


BEYONDPLANCK

XI. Bayesian CMB analysis with sample-based end-to-end error propagation

L. P. L. Colombo^{1,3} , J. R. Eskilt², S. Paradiso^{1,3}, H. Thommesen², K. J. Andersen², R. Aurlien², R. Banerji², A. Basyrov², M. Bersanelli^{1,3,4}, S. Bertocco⁵, M. Brilenkov², M. Carbone⁶, H. K. Eriksen², M. K. Foss², C. Franceschet^{1,3}, U. Fuskeland², S. Galeotta⁵, M. Galloway², S. Gerakakis⁶, E. Gjerløw², B. Hensley⁷, D. Herman², M. Iacobellis⁶, M. Ieronymaki⁶, H. T. Ihle², J. B. Jewell⁸, A. Karakci², E. Keihänen^{9,10}, R. Keskitalo¹¹, G. Maggio⁵, D. Maino^{1,3,4}, M. Maris⁵, B. Partridge¹², M. Reinecke¹³, A.-S. Suur-Uski^{9,10}, T. L. Svalheim², D. Tavagnacco^{5,14}, D. J. Watts², I. K. Wehus², and A. Zacchei⁵

¹ Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria, 16, Milano, Italy
e-mail: loris.colombo@unimi.it

² Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo, Norway

³ INFN, Sezione di Milano, Via Celoria 16, Milano, Italy

⁴ INAF-IASF Milano, Via E. Bassini 15, Milano, Italy

⁵ INAF – Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, Trieste, Italy

⁶ Planetek Hellas, Leoforos Kifisias 44, Marousi 151 25, Greece

⁷ Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

⁸ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, USA

⁹ Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2, Helsinki, Finland

¹⁰ Helsinki Institute of Physics, University of Helsinki, Gustaf Hällströmin katu 2, Helsinki, Finland

¹¹ Computational Cosmology Center, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

¹² Haverford College Astronomy Department, 370 Lancaster Avenue, Haverford, PA, USA

¹³ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany

¹⁴ Dipartimento di Fisica, Università degli Studi di Trieste, Via A. Valerio 2, Trieste, Italy

Received 28 July 2022 / Accepted 17 January 2023

ABSTRACT

We present posterior sample-based cosmic microwave background (CMB) constraints from *Planck* LFI and WMAP observations as derived through global end-to-end Bayesian processing within the BEYONDPLANCK framework. We first used these samples to study correlations between CMB, foreground, and instrumental parameters. We identified a particularly strong degeneracy between CMB temperature fluctuations and free-free emission on intermediate angular scales ($400 \lesssim \ell \lesssim 600$), mitigated through model reduction, masking, and resampling. We compared our posterior-based CMB results with previous *Planck* products and found a generally good agreement, however, with notably higher noise due to our exclusion of *Planck* HFI data. We found a best-fit CMB dipole amplitude of $3362.7 \pm 1.4 \mu\text{K}$, which is in excellent agreement with previous *Planck* results. The quoted dipole uncertainty is derived directly from the sampled posterior distribution and does not involve any ad hoc contributions for *Planck* instrumental systematic effects. Similarly, we find a temperature quadrupole amplitude of $\sigma_2^{TT} = 229 \pm 97 \mu\text{K}^2$, which is in good agreement with previous results in terms of the amplitude, but the uncertainty is one order of magnitude greater than the naive diagonal Fisher uncertainty. Concurrently, we find less evidence of a possible alignment between the quadrupole and octopole than previously reported, due to a much larger scatter in the individual quadrupole coefficients that is caused both by marginalizing over a more complete set of systematic effects – as well as by requiring a more conservative analysis mask to mitigate the free-free degeneracy. For higher multipoles, we find that the angular temperature power spectrum is generally in good agreement with both *Planck* and WMAP. At the same time, we note that this is the first time that the sample-based, asymptotically exact Blackwell-Rao estimator has been successfully established for multipoles up to $\ell \leq 600$. It now accounts for the majority of the cosmologically important information. Overall, this analysis demonstrates the unique capabilities of the Bayesian approach with respect to end-to-end systematic uncertainty propagation and we believe it can and should play an important role in the analysis of future CMB experiments. Cosmological parameter constraints are presented in a companion paper.

Key words. cosmic background radiation – cosmology: observations – methods: statistical – cosmology: miscellaneous

1. Introduction

Detailed measurements of the cosmic microwave background (CMB) have revolutionized modern cosmology over the last three decades. Offering a unique and clear view of the

baby Universe only 380 000 years after the Big Bang (e.g., Bennett et al. 2013; Planck Collaboration I 2020), its minuscule temperature fluctuations allow scientists to measure a range of cosmological parameters with subpercent accuracy and this work has culminated in a tremendously successful standard

model of cosmology known as Λ CDM (e.g., [Hinshaw et al. 2013](#); [Planck Collaboration VI 2020](#)). According to this model, the Universe began with a hot Big Bang some 13.8 billion years ago: it was filled with Gaussian random density fluctuations during a cataclysmic quantum mechanical process called inflation taking place only some 10^{-34} s after the beginning, when it grew exponentially in size. Today it is populated by about 65% dark energy (Λ), 30% cold dark matter (CDM), and only 5% ordinary baryonic matter.

While this model is extremely successful in terms of predicting cosmological observations quantitatively, it leaves fundamental questions unanswered, for instance, regarding what dark matter is made of and what dark energy is. Additionally, the biggest question of all might simply pertain to what exactly happened during the very first moments of the Big Bang. As of 2022, cosmic inflation (e.g., [Kamionkowski & Kovetz 2016](#)) has represented a basic paradigm for this process that is widely accepted by the community, simply because it is able to heuristically explain a range of important observations in modern cosmology, including cosmological isotropy, flatness, and the absence of topological defects; also, its predictions are largely consistent with CMB measurements, such as Gaussianity and a nearly scale-invariant, but slightly tilted, spectrum of initial perturbations. At the same time, inflation as a general concept is both heavily criticized for being overly flexible (e.g., [Penrose 1989](#); [Ijjas et al. 2014](#)), to the extent that there is an ongoing debate over whether it has any predictive power and it has also been said to lack a robust theoretical foundation, which may require a proper theory of quantum gravity.

To make progress in this area, more data are desperately needed. And the most promising path to such is through deep measurements of large-scale CMB polarization (e.g., [Kamionkowski & Kovetz 2016](#)). A firm prediction of the inflationary paradigm is that there ought to be evidence of a background of primordial gravitational waves that were excited during the period of exponential expansion. If this were so, these super-horizon gravitational waves would also be expected to make an imprint on the CMB field in the form of so-called *B*-mode polarization. The amplitude of this signal is typically measured in terms of the tensor-to-scalar ratio, r , and different inflationary models (corresponding to different inflationary potentials) predict different values for r , with typical values varying between 10^{-4} and 0.1 for large model spaces. The strongest upper limit today is $r < 0.032^1$ at 95% confidence, as measured by the combination of BICEP2/Keck and *Planck* ([Tristram et al. 2022](#)). A robust positive detection of $r > 0$ would rank among the greatest discoveries in cosmology, providing a unique signature of ultra-high energy physics almost at the *Planck* energy scale. As a reflection of the fundamental importance of such a detection, billions of dollars, euros, and yen are currently being invested in efforts to detect this signal ([Gerakakis et al. 2023](#)).

However, the technical challenges involved in making such a discovery are massive. For a typical value of $r \sim 10^{-3}$, the amplitude of the *B*-mode polarization signal will not be more than a few tens of nanokelvins on large angular scales. All sources of systematic errors must therefore be controlled to unprecedented levels, no matter whether they are of instrumental or astrophysical origin, and the corresponding uncertainties must be accurately propagated throughout the entire analysis process. Underestimating the integrated uncertainty on r by, say, a factor

of two could turn an innocent 2.5σ fluke into a fatal 5σ false claim of new physics.

