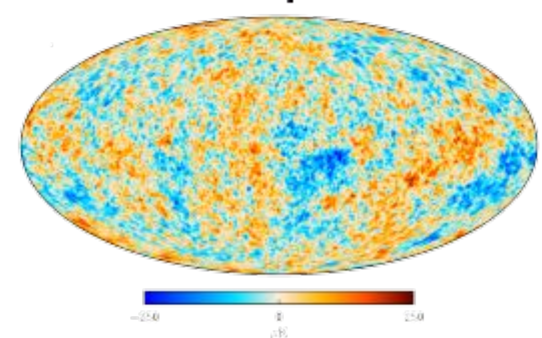
Sky



+



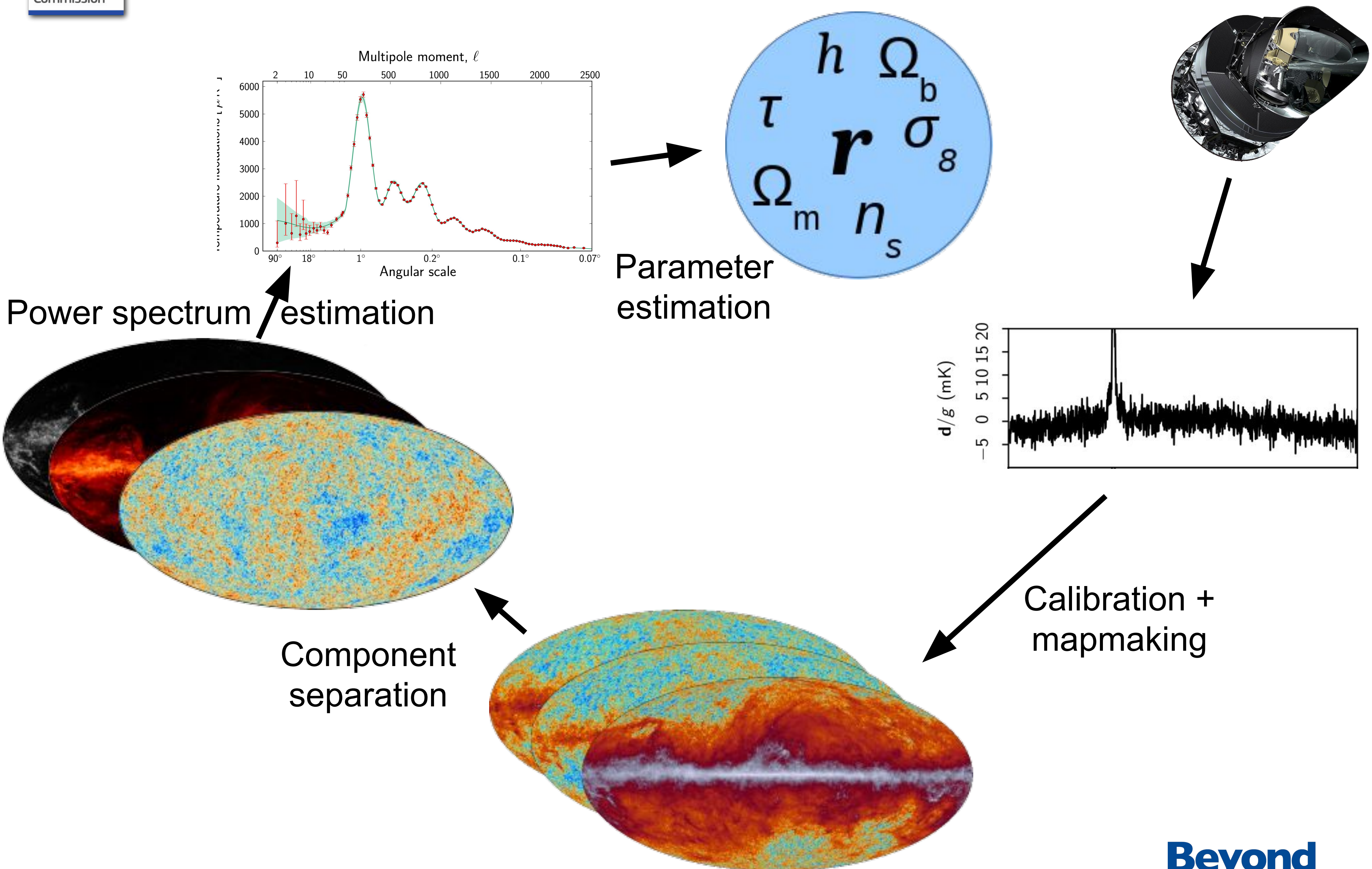
+



… but also need to know the sky in order to calibrate the instrument!

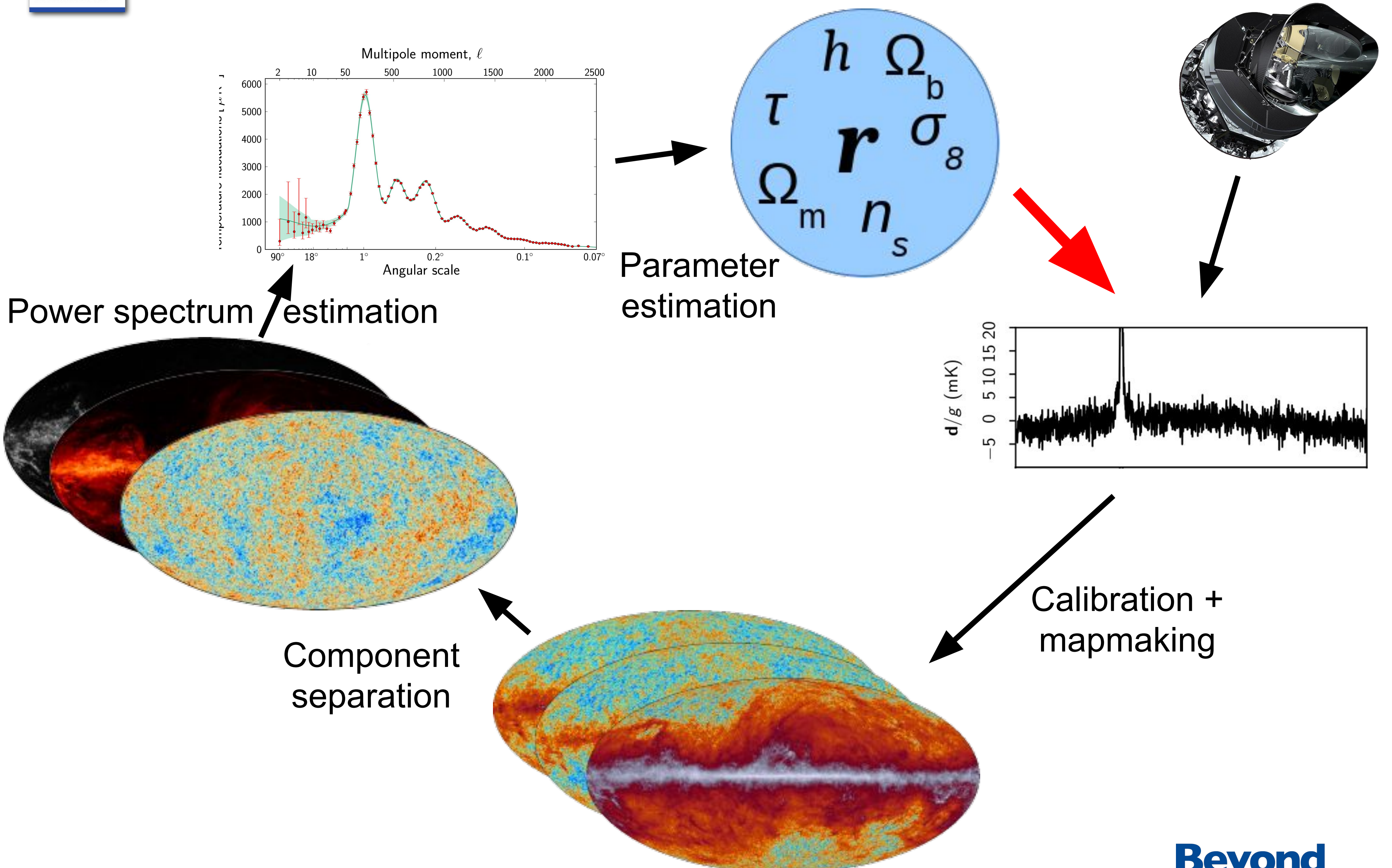
European Commission

Classic CMB analysis

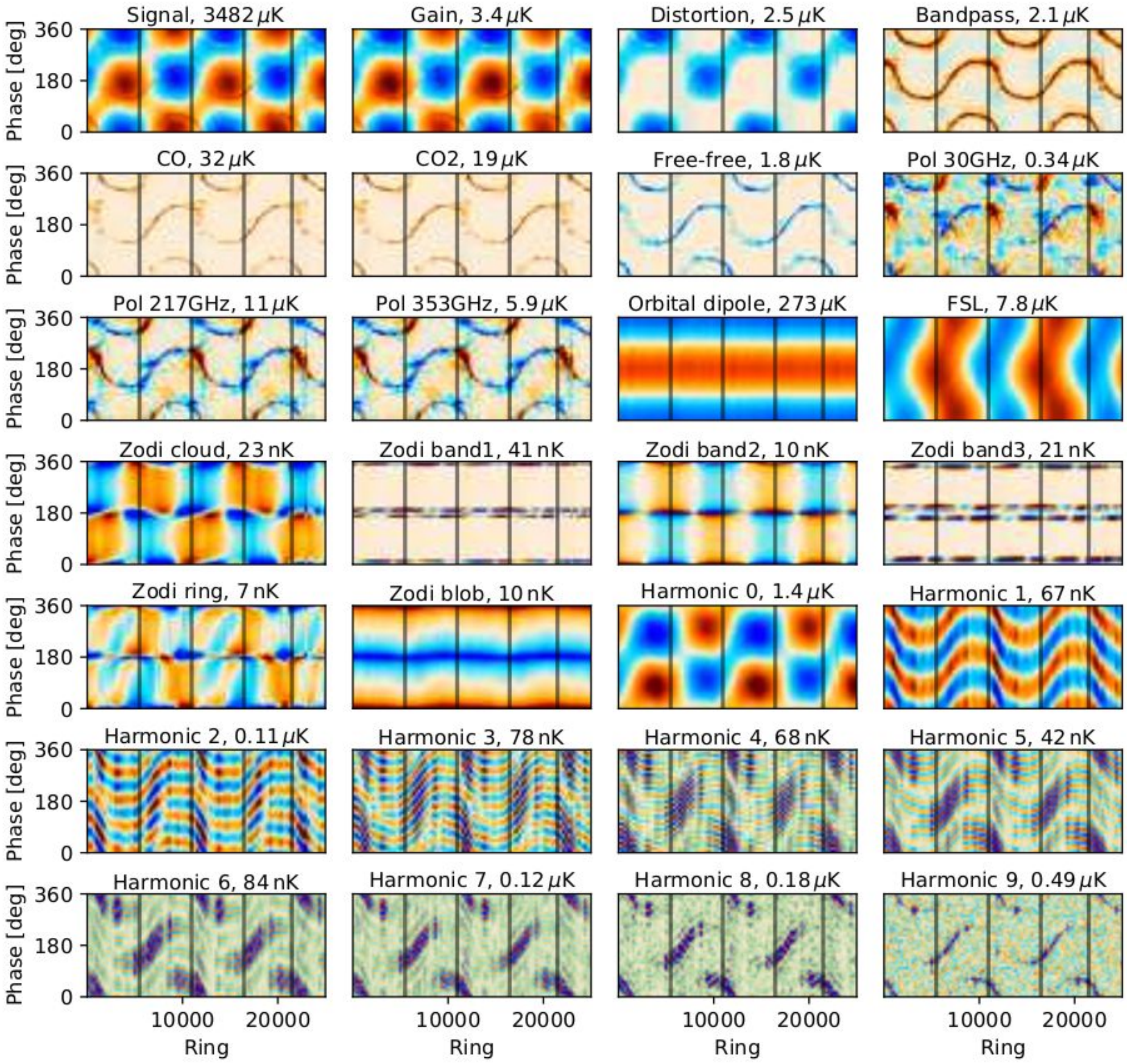


Beyond PLANCK

End-to-end iterative analysis



NPIPE (Planck PR4) and SROLL2



NPIPE = Joint low-level calibration and mapmaking through generalized linear regression of TOD templates

SROLL2 (Delouis et al. 2019) is a similar initiative, adopting a different gain+ADC model

Planck Int. LVII (2020), A&A, 643, A42

Starting point for BeyondPlanck

Could we solve some of the outstanding issues with Planck LFI 2018 by:

1. speeding up the iteration process, and perform hundreds of component separation + calibration iterations, not just four?

2. break internal Planck-specific degeneracies using external data, in particular WMAP?

The name BeyondPlanck was chosen to

- recognize that this work builds on, and is a natural continuation of, the official Planck analysis effort
- emphasize that this involves not only Planck, but also other data sets

Main goals for BeyondPlanck

Main goals of the BeyondPlanck project:

- Implement an end-to-end analysis framework for current and future CMB experiments using Planck experience
- Demonstrate this framework with Planck LFI data
- Make software and results publicly available under an OpenSource license

The BeyondPlanck pipeline in one slide

1. Write down an explicit parametric model for the observed data:

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

Let ω = {all free parameters}

2. Derive the joint posterior distribution with Bayes' theorem:

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

3. Map out $P(\omega \mid \boldsymbol{d})$ with standard Markov Chain Monte Carlo (MCMC) methods

The BeyondPlanck data model

European Commission

Data

Pointing

Bandpass

Sidelobe pickup

White noise

$d_{j,t} =$ ... ${}^{\mathrm{w}}_{j,t}$.