Most pre-*Planck* and early *Planck* CMB analysis pipelines have effectively relied on systematic errors being relatively small compared both with the target signal and the noise level of the given experiment (e.g., [Bennett et al. 2013](#); [Planck Collaboration I 2014](#)). In many cases, it has been an acceptable approximation to account primarily for (correlated and white) noise uncertainties on the instrument side and sample and cosmic variance on the CMB side. The impact of astrophysical foregrounds, whether caused by Milky Way or extra-galactic sources, has typically been minor and could often be accounted for through simple template fitting or internal linear combination methods (e.g., [Bennett et al. 2003](#); [Planck Collaboration XII 2014](#)). However, as the signal-to-noise ratio of a given dataset increases, the relative importance of systematic errors increases, to the point that they eventually totally dominate the error budget. A key example of this is the strong coupling between calibration and astrophysical foregrounds; since high-sensitivity CMB experiments, such as *Planck*, directly exploit the CMB dipole to estimate their gain, it is key to establish a robust model of any Galactic foreground that may obscure this signal. At the same time, such a foreground model can only be derived from the same high-sensitivity dataset, leading to a highly nonlinear analysis problem. For *Planck*, this insight eventually led to the development of highly integrated analysis pipelines ([Delouis et al. 2019](#); [Planck Collaboration Int. LVII 2020](#)) that jointly fit both instrumental and astrophysical parameters as part of the mapmaking process. It is safe to assume that similar integrated approaches will be even more important for next-generation inflationary *B*-mode experiments, due to their extreme precision requirements.

To understand how error propagation may be improved for next-generation experiments, it is worth noting that two fundamentally different modes of operations have seen widespread applications in the CMB field up through our day, corresponding either to the use of simulations or Bayesian statistics, respectively. In the simulation approach, it is assumed that we precisely know the cosmological model, the astrophysical foregrounds, and the instrument, and then we would derive as many realistic time-ordered data (TOD) simulations as possible of the dataset in question (e.g., [Planck Collaboration XII 2016](#)). Each simulation is then processed with exactly the same algorithms as the real data and the scatter in the final quantity is taken as the uncertainty of the point estimate derived from the data. This mode of operation has traditionally dominated all lower-level aspects of CMB data processing, including calibration, mapmaking, and component separation (e.g., [Planck Collaboration VII 2016](#); [Planck Collaboration IV 2020](#)).

In contrast, the key elements in the Bayesian approach consist of explicit models for the data and likelihood in question, and the analysis process simply amounts to mapping out the corresponding posterior distribution. In practice, this is typically done using modern Markov chain Monte Carlo (MCMC) methods, due to the high dimensionality of the data model. This approach is typically preferred for the high-level aspects of the analysis and, in particular, for cosmological parameter estimations (e.g., [Lewis & Bridle 2002](#)). One main reason for this is that it allows for a more natural and efficient exploration of degeneracies between parameters. For *Planck*, the Bayesian approach was used for the final cosmological parameter stage, integrating a limited instrumental and astrophysical model directly into the corresponding likelihood ([Planck Collaboration V 2020](#)), allowing for a joint exploration of

¹ Evaluated at a pivot scale of 0.05 Mpc^{-1} .

a few hundred free parameters (Planck Collaboration VI 2020). However, most instrumental uncertainties were still estimated using the low-level simulation approach.

In this paper, we consider error propagation within the context of a novel end-to-end Bayesian analysis framework called BEYONDPLANCK (BeyondPlanck Collaboration 2023). This pipeline is, in principle, equivalent to the Bayesian cosmological parameter approach described above, but with one critical difference: In BEYONDPLANCK, the entire pipeline is integrated into the core Monte Carlo sampler (BeyondPlanck Collaboration 2023). As such, the number of free model parameters does not amount to hundreds, but billions, in fact – and there is no separation between low-level and high-level analysis. Two key advantages of this global integrated approach are: firstly, a joint exploration of all free parameters and, secondly, seamless end-to-end error propagation. In short, it is the ultimate logical extension of *Planck*'s approach of adding a handful of critical instrumental and astrophysical parameters to the CMB likelihood. It is also interesting to note that this approach was in fact first suggested almost 20 years ago by Jewell et al. (2004) and Wandelt et al. (2004). It took almost two decades of algorithmic and computer developments before it could be realized in practice.

In this paper, we present CMB results derived from within the BEYONDPLANCK pipeline, while a series of companion papers describes the individual instrumental (Herman et al. 2023a; Ihle et al. 2023; Gjerløw et al. 2023; Galloway et al. 2023a; Svalheim et al. 2023a) and astrophysical components (Andersen et al. 2023; Svalheim et al. 2023b; Herman et al. 2023b). An important common feature in all of these papers, however, is the fact that each free parameter is quantified in the form of a set of “samples” drawn from the joint posterior distribution. At first glance, these look very similar to the simulations produced in the traditional low-level approach, but they have a fundamentally different statistical interpretation: While a simulation represents one possible instrument configuration in a random universe, unconstrained by the actual measurements, a posterior sample represents one possible instrument configuration in our universe, as constrained by the actual measurements.

An important consequence of this difference is that the two approaches cover different aspects of the full analysis problem in which they excel. For questions that may be formulated in terms of numerical parameter estimates that require a robust error assessment, for instance “the best-fit value of r ”, the Bayesian approach is ideal. For questions that may be formulated in terms of statistical agreement with a general paradigm, such as “how likely it is for the CMB Cold Spot to appear in a Gaussian and isotropic universe”, the simulation-based approach is ideal. That is not to say that either of the two methods cannot address questions in the other category – but they are complementary and generally better suited to answer specific questions. Going forward, we consider it very likely that most next-generation experiments will strive to implement both pipeline types and cross-validate their results between them.

In this paper, we demonstrate the use of these novel posterior samples for several classic CMB analysis applications, including the CMB dipole estimation, power spectrum estimation, and low- ℓ anomaly studies, with special attention paid to robust error propagation. However, we stress that the current BEYONDPLANCK processing primarily focuses on *Planck* LFI data and, in particular, it does not include *Planck* HFI observations in the 100–217 GHz range (BeyondPlanck Collaboration 2023). The results are therefore significantly less sensitive than the main *Planck* results in most respects. In general, the main purpose of the current paper is to demonstrate the sample-based CMB analysis from an algorithmic

point-of-view, while leaving the full integration of additional state-of-the-art datasets to a future work.

The paper is organized as follows. In Sect. 2, we briefly review the BEYONDPLANCK data model and show how CMB samples are derived within this framework. In Sect. 3, we inspect the raw outputs from the algorithm in the form of posterior samples, quantify correlations among the various parameters, and identify one particularly strong degeneracy with respect to free-free emission. In Sect. 4, we consider posterior mean maps and power spectra and compare their properties with those presented by earlier analyses. In Sect. 5, we present the first fully Bayesian estimate of the CMB Solar dipole from *Planck* data, before we revisit selected low- ℓ anomalies in Sect. 6. We present our conclusions in Sect. 7.