$$\omega \equiv \{g, \Delta_{\mathrm{bp}}, \boldsymbol{n}_{\mathrm{corr}}, \boldsymbol{a}_i, \beta_i, C_\ell, \ldots\}$$

$+ \sum_{j=1} \boldsymbol{a}^{j}_{\mathrm{src}} \left(\frac{}{\nu_{0,\mathrm{src}}} \right)$

Point sources

Beyond PLANCK

The posterior distribution

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

Gaussian noise

$$\mathcal{L}(\omega) = \frac{e^{-\frac{1}{2}(\boldsymbol{d}-\boldsymbol{s}(\omega))^t \mathrm{N}_{\mathrm{wn}}^{-1}(\boldsymbol{d}-\boldsymbol{s}(\omega))}}{\sqrt{|\mathrm{N}_{\mathrm{wn}}|}}$$

- $P(f_{\mathrm{knee}})$ = lognorm(DPC, 0.1)
- $P(\beta_{\mathrm{synch}})$ = -3.1 ± 0.1
- $P(T_{\mathrm{dust}})$ = $\delta(T_{\mathrm{dust}} - T_{\mathrm{dust,\,HFI}})$
- $P(a_{\mathrm{ff}})$ = $N(a_{\mathrm{ff,Planck}}, \sigma^2_{l,\mathrm{ff}})$
- $P(a_{\mathrm{ame}})$ = $N(\alpha \cdot m_{857}, \sigma^2_{l,\mathrm{ame}})$

⋮

How to sample from *big* distributions?

European Commission

Not logged in Talk Contributions Create account Log in

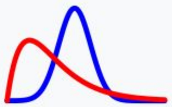
Article Talk Read Edit View history Search Wikipedia

Gibbs sampling

From Wikipedia, the free encyclopedia

In statistics, **Gibbs sampling** or a **Gibbs sampler** is a Markov chain Monte Carlo (MCMC) algorithm for obtaining a sequence of observations which are approximated from a specified multivariate probability distribution, when direct sampling is difficult. This sequence can be used to approximate the joint distribution (e.g., to generate a histogram of the distribution); to approximate the marginal distribution of one of the variables, or some subset of the variables (for example, the unknown parameters or latent variables); or to compute an integral (such as the expected value of one of the variables). Typically, some of the variables correspond to observations whose values are known, and hence do not need to be sampled.

Part of a series on
Bayesian statistics



Theory
Admissible decision rule · Bayesian efficiency · Bayesian probability · Probability interpretations · Bayes' theorem · Bayes factor · Bayesian inference · Bayesian network · Prior · Posterior · Likelihood · Conjugate prior · Posterior predictive · Hyperparameter · Hyperprior ·

WIKIPEDIA
The Free Encyclopedia

- Main page
- Contents
- Current events
- Random article
- About Wikipedia
- Contact us
- Donate

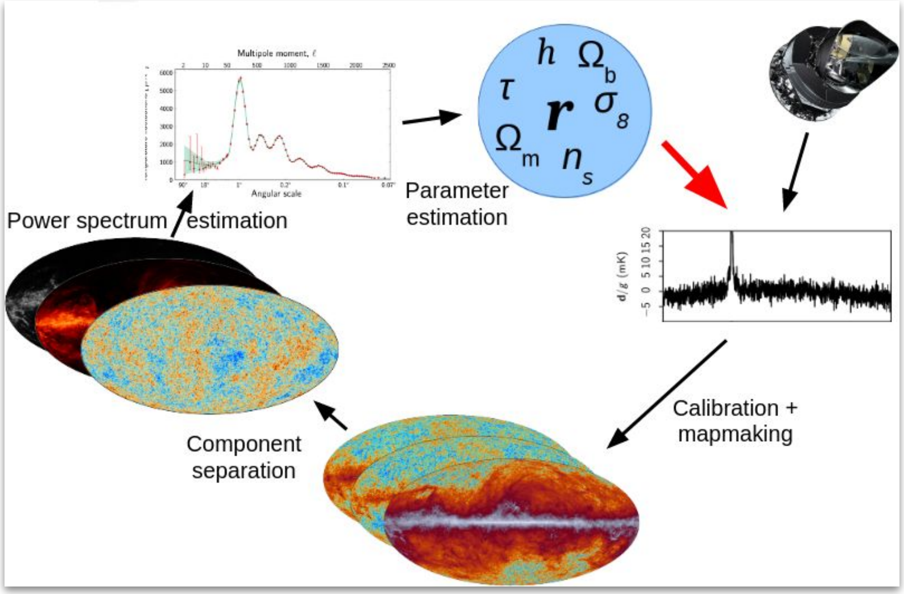
Contribute

- Help
- Learn to edit
- Community portal
- Recent changes
- Upload file

Beyond PLANCK

The BeyondPlanck Gibbs sampler

What we want to do:

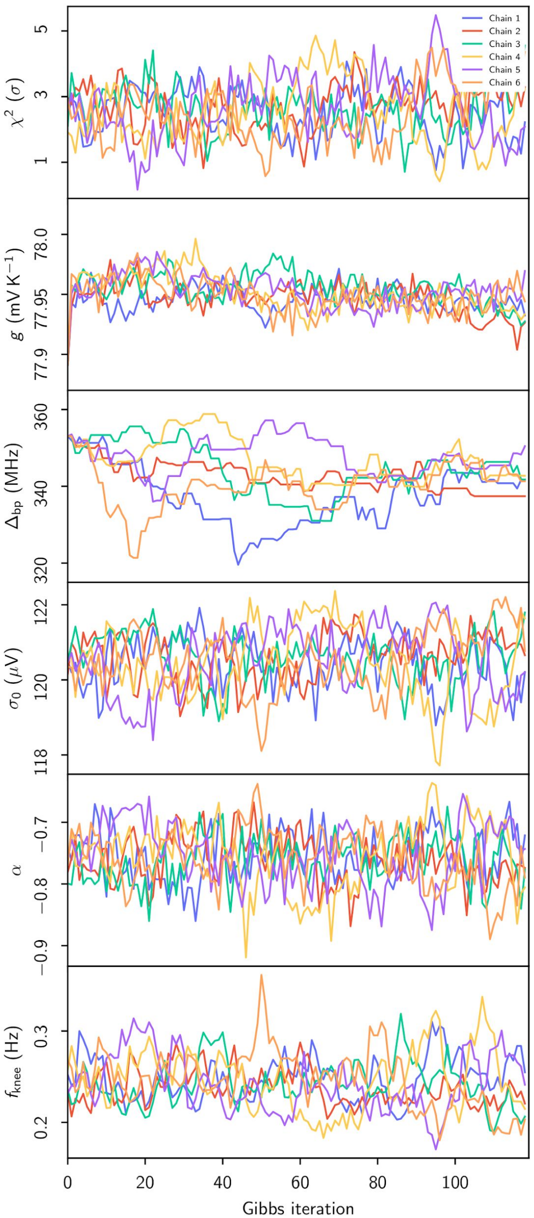


How we actually do it:

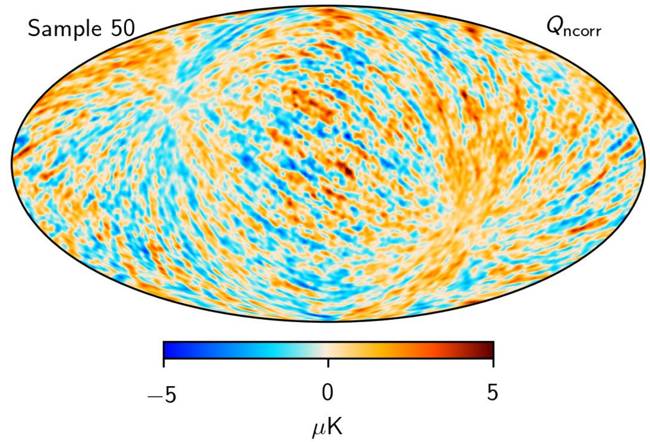
$$
\begin{array}{rcll}
\boldsymbol{g} & \leftarrow & P(\boldsymbol{g} & \mid \boldsymbol{d}, \qquad\qquad\qquad \xi_n, \Delta_{\mathrm{bp}}, \boldsymbol{a}, \beta, C_\ell) \\
\boldsymbol{n}_{\mathrm{corr}} & \leftarrow & P(\boldsymbol{n}_{\mathrm{corr}} & \mid \boldsymbol{d}, \boldsymbol{g}, \qquad\qquad \xi_n, \Delta_{\mathrm{bp}}, \boldsymbol{a}, \beta, C_\ell) \\
\xi_n & \leftarrow & P(\xi_n & \mid \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \qquad \Delta_{\mathrm{bp}}, \boldsymbol{a}, \beta, C_\ell) \\
\Delta_{\mathrm{bp}} & \leftarrow & P(\Delta_{\mathrm{bp}} & \mid \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \xi_n, \qquad \boldsymbol{a}, \beta, C_\ell) \\
\beta & \leftarrow & P(\beta & \mid \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \xi_n, \Delta_{\mathrm{bp}}, \qquad C_\ell) \\
\boldsymbol{a} & \leftarrow & P(\boldsymbol{a} & \mid \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \xi_n, \Delta_{\mathrm{bp}}, \quad \beta, C_\ell) \\
C_\ell & \leftarrow & P(C_\ell & \mid \boldsymbol{d}, \boldsymbol{g}, \boldsymbol{n}_{\mathrm{corr}}, \xi_n, \Delta_{\mathrm{bp}}, \boldsymbol{a}, \beta \qquad)
\end{array}
$$