2. BEYONDPLANCK and end-to-end CMB analysis

2.1. General model and Gibbs sampling scheme

The starting point of the LFI-oriented Bayesian BEYONDPLANCK analysis framework is an explicit parametric model of the time-ordered data of the following form (BeyondPlanck Collaboration 2023),

$$d_{j,t} = g_{j,t} \mathbf{P}_{tp,j} \left[\mathbf{B}_{pp',j}^{\text{symm}} s_{p',j}^{\text{sky}} + s_{j,t}^{\text{orb}} + s_{j,t}^{\text{fsl}} \right] + s_{j,t}^{1\text{Hz}} + n_{j,t}^{\text{corr}} + n_{j,t}^{\text{w}}, \quad (1)$$

$$\equiv s^{\text{tot}} + n_{j,t}^{\text{w}}. \quad (2)$$

In this expression, j is a detector index, t is a time index, and p is a sky pixel index. Furthermore, g represents the time-variable instrumental gain, \mathbf{P} is a matrix that describes the satellite pointing, \mathbf{B}^{symm} denotes a (assumed azimuthally symmetric) beam convolution operator, s^{sky} represents the total astrophysical sky signal, s^{orb} is the orbital CMB dipole, s^{fsl} are the far sidelobe corrections, and $s^{1\text{Hz}}$ represents electronic 1 Hz spike corrections, while n^{corr} and n^{w} represent the correlated and white noise, respectively. Both noise terms are well described by zero-mean Gaussian distributions with different covariances, reflecting their different physical origins. In particular, the main contribution to n^{corr} are gain fluctuations. Splitting the total instrumental noise into correlated and white parts allows for a particularly efficient algorithm that jointly samples g and n^{corr} (BeyondPlanck Collaboration 2023; Ihle et al. 2023). For the convenience of later notation, the last line defines all time-ordered data components, except the white noise, as s^{tot} .

The total sky signal may be decomposed into individual astrophysical emission mechanisms and we assume the following expression in the current analysis,

$$s^{\text{sky}} = \left(a^{\text{CMB}} + a^{\text{quad}}(\nu) \right) \frac{x^2 e^x}{(e^x - 1)^2}, \quad (3)$$

$$+ a^{\text{s}} \left(\frac{\nu}{\nu_{0,\text{s}}} \right)^{\beta_{\text{s}}}, \quad (4)$$

$$+ a^{\text{ff}} \left(\frac{\nu_{0,\text{ff}}}{\nu} \right)^2 \frac{g_{\text{ff}}(\nu; T_{\text{e}})}{g_{\text{ff}}(\nu_{0,\text{ff}}; T_{\text{e}})}, \quad (5)$$

$$+ a^{\text{ame}} \left(\frac{\nu_{0,\text{ame}}}{\nu} \right)^2 \frac{f_{\text{ame}} \left(\nu \cdot \frac{30.0\text{GHz}}{\nu_{\text{p}}} \right)}{f_{\text{ame}} \left(\nu_{0,\text{ame}} \cdot \frac{30.0\text{GHz}}{\nu_{\text{p}}} \right)}, \quad (6)$$

$$+ a^{\text{d}} \left(\frac{\nu}{\nu_{0,\text{d}}} \right)^{\beta_{\text{d}}+1} \frac{e^{h\nu_{0,\text{d}}/k_{\text{B}}T_{\text{d}}} - 1}{e^{h\nu/k_{\text{B}}T_{\text{d}}} - 1}, \quad (7)$$

$$+ U_{\text{mJy}} \sum_{j=1}^{N_{\text{src}}} a_{j,\text{src}} \left(\frac{\nu}{\nu_{0,\text{src}}} \right)^{\alpha_{j,\text{src}}-2}, \quad (8)$$

Table 1. Summary of main BEYONDPLANCK model free parameter families.

Symbol	Parameter	$n_{\text{d.o.f.}}$	Reference
n^{corr}	Correlated noise time-domain amplitude	$n_{\text{det}} \cdot n_{\text{tod}} \sim \mathcal{O}(10^{11})$	Ihle et al. (2023)
ξ_{n}	Noise power spectral density	$3 \cdot n_{\text{det}} \cdot n_{\text{PID}} \sim \mathcal{O}(10^7)$	Ihle et al. (2023)
g	Instrumental gain	$n_{\text{det}} \cdot n_{\text{PID}} \sim \mathcal{O}(10^6)$	Gjerløw et al. (2023)
Δ_{bp}	Detector bandpass correction	$n_{\text{det}} \sim \mathcal{O}(10^1)$	Svalheim et al. (2023a)
a^{CMB}	CMB pixel amplitude map (I, Q, U)	$n_{\text{pol}} \cdot n_{\text{pix}} \sim \mathcal{O}(10^8)$	This work, Paradiso et al. (2023)
C_{ℓ}	CMB angular power spectrum	$\ell_{\text{max}} \sim \mathcal{O}(10^3)$	This work, Paradiso et al. (2023)
a^{s}	Synchrotron pixel amplitude map (I, Q, U)	$n_{\text{pol}} \cdot n_{\text{pix}} \sim \mathcal{O}(10^8)$	Andersen et al. (2023), Svalheim et al. (2023b)
β_{s}	Synchrotron spectral index	5	Andersen et al. (2023), Svalheim et al. (2023b)
a^{ff}	Free-free pixel amplitude map (I)	$n_{\text{pix}} \sim \mathcal{O}(10^7)$	Andersen et al. (2023)
a^{ame}	AME pixel amplitude map (I)	$n_{\text{pix}} \sim \mathcal{O}(10^7)$	Andersen et al. (2023)
ν_{p}	AME characteristic frequency	1	Andersen et al. (2023)
a^{d}	Thermal dust pixel amplitude map (I, Q, U)	$n_{\text{pol}} \cdot n_{\text{pix}} \sim \mathcal{O}(10^8)$	Andersen et al. (2023), Svalheim et al. (2023b)
β_{d}	Thermal dust spectral index	2	Andersen et al. (2023), Svalheim et al. (2023b)
$a_{j,\text{src}}$	Radio source amplitudes and spectral indices	$2 \cdot n_{\text{src}} \sim \mathcal{O}(10^4)$	Andersen et al. (2023)

Notes. For the sake of notation, some symbols denote a collection of heterogeneous parameters defining the corresponding quantity, for instance, the noise power spectral density, ξ^{n} , is minimally specified by a normalization, σ_0 , a characteristic frequency, f_{knee} , and a slope, α . See the specific papers for additional details on each parameter family.

where h is Planck's constant, k_{B} is Boltzmann's constant, and $x \equiv hv/k_{\text{B}}T_{\text{CMB}}$, where $T_{\text{CMB}} = 2.7255$ K is the mean CMB temperature (Fixsen 2009). Each line in this expression represents one specific astrophysical component, each of which is defined in terms of an amplitude map, a , and a spectral energy density, $f(\nu; \beta)$, that describe the strength of the component as a function of frequency, relative to some reference frequency, ν_0 , and some set of free spectral parameters, β . From top to bottom, the six lines describe respectively CMB (including a relativistic quadrupole correction), synchrotron, free-free, AME, and thermal dust emission, and, finally, a discrete set of point sources. We assume that only the CMB, synchrotron, and thermal dust emission are polarized. For further information regarding any of these astrophysical foreground components, we refer to Andersen et al. (2023) and Svalheim et al. (2023b). In practice, each of these terms is integrated separately with respect to the instrumental bandpass of each detector, which itself also is associated with a free correction parameter, Δ_{bp} , as discussed by Svalheim et al. (2023a).

It is convenient to decompose the CMB sky map into spherical harmonics:

$$a^{\text{CMB}}(\hat{n}) = \sum_{\ell=0}^{\ell_{\text{max}}} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n}), \quad (9)$$

where ℓ_{max} denotes an harmonic-space bandwidth limit and $a_{\ell m}$ are the spherical harmonics coefficients. It is common to assume that the CMB field is statistically isotropic, in the case of which the CMB covariance matrix may be defined as:

$$S_{\ell m, \ell' m'}^{\text{CMB}} \equiv \langle a_{\ell m} a_{\ell' m'}^* \rangle \equiv C_{\ell} \delta_{\ell \ell'} \delta_{m m'}, \quad (10)$$

where the brackets indicate an ensemble average and C_{ℓ} is called the angular power spectrum. We note that for simplicity, this notation applies only to CMB temperature analysis, but the generalization to polarization is straightforward and described by Zaldarriaga & Seljak (1997). The angular power spectrum plays a particularly important role in CMB analysis, as this provides a computationally efficient path to cosmological parameter estimations (e.g., Lewis & Bridle 2002). Estimating the power spectrum distribution $P(C_{\ell} | \mathbf{d})$, marginalized over all relevant

systematic effects, may in fact be considered the single most important goal of any CMB analysis pipeline.