Main product: Ensemble of full sample sets

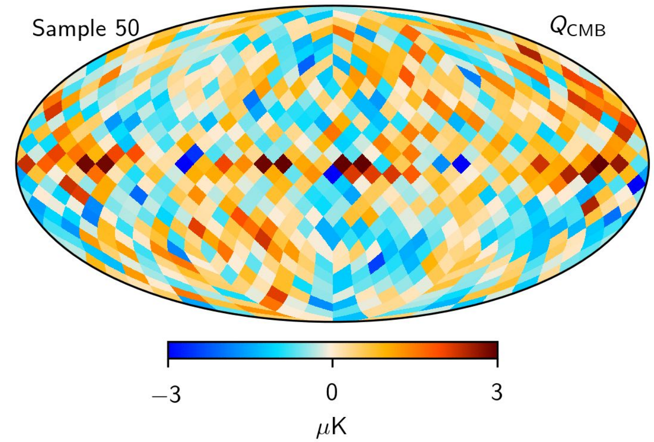
Instrument



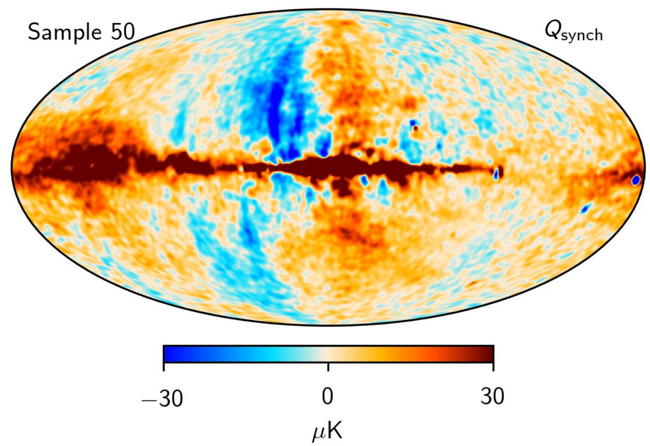
Correlated noise



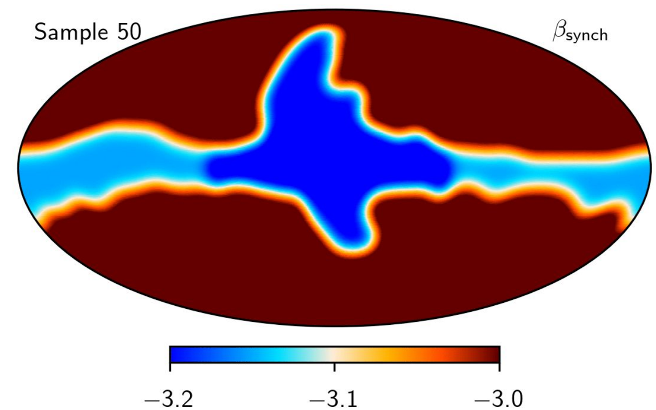
CMB Stokes Q



Synch Stokes Q

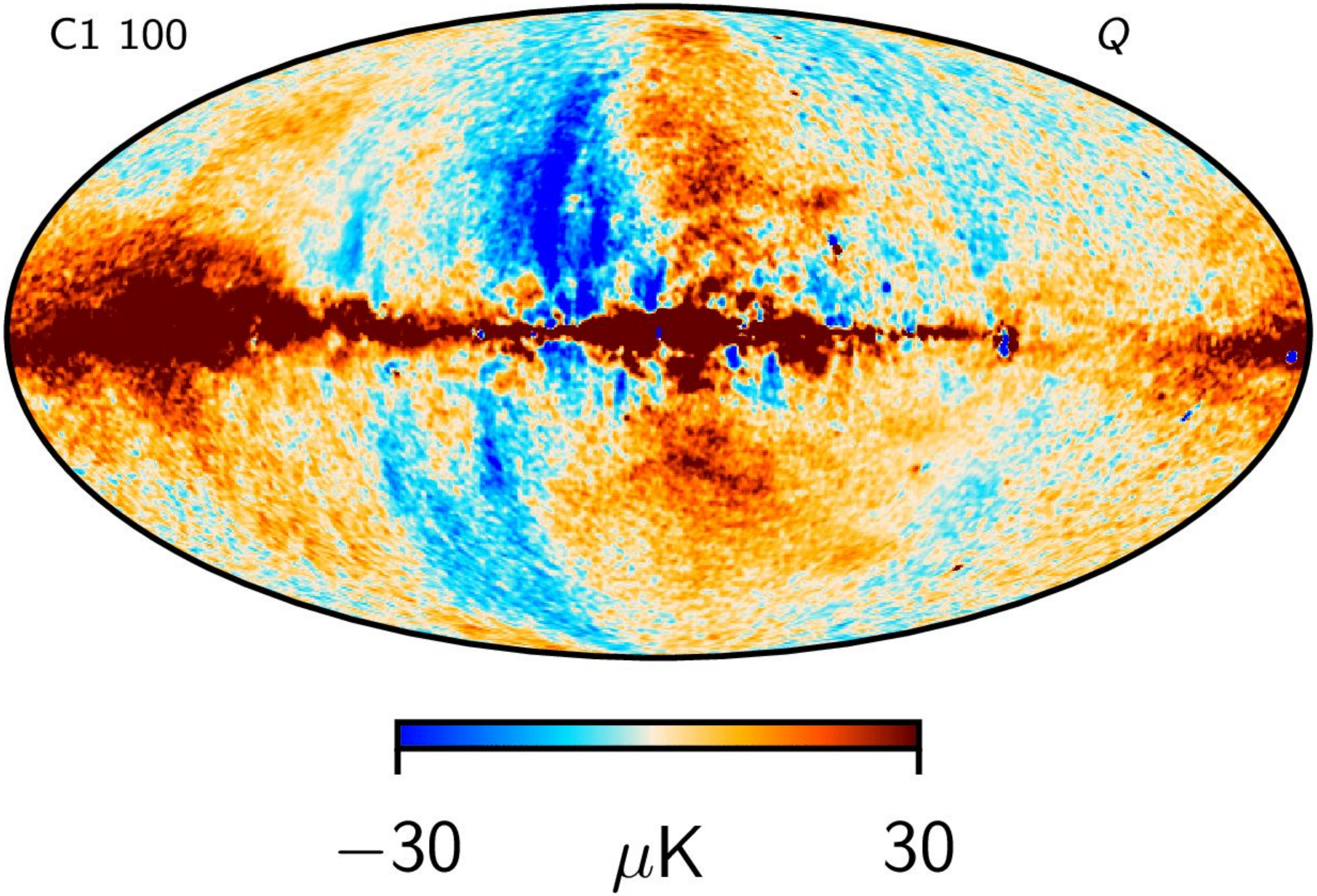


Synch pol β



...

Frequency maps: 30 GHz Stokes Q

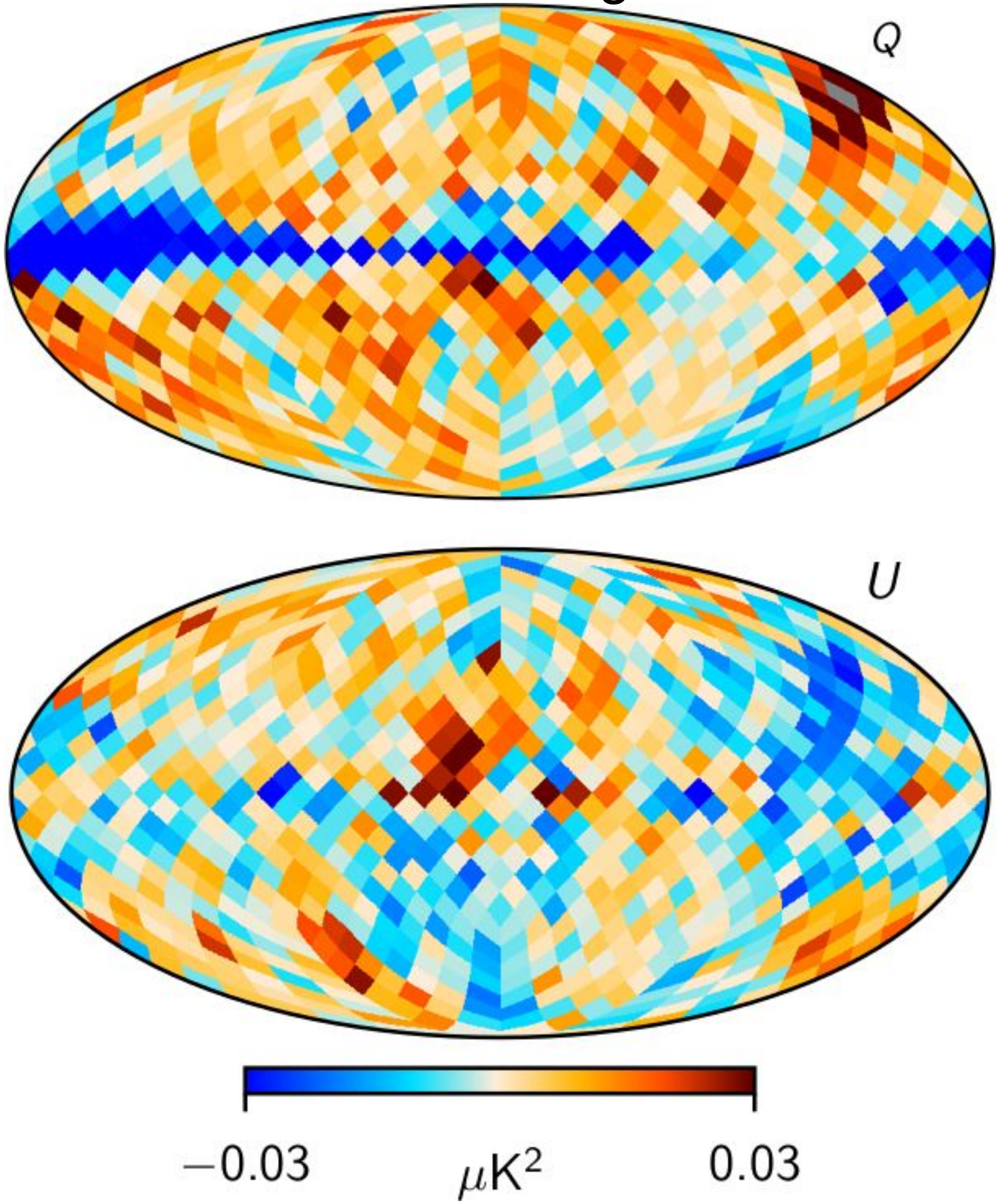


Suur-Uski et al. (2020)