Given this parametric signal model, the BEYONDPLANCK approach to CMB analysis follows well-established Bayesian methods. We start by defining $\omega \equiv \{a, \beta, g, \Delta_{\text{bp}}, n^{\text{corr}}, C_{\ell}, \dots\}$ to be the set of all free parameters in the model; instrumental, astrophysical, and cosmological. By Bayes' theorem, the joint posterior distribution may then be written as:

$$P(\omega | \mathbf{d}) = \frac{P(\mathbf{d} | \omega)P(\omega)}{P(\mathbf{d})} \propto \mathcal{L}(\omega)P(\omega), \quad (11)$$

where $\mathcal{L}(\omega) \equiv P(\mathbf{d} | \omega)$ is called the likelihood and $P(\omega)$ is a set of user-specified priors. The likelihood is defined simply by noting that the white noise, which is equal to $\mathbf{d} - \mathbf{s}^{\text{tot}}$ (Eq. (2)), is assumed to be Gaussian and therefore:

$$-2 \ln \mathcal{L}(\omega) = (\mathbf{d} - \mathbf{s}^{\text{tot}}(\omega))^t \mathbf{N}_{\text{wn}}^{-1} (\mathbf{d} - \mathbf{s}^{\text{tot}}(\omega)), \quad (12)$$

where \mathbf{N}_{wn} is the white noise covariance matrix. The prior, $P(\omega)$, is less well defined, and must be specified by the user. Table 1 lists the main BEYONDPLANCK free parameter families. Other model parameters not included in the table are kept fixed to a reference value. We refer to the individual papers for more detail on specific parameter families, while for a summary of the priors adopted in the current analysis, we refer to BeyondPlanck Collaboration (2023).

It is important to note that ω includes a vast number of parameters with different impact on the final posterior. For instance, the correlated noise, n^{corr} , contains in principle billions of degrees of freedom, one for each time sample. However, n^{corr} represents a non-stationary sky signal that is significantly constrained by a power spectral density prior, and each n^{corr} sample therefore affects higher-level quantities almost negligibly. In contrast, each astrophysical sky map contains millions of degrees of freedom on orders of magnitude fewer than the correlated noise, but these are all sky-stationary, and they therefore have a larger effect on the full posterior. Finally, there are a handful of global parameters that affect every single sky pixel at once, such as the absolute gain and bandpass corrections, and these have a massive impact on almost all other model parameters.

Both the vast number of free parameters and their complex relationships make it a significant computational challenge to map out the posterior distribution efficiently. The only computationally feasible approach suggested to date is Gibbs sampling (Geman & Geman 1984), which allows the user to draw samples from a joint distribution by iterating over all corresponding conditional distributions. For BEYONDPLANCK, this process may be formally written in terms of the following Gibbs chain:

$$\mathbf{g} \leftarrow P(\mathbf{g} \mid \mathbf{d}, \xi_n, \Delta_{\text{bp}}, \mathbf{a}, \beta, C_\ell), \quad (13)$$

$$\mathbf{n}_{\text{corr}} \leftarrow P(\mathbf{n}_{\text{corr}} \mid \mathbf{d}, \mathbf{g}, \xi_n, \Delta_{\text{bp}}, \mathbf{a}, \beta, C_\ell), \quad (14)$$

$$\xi_n \leftarrow P(\xi_n \mid \mathbf{d}, \mathbf{g}, \mathbf{n}_{\text{corr}}, \Delta_{\text{bp}}, \mathbf{a}, \beta, C_\ell), \quad (15)$$

$$\Delta_{\text{bp}} \leftarrow P(\Delta_{\text{bp}} \mid \mathbf{d}, \mathbf{g}, \mathbf{n}_{\text{corr}}, \xi_n, \mathbf{a}, \beta, C_\ell), \quad (16)$$

$$\beta \leftarrow P(\beta \mid \mathbf{d}, \mathbf{g}, \mathbf{n}_{\text{corr}}, \xi_n, \Delta_{\text{bp}}, C_\ell), \quad (17)$$

$$\mathbf{a} \leftarrow P(\mathbf{a} \mid \mathbf{d}, \mathbf{g}, \mathbf{n}_{\text{corr}}, \xi_n, \Delta_{\text{bp}}, \beta, C_\ell), \quad (18)$$

$$C_\ell \leftarrow P(C_\ell \mid \mathbf{d}, \mathbf{g}, \mathbf{n}_{\text{corr}}, \xi_n, \Delta_{\text{bp}}, \mathbf{a}, \beta), \quad (19)$$

where \leftarrow indicates sampling from the distribution on the right hand side and \mathbf{a} stands for the collection of all map amplitude degrees of freedom (e.g., $a^{\text{CMB}}, a^{\text{ff}}, \dots$), which are sampled jointly for better efficiency (see Sect. 2.2). The Bayesian CMB analysis as implemented within the BEYONDPLANCK framework is nothing but repeated sampling from each of these distributions and the main product from this process is a discrete set of samples drawn from the joint posterior distribution, $P(\omega \mid \mathbf{d})$, which naturally and seamlessly allows for detailed instrumental systematics and astrophysical foreground marginalization.

The CMB sky map and power spectrum estimation are accounted for in the above Gibbs loop in Eqs. (18) and (19), respectively, and explicit expressions for these were first derived by Jewell et al. (2004) and Wandelt et al. (2004). All other steps describe either instrumental or astrophysical effects, and only affect the CMB estimates indirectly. Sampling algorithms for each of those distributions are described in detail in BeyondPlanck Collaboration (2023) and references therein, and in the following we only briefly review the sampling algorithms for Eqs. (18) and (19).

2.2. CMB sky map sampling, $P(a^{\text{CMB}} \mid \mathbf{d}, \mathbf{g}, \beta, C_\ell, \dots)$

To derive an appropriate sampling algorithm for $P(a^{\text{CMB}} \mid \mathbf{d}, \mathbf{g}, \beta, \dots)$, we start with the general data model defined in Eqs. (2) and (3) and aim to isolate the component amplitude, \mathbf{a} , parameter. In principle, we could even isolate the a^{CMB} parameter alone, but since it generally leads to a shorter Monte Carlo correlation length to sample partially degenerate components jointly, we aim to derive a joint sampling step for both CMB and astrophysical foreground amplitudes; we refer to Andersen et al. (2023) for further details.

The first step in the algorithm corresponds essentially to mapmaking. Because all instrumental parameters are conditioned on in this distribution, we may deterministically form the following residual:

$$r_{j,t} \equiv \frac{d_{j,t} - (s_{j,t}^{\text{1Hz}} + n_{j,t}^{\text{corr}})}{g_{j,t}} - \mathbf{P}_{tp,j} [s_{j,t}^{\text{orb}} + s_{j,t}^{\text{fsl}}], \quad (20)$$

$$= \mathbf{P}_{tp,j} \mathbf{B}_{pp',j} s_{p',j}^{\text{sky}} + n_{j,t}^{\text{w}} / g_{j,t}, \quad (21)$$

which now represents TOD that contain only astrophysical signal convolved with an (implicitly assumed azimuthally symmetric, see BeyondPlanck Collaboration 2023, for an in depth discussion) beam and white noise, all in calibrated brightness temperature units. This residual may be compressed nearly

losslessly into a pixelized sky map, m_ν , by solving the so-called mapmaking equation,

$$\left(\sum_j \mathbf{P}_j^t \mathbf{N}_{j,\text{wn}}^{-1} \mathbf{P}_j \right) \mathbf{m}_\nu = \sum_j \mathbf{P}_j^t \mathbf{N}_{j,\text{wn}}^{-1} \mathbf{r}_j. \quad (22)$$

For *Planck*, this equation may be solved pixel-by-pixel, and it is therefore computationally very fast.