Sample-based low-l likelihood

European Commission

Slice through N



Compute low-resolution CMB map and covariance matrix directly from samples:

$$\hat{\mathbf{s}}_{\mathrm{CMB}} = \left\langle \mathbf{s}^{i}_{\mathrm{CMB}} \right\rangle$$

$$\mathrm{N} = \left\langle (\mathbf{s}^{i}_{\mathrm{CMB}} - \hat{\mathbf{s}}_{\mathrm{CMB}})(\mathbf{s}^{i}_{\mathrm{CMB}} - \hat{\mathbf{s}}_{\mathrm{CMB}})^{t} \right\rangle$$

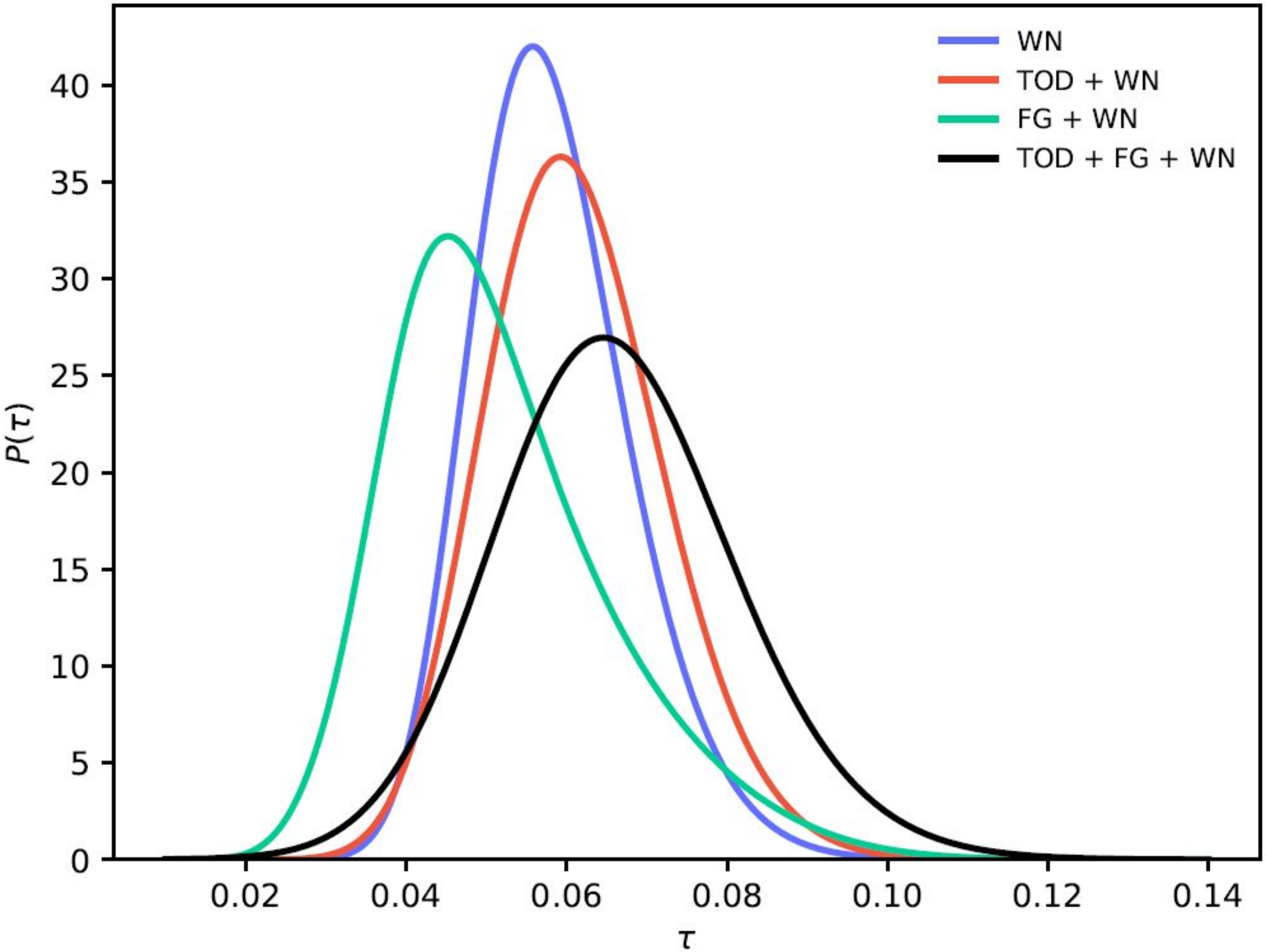
This is the first time uncertainties from **gain, bandpass and a fine-grained foreground model** have been consistently propagated into **CMB low-l likelihood** inputs

$$P(C_\ell \mid \hat{\mathbf{s}}_{\mathrm{CMB}}) \propto \frac{e^{-\frac{1}{2}\hat{\mathbf{s}}^{t}_{\mathrm{CMB}}(\mathrm{S}(C_\ell)+\mathrm{N})^{-1}\hat{\mathbf{s}}_{\mathrm{CMB}}}}{\sqrt{|\mathrm{S}(C_\ell)+\mathrm{N}|}}$$

Colombo et al. (2020)

Beyond PLANCK

Uncertainties on the optical depth of reionization



Paradiso et al. (2020)

Computational resource requirements

Item	30 GHz	44 GHz	70 GHz	Sum
Data volume				
Uncompressed data volume	761 GB	1 633 GB	5 522 GB	7 915 GB
Compressed data volume/RAM requirements	86 GB	178 GB	597 GB	**861 GB**
Processing time (cost per run)				
TOD initialization/IO time	176 sec	288 sec	753 sec	1217 sec
Other initializati				663 sec
Total initializa				**1880 sec**
Gibbs sampling s				
Data decompre				393 sec
TOD projectio				330 sec
Sidelobe evalu				480 sec
Orbital dipole (				449 sec
Gain sampling				94 sec
Correlated nois				3138 sec
TOD binning (				498 sec
Loss due to po				502 sec
Sum of other T				306 sec
TOD processing cost per sample	[illegible]	[illegible]	[illegible]	**6396 sec**
Amplitude sampling, $P(\boldsymbol{a} \mid \boldsymbol{d}, \omega \setminus \boldsymbol{a})$				527 sec
Spectral index sampling, $P(\beta \mid \boldsymbol{d}, \omega \setminus \beta)$				1080 sec
Other steps				149 sec
Total cost per sample				$\boxed{\textbf{8168 sec}}$

2.3 hours/sample

on

72-core node with 1.5 TB RAM

Galloway et al. (2020)

- **Six independent Gibbs chains of each 200 samples** were generated on 6 compute nodes
- Total wall production time for main run was **3 weeks**
- Total CPU cost for main run was **220,000 CPU hours**
  - For comparison, simulating one single traditional Planck Full Focal Plane 70 GHz realization costs O(10^4) CPU hours (Planck Collaboration 2016, A&A, 596, A12)

One-minute summary

1. **Instrumental and astrophysical uncertainties are intimately coupled!**
    - Cannot know your instrument without knowing the sky, and cannot know the sky without knowing your instrument
    - Assertion: A global instrument+foreground analysis strategy is a strict prerequisite for reaching $r \lesssim 10^{-2}$, not a "nice-thing-to-have"
2. **Know your instrument before commissioning!**
    - Measure 4π beams, bandpasses, ADCs etc. *accurately* on ground!
    - Develop realistic and practical uncertainty models, and integrate these into your analysis pipeline
3. **Degeneracies should be broken by additional data, not "clever algorithms"**
    - Design your own experiment to have as few "blind spots" as possible
    - Use results from earlier experiments -- but do propagate uncertainties properly!

- **BeyondPlanck has implemented the first integrated end-to-end Bayesian CMB analysis pipeline, called Commander3**
    - This has been successfully applied to Planck LFI
    - Computational costs are competitive with the traditional frequentist/simulation approach

Online resources

BeyondPlanck project

Main webpage: https://beyondplanck.science
Products: https://products.beyondplanck.science
https://pla.esac.esa.int (subset; when papers are accepted)
Papers: https://beyondplanck.science/products/publications
Discussion forum: https://forums.beyondplanck.science

Commander

Source code : https://github.com/cosmoglobe/Commander
Documentation: https://docs.beyondplanck.science

Cosmoglobe

Main webpage: http://cosmoglobe.uio.no

Planck Legacy Archive (selected BeyondPlanck products coming soon)

Link: https://pla.esac.esa.int

The BeyondPlanck collaboration

EU-funded institutions

Kristian Joten Andersen
Ragnhild Aurlien
Ranajoy Banerji
Maksym Brilenkov
Hans Kristian Eriksen
Johannes Røsok Eskilt
Marie Kristine Foss
Unni Fuskeland
Eirik Gjerløw
Mathew Galloway
Daniel Herman
Ata Karakci
Håvard Tveit Ihle
Metin San
Trygve Leithe Svalheim
Harald Thommesen
Duncan Watts
Ingunn Kathrine Wehus

Marco Bersanelli
Loris Colombo
Cristian Franceschet
Davide Maino
Aniello Mennella
Simone Paradiso

Sara Bertocco
Samuele Galeotta
Gianmarco Maggio
Michele Maris
Daniele Tavagnacco
Andrea Zacchei

Elina Keihänen
Anna-Stiina Suur-Uski

Stelios Bollanos
Stratos Gerakakis
Maria Ieoronymaki
Ilias Ioannou

External collaborators

Brandon Hensley

Jeff Jewell

Reijo Keskitalo

Bruce Partridge

Martin Reinecke

Funding

European Commission

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 776282

- “*BeyondPlanck*”
  - COMPET-4 program
  - PI: Hans Kristian Eriksen
  - Grant no.: 776282
  - Period: Mar 2018 to Nov 2020

Collaborating projects:

- “*bits2cosmology*”
  - ERC Consolidator Grant
  - PI: Hans Kristian Eriksen
  - Grant no: 772 253
  - Period: April 2018 to March 2023
- “*Cosmoglobe*”
  - ERC Consolidator Grant
  - PI: Ingunn Wehus
  - Grant no: 819 478
  - Period: June 2019 to May 2024

Beyond PLANCK

Questions?

European Commission

Beyond PLANCK

Commander

UNIVERSITAS OSLOENSIS MDCCCXI

UNIVERSITAS STUDIORUM MEDIOLANENSIS

INAF ISTITUTO NAZIONALE DI ASTROFISICA OSSERVATORIO ASTRONOMICO DI TRIESTE

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

planetek hellas

JPL

BERKELEY LAB

HAVERFORD COLLEGE

THE FRAMEWORK PROGRAMME FOR RESEARCH AND INNOVATION
HORIZON 2020

erc

Cosmoglobe

Beyond PLANCK