The second step in the algorithm corresponds essentially to component separation. Given the above frequency maps, the data model described by Eqs. (2) and (3) may now be rewritten compactly in terms of sky maps:

$$\mathbf{m}_\nu = \mathbf{B}_\nu^{\text{symm}} \mathbf{M}_{\nu,c} \mathbf{a} + \mathbf{n}_\nu^{\text{wn}}, \quad (23)$$

where $\mathbf{M}_{\nu,c}$ is called the mixing matrix and encodes the bandpass-integrated SEDs for the various astrophysical components in each column; when multiplied by the amplitude vector, this matrix generates the full sky signal at frequency, ν , in the appropriate units for that channel.

It is now straightforward to sample \mathbf{a} , again based on the observation that the white noise component is Gaussian, and, therefore, that $\mathbf{m}_\nu - \mathbf{B}_\nu^{\text{symm}} \mathbf{M}_{\nu,c} \mathbf{a}$ is Gaussian with the same covariance. The necessary sampling equation for this step is therefore structurally identical to the mapmaking equation in Eq. (22), except that it has an additional fluctuation term in order to propagate the noise uncertainties:

$$\left(\sum_\nu \mathbf{B}_\nu^t \mathbf{M}_\nu^t \mathbf{N}_{\nu,\text{wn}}^{-1} \mathbf{M}_\nu \mathbf{B}_\nu \right) \mathbf{a} = \sum_\nu \mathbf{B}_\nu^t \mathbf{M}_\nu^t \mathbf{N}_{\nu,\text{wn}}^{-1} \mathbf{m}_\nu + \sum_\nu \mathbf{B}_\nu^t \mathbf{M}_\nu^t \mathbf{N}_{\nu,\text{wn}}^{-\frac{1}{2}} \eta_\nu, \quad (24)$$

where η_ν is a random vector of $N(0, 1)$ stochastic variates; for a full derivation of this equation, we refer to Appendix A in BeyondPlanck Collaboration (2023). A computationally efficient Conjugate Gradient (CG) solver for this equation was presented by Seljebotn et al. (2019).

Equation (24) does not account for priors on \mathbf{a} . We support Gaussian priors in our analyses, as defined in terms of some mean map, μ , and a corresponding prior covariance matrix, \mathbf{S} . The purpose of this prior is two-fold; firstly, for the CMB component it directly defines the connection to the angular power spectrum and cosmological parameters, as described by the CMB covariance matrix in Eq. (10). Secondly, for astrophysical foregrounds it both allows us to introduce useful information in the form of prior knowledge from other datasets to break particularly difficult degeneracies and it allows us to impose smoothness on small angular scales. With such a Gaussian prior in place, the full sampling equation for \mathbf{a} is expressed as:

$$\left(\mathbf{S}^{-1} + \sum_\nu \mathbf{B}_\nu^t \mathbf{M}_\nu^t \mathbf{N}_{\nu,\text{wn}}^{-1} \mathbf{M}_\nu \mathbf{B}_\nu \right) \mathbf{a} = \sum_\nu \mathbf{B}_\nu^t \mathbf{M}_\nu^t \mathbf{N}_{\nu,\text{wn}}^{-1} \mathbf{m}_\nu + \mathbf{S}^{-1} \mu + \sum_\nu \mathbf{B}_\nu^t \mathbf{M}_\nu^t \mathbf{N}_{\nu,\text{wn}}^{-\frac{1}{2}} \eta_\nu + \mathbf{S}^{-\frac{1}{2}} \eta_0. \quad (25)$$

A derivation of this expression is given in Appendix A in BeyondPlanck Collaboration (2023) and a detailed discussion of foreground priors in BEYONDPLANCK is presented in Andersen et al. (2023).

It is worth noting that all of the above equations are general in terms of basis sets and can be applied equally well to objects defined in terms of pixels or spherical harmonics or

any other complete basis on the sphere. In practice, our current codes model all diffuse components in terms of spherical harmonics up to some band limit ℓ_{\max} . The main reason for this choice is simply that harmonics are more efficient in terms of the number of free parameters than pixels; for a HEALPix² map (Górski et al. 2005) with a given N_{side} resolution parameter, there are $12 N_{\text{side}}^2$ pixels, while for a typical maximum band limit of $\ell_{\max} = 3 N_{\text{side}}$, there are only $(\ell_{\max} + 1)^2 \approx 9 N_{\text{side}}^2$ spherical harmonic coefficients. In addition, it is easier to impose additional smoothness on a given foreground component in harmonic space, simply by reducing ℓ_{\max} for that component. We emphasize, however, that this is only a practical choice, not a fundamental one; the algorithm works equally well with both bases.

2.3. Angular power spectrum sampling, $P(C_\ell | \mathbf{d}, \mathbf{a}, \dots)$

Next, we need to derive a sampling algorithm for the power spectrum distribution, $P(C_\ell | \mathbf{d}, \mathbf{a}, \dots)$. This was first presented by Wandelt et al. (2004) and here we only review the main steps in the following in brief.

Firstly, we make the trivial observation that $P(C_\ell | \mathbf{d}, \mathbf{a}, \dots) = P(C_\ell | \mathbf{a}^{\text{CMB}})$; if we already know the CMB map, \mathbf{a}^{CMB} , with infinite precision, no further instrumental or astrophysical knowledge can possibly provide more information regarding the CMB power spectrum. Secondly, in the previous sampling step we assumed only that the CMB SED is defined by a blackbody spectrum; in this step, we additionally assume that the CMB is statistically isotropic, namely, that its harmonic space covariance matrix is diagonal and given by C_ℓ and that it is Gaussian-distributed. With these additional assumptions, the relevant distribution may be written as follows:

$$P(C_\ell | \mathbf{a}) \propto \frac{e^{-\frac{1}{2} \mathbf{a}^T \mathbf{S}^{-1} \mathbf{a}}}{\sqrt{|\mathbf{S}|}}, \quad (26)$$

$$= \frac{e^{-\frac{1}{2} \sum_\ell \frac{a_{\ell m}^2}{C_\ell}}}{\prod_\ell C_\ell^{\frac{2\ell+1}{2}}}, \quad (27)$$

$$= \prod_\ell \frac{e^{-\frac{2\ell+1}{2} \frac{\sigma_\ell}{C_\ell}}}{C_\ell^{\frac{2\ell+1}{2}}}, \quad (28)$$

where $\sigma_\ell = \frac{1}{2\ell+1} |a_{\ell m}|^2$ is the observed power spectrum of our specific universe.

The distribution inside the product in Eq. (28) is called an inverse Gamma distribution with $2\ell + 1$ degrees of freedom. Its multivariate generalization of this, which is needed for the polarization analysis, is called the inverse Wishart distribution (Larson et al. 2007). Sampling from an inverse Gamma distribution is trivial; simply draw $2\ell - 1$ independent Gaussian random variates, η_i , and set $C_\ell = \sigma_\ell / \sum_i \eta_i^2$ (Wandelt et al. 2004).

Unfortunately, as discussed by Eriksen et al. (2004a), this strict Gibbs sampling algorithm for \mathbf{a} and C_ℓ has a significant drawback in the low signal-to-noise regime, namely, that the Monte Carlo step size between two consecutive C_ℓ samples is determined by cosmic variance alone, while the full posterior width is defined by both cosmic variance, sample variance (i.e., masking), and instrumental noise. In practice, this algorithm therefore has a prohibitively long correlation length at high multipoles. This problem was addressed and solved by Jewell et al. (2009) and Racine et al. (2016), who proposed a joint sampling step for $\{\mathbf{a}, C_\ell\}$ that moves quickly in the low

signal-to-noise regime. Unfortunately, this step has not yet been fully implemented in the latest version of the Commander code (Galloway et al. 2023b) and it is therefore not used in the BEYONDPLANCK processing. This work is, however, on-going, and will be available in the near future for other projects. An immediate result of this, however, is that in the following, we only present a BEYONDPLANCK temperature power spectrum up to $\ell_{\max} = 600$, while higher multipoles are, when needed, taken from the official *Planck* processing (Planck Collaboration V 2020).

2.4. BEYONDPLANCK data selection

As described by BeyondPlanck Collaboration (2023), the BEYONDPLANCK program has two main goals. The first goal is to implement and demonstrate the world's first end-to-end Bayesian sampling algorithm for CMB observations. The second goal is to resolve a number of outstanding questions regarding the *Planck* LFI data that remained after the conclusion of the official *Planck* consortium. For both of these reasons, the BEYONDPLANCK processing includes significantly less data than if the primary goal had been to establish a new state-of-the-art sky model and CMB sky map. Explicitly, we only include *Planck* LFI 30, 44, and 70 GHz data in the time-domain, which are the main target of the entire analysis; WMAP *Ka*, *Q*, and *V* data to constrain low-frequency foregrounds and poorly observed *Planck* modes; Haslam 408 MHz measurements to constrain synchrotron emission; and *Planck* PR4 measurements at 857 GHz (in temperature) and 353 GHz (in polarization) to constrain the thermal dust emission.

Critically, the CMB-dominated HFI and the WMAP *K*-band data are not included: even though they would clearly result in a better and less degenerate sky model, they would also obscure the impact of the new algorithm because of their high signal-to-noise ratios and they could also potentially introduce unmodeled systematic errors into the LFI results. Instead, a gradual integration of these datasets falls within the scope of the COSMOGLOBE³ framework (Gerakakis et al. 2023; Watts et al. 2023), which is aimed at applying these methods to a broad range of state-of-the-art datasets in the field.

2.5. Masking, degeneracies, and resampling

For an ideal dataset and a well-constrained model, the above algorithm could, in principle, be applied without additional modifications. However, for real-world data, there are several challenges that must be addressed. The first of these is masking: despite the notable complexity of the astrophysical data model described by Eqs. (3)–(8), this is by no means adequate for modeling the actual sky to the statistical precision of the *Planck* data. As a result, we have to remove parts of the sky, in particular the Galactic plane and bright point sources, before actually estimating the CMB power spectrum.

The CMB confidence mask used for the current BEYONDPLANCK processing is shown in Fig. 1 and is generated in a two-step process. First, we compute data-minus-signal residual maps for each CMB-dominated frequency. These are smoothed to 1° angular resolution and thresholded in absolute amplitude. These maps serve a similar purpose as absolute goodness-of-fit tracers as the total χ^2 map that was used to define the Commander confidence mask in for instance the *Planck* 2018 analysis. However, the total χ^2 does not provide information on the quality of the individual components, but only on the

² <https://healpix.jpl.nasa.gov>

³ <https://cosmoglobe.uio.no>

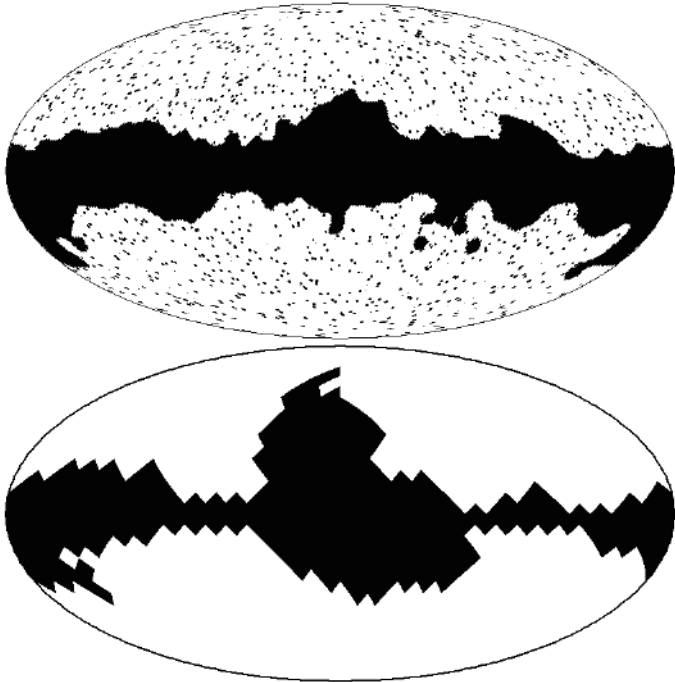


Fig. 1. Temperature (*top*) and polarization (*bottom*) confidence masks used for BEYONDPLANCK CMB analysis. The masks allow, respectively, for a sky fraction of 69% and 68%.

capability of the model to describe the full set of frequencies. By considering only residual maps for the CMB-dominated frequencies, we instead exclude potential modeling issues that affect only foreground reconstruction but are irrelevant for CMB estimation. These partial single-frequency masks are multiplied together to form an overall confidence mask.

The second mask generation step accounts for the resampling procedure after excluding free-free emission from the model. As discussed in greater detail in Sect. 3 below, sample fluctuations in CMB maps are correlated with those in the free-free component at intermediate scales. Intuitively, we do not trust any pixel for which the CMB map is significantly different depending on whether free-free emission is modeled or not. This idea is implemented in practice by generating a single resampled constrained realization; computing the difference between this constrained realization (without free-free emission in the model) and the corresponding CMB map from the main Gibbs analysis (with free-free emission in the model); computing the absolute value, and smoothing to 4° FWHM; and exclude all pixels above a $10 \mu\text{K}$ threshold, corresponding to a $\sim 3\sigma$ fluctuation for the difference map defined above. The resulting mask is median filtered with a 4 deg radius to exclude isolated “islands” inside the Galactic plane; and, finally, we exclude point sources using the *Planck* LFI template point source mask. For polarization, we consider the same *Planck* LFI set of masks described in *Planck Collaboration V* (2020) and adopt the R1.8 mask (with $f_{\text{sky}} = 0.68$) for the polarization cosmological analysis; we refer to *Paradiso et al.* (2023) for a discussion.

Formally speaking, applying a confidence mask in the sampling algorithms described in Sects. 2.2 and 2.3 is trivial; one simply sets the masked pixels in the inverse frequency covariance matrix, N_{ν}^{-1} to zero, thereby assigning the removed pixels infinite noise. In practice, however, this also carries a high computational cost for solving Eq. (24) by CG, as it massively increases the condition number of the coefficient matrix on the left-hand side

(*Seljebotn et al.* 2019). At the same time, the Galactic plane region is critically important to estimate other parameters in the full data model, for instance, the bandpass corrections (*Svalheim et al.* 2023a), and simply excluding these regions entirely from the analysis is therefore both computationally expensive and wasteful in terms of discarding useful information.

A second complication regards degeneracies between the various astrophysical components on small angular scales. As discussed by *Andersen et al.* (2023), the BEYONDPLANCK dataset (comprising *Planck* LFI, WMAP, two HFI channels, and Haslam 408 MHz observations) simply is not able to robustly constrain the astrophysical model on its own on multipoles above $\ell \gtrsim 300$; on these scales, the LFI 30 GHz and WMAP *Ka*-band beams start to drop off exponentially, and their effective signal-to-noise ratio falls quickly. Leaving only intermediate frequencies to constrain the model, one observes a very strong degeneracy between CMB, AME, and free-free emission. To solve this problem, *Andersen et al.* (2023) introduce informative Gaussian priors on the free-free and AME components, effectively using information from *Planck* HFI to constrain the spatial morphology of these components on small angular scales. The impact of these priors on the CMB component are explored in Sect. 3 in this paper.

To simultaneously mitigate both the masking-induced computational expense and the degeneracy challenges, we introduce two small but important additions to the Gibbs chain described in Eqs. (13)–(19) that we refer to as “resampling”. The first step of this process is to run the algorithm as described in the previous sections, but without imposing either a confidence mask or the Gaussian prior on the CMB component. The outputs from this process are thus full-sky CMB and astrophysical component maps, together with a full characterization of the various instrumental parameters. These preliminary CMB maps are, however, not suitable for high-precision temperature-based power spectrum and cosmological parameter analysis because of unmasked foreground residuals and the free-free degeneracy on intermediate scales discussed above; they can, however, be used for large-angle polarization analysis, as free-free emission is not expected to be significantly polarized.

To establish CMB intensity maps that actually can be used for cosmological analysis, we resampled the original chain; that is, for each sample in the main chain, we resampled the CMB component while conditioning on the instrumental and nonlinear astrophysical parameters derived in the first main sampling phase. In practice, this amounts to repeating the sampling steps defined by Eqs. (18) and (19) for each main chain samples, while keeping all other parameters fixed at their main chain value for that same sample. The resulting resampled distribution are therefore marginalized over non-CMB degrees of freedom in a manner analogous to main chain samples. During this process, we make two important changes to the data model: We first apply the confidence mask, as defined above, to suppress the majority of the residual foreground contamination. Secondly, we remove the free-free component in its entirety from the model, leaving only the synchrotron, AME, thermal dust, and point sources to account for any non-masked signal at the unmasked high Galactic latitudes. Since the free-free emission is generally more localized on the sky than synchrotron or thermal dust emission (*Planck Collaboration Int. XLVI* 2016; *Andersen et al.* 2023), it is possible to eliminate most of this signal by masking. On the other hand, the confidence mask does have to be considerably larger than if free-free emission had been explicitly modeled and this is the main reason that our temperature confidence mask, as defined above, has a relatively low accepted sky fraction of only

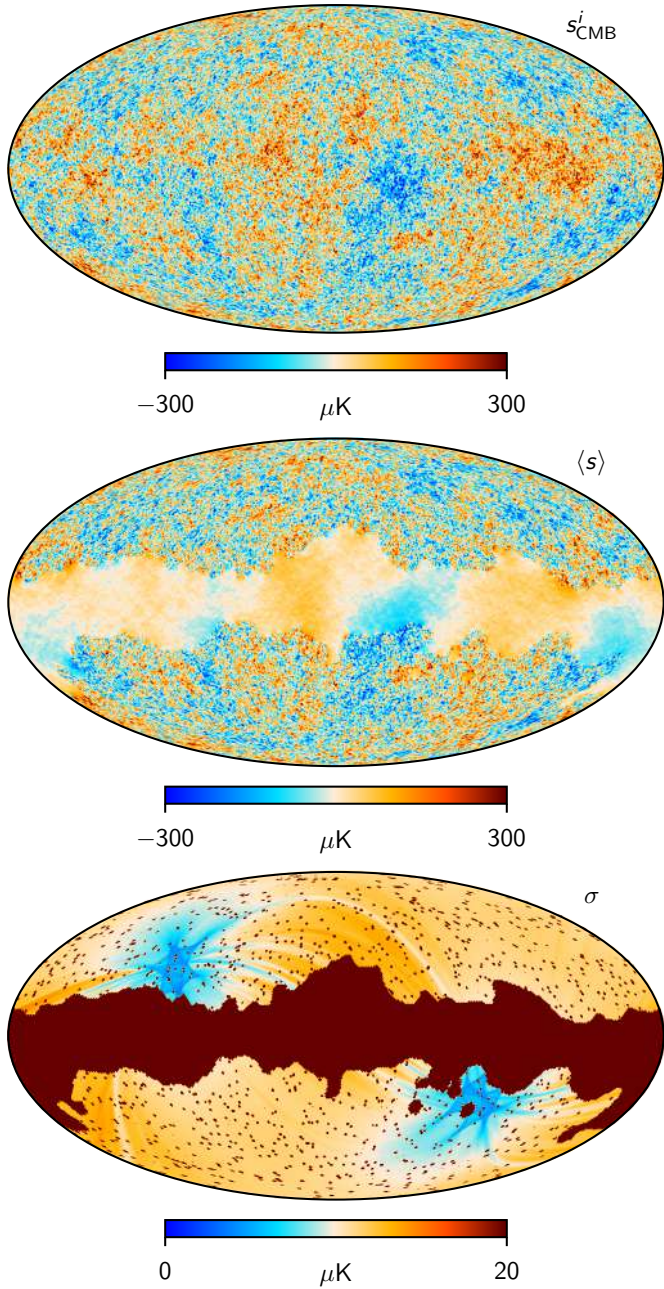


Fig. 2. Full-resolution CMB temperature constrained realization maps. *Top:* single constrained realization, s^i , drawn from $P(s \mid \mathbf{d}, C_\ell, \dots)$. *Middle:* posterior mean map, $\langle s \rangle$, as evaluated from the ensemble of constrained CMB realizations; note that the small-scale signal amplitude inside the mask decreases smoothly to zero with increasing distance from the edge of the mask. *Bottom:* CMB posterior standard deviation map, as evaluated pixel-by-pixel from the ensemble of constrained CMB realizations. This map is dominated by instrumental noise outside the mask and by random fluctuations informed by the assumptions of isotropy and Gaussianity inside the mask. The CMB Solar dipole has been removed from the top two panels.

$f_{\text{sky}} = 0.64$. To account for possible unmasked residual free-free emission at high latitudes, we also resampled the AME component amplitude jointly with the CMB component, such that the resulting AME component at this stage in reality becomes an “AME-plus-free-free” component. This is conceptually similar to the single “low-frequency foreground” component used in the *Planck* 2018 Commander analysis, except that in that case, the

synchrotron emission was also included. We note, however, that this “AME-plus-free-free” component is never used in any further analysis, but it is only a pure phenomenological nuisance parameter as far as the CMB component is concerned. A single resampled CMB constrained realization sample is shown in the top panel of Fig. 2. The middle panel shows the corresponding Wiener filter solution alone, in which structures within the Galactic plane mask may be partially reconstructed due to the assumptions of statistical isotropy and Gaussianity. The bottom panel shows the posterior standard deviation in each pixel.

Finally, we also performed an extra resampling step for the CMB polarization analysis. In this case, we once again conditioned the instrumental and astrophysical parameters from each sample in the main Gibbs chain; however, in this case we performed $N = 50$ additional amplitude-sampling steps for each main sample, as defined by Eq. (18), each of which is computationally much cheaper than a full sample. This resampling step thus involves no fundamental modifications of either the algorithm or data model as such, but is just a computationally convenient way of marginalizing over white noise and thereby converging faster at a modest additional computational cost. A detailed assessment of the impact of this procedure on the convergence of cosmological parameter estimates is presented in [Paradiso et al. \(2023\)](#).

3. Markov chains and correlations

The full BEYONDPLANCK Gibbs sampler and data configuration are summarized in [BeyondPlanck Collaboration \(2023\)](#). The main products from this process are a set of 4000 end-to-end samples evenly distributed over four chains. The first 200 samples in each chain are conservatively rejected as burn-in, although we have not identified strong evidence for non-stationary behavior after the first few tens of samples. A total of 3200 main Gibbs samples were retained for science exploitation and we produced one resampled high- ℓ temperature sample and 50 low- ℓ polarization samples per main Gibbs sample. The production of a single main chain sample requires ~ 164 CPU-h, of which approximately 60% is TOD processing and 40% is component separation, and ~ 1.5 TB RAM. The latter requirement is driven by the necessity of storing the TOD in memory⁴, avoiding the need to reload the TOD from disk at each Gibbs iteration. The resampling process is dominated by the high- ℓ temperature map production and requires an additional ~ 47 CPU-h per sample, while the low- ℓ polarization resampling step requires only a few CPU-h per iteration. The total computational cost of the full analysis is about 800 kCPU-h ([Galloway et al. 2023b](#)). The full sample set is made publicly available through the COSMOGLOBE⁵ web page.

A quantitative assessment of the convergence of the main chain CMB maps is provided by the Gelman-Rubin convergence statistic, R , ([Gelman & Rubin 1992](#)) which measures the ratio between the intra-chain and inter-chain variances, and values of $R < 1.01$ typically indicate good convergence. Figure 3 shows $R - 1$ for the TT power spectrum of main chain CMB map samples. Power spectra have been computed outside the BEYONDPLANCK temperature confidence mask, using PolSpice⁶, a fast pseudo- C_ℓ based estimator which leverages pixel-space correlation functions to speed up computations.

⁴ The full LFI TOD is stored in a compressed format requiring ~ 900 GB, almost an order of magnitude reduction from the ~ 8 TB uncompressed data volume.

⁵ <https://cosmoglobe.uio.no>

⁶ <http://www2.iap.fr/users/hivon/software/PolSpice/>

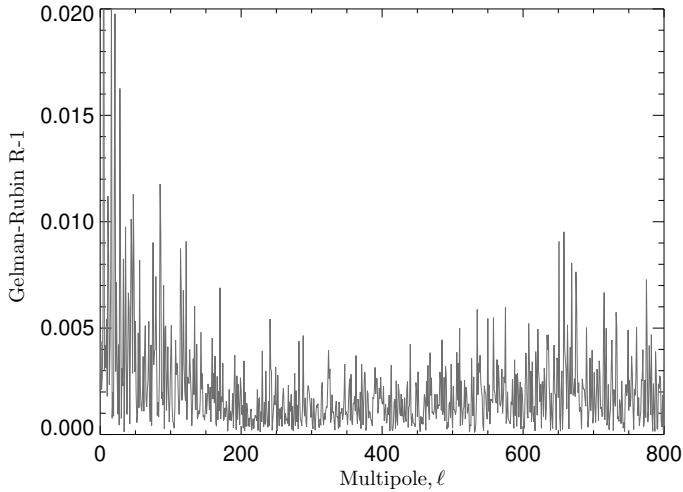


Fig. 3. Gelman-Rubin convergence diagnostic, $R - 1$, for main chain CMB TT power spectrum, computed on the sky fraction allowed by BEYONDPLANCK temperature analysis confidence mask.

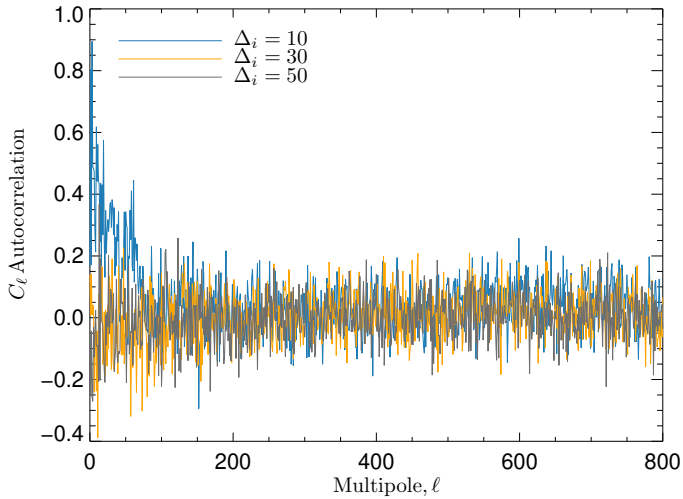


Fig. 4. Main chain CMB TT power spectrum autocorrelation coefficient for different sample separation values, Δ_i , averaged over all chains.

Most multipoles satisfy $R - 1 < 0.01$, with a few exceptions displaying values of $R - 1 \lesssim 0.02$, at the lowest multipoles $\ell < 50$, where pseudo- C_ℓ based methods are suboptimal (e.g., Gerbino et al. 2020).

As an additional diagnostic, we can compute the C_ℓ autocorrelation coefficient as a function of sample separation:

$$A_\ell(\Delta_i) = \frac{\sum_{i=1}^{i=n_{\text{sample}}-\Delta_i} (C_{\ell,i} - \langle C_{\ell,i} \rangle)(C_{\ell,i+\Delta_i} - \langle C_{\ell,i} \rangle)}{\sum_{i=1}^{i=n_{\text{sample}}} (C_{\ell,i} - \langle C_{\ell,i} \rangle)^2}, \quad (29)$$

where the index, i , labels the samples within a single chain after burnin, and $\langle \dots \rangle$ denotes sample averaging within that chain. Figure 4 shows TT autocorrelation coefficients, averaged between all chains, for three different lags, $\Delta_i = 10, 30, 50$. We can see that for $\Delta_i = 10$ the lowest multipoles exhibit a significant amount of correlation, which can exceed 50% in a few isolated cases, decreasing to a few percent level at multipoles $\ell > 100$. Correlation level rapidly decreases with increased sample separation and, for $\Delta_i = 50$, the autocorrelation level drops to $A_\ell = 0.00 \pm 0.07$ at $\ell \in [2, 800]$. We stress that main chain temperature maps are not directly used for cosmological analy-

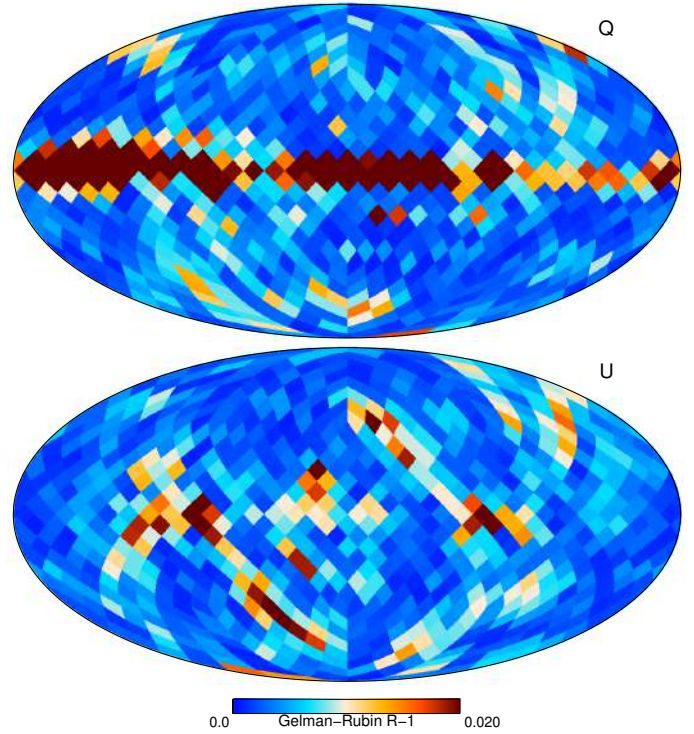


Fig. 5. Gelman-Rubin convergence diagnostic for main chain CMB Q , U maps. Slow convergence pixels in the Galactic plane are excluded by the BEYONDPLANCK confidence mask, while high R stripes outside the plane reflect correlated noise features accounted for by the low-resolution noise covariance matrix (see Sect. 4).

sis, but rather form the basis for the resampling runs responsible for the production of the final CMB maps and parameter estimation, which will be discussed in greater detail in Sect. 4 (also see Paradiso et al. 2023). In addition, cosmological parameters depend on coherent variations over several tens (or hundreds) of multipoles, so each individual C_ℓ has only a minor impact on the end science results.

Planck LFI polarization measurements are strongly dominated by instrumental noise, and CMB E -mode signal-to-noise ratio rapidly drops below 1 at multipoles $\gtrsim 10$. As such, all the cosmologically relevant modes can be properly supported on HEALPIX $N_{\text{side}} = 8$ resolution maps, making it more convenient (and numerically feasible) to analyze BEYONDPLANCK CMB maps directly in pixel space. Figure 5 shows the Gelman-Rubin statistics for the low-resolution Q , U map pixels. Most regions of poor convergence are localized within the Galactic Plane and excluded by the polarization confidence mask. We can also see some stripes of slow convergence at higher Galactic latitudes, due to correlated gain and noise fluctuations, modulated by *Planck* scanning strategy. In the final cosmological analysis, these features are properly accounted for by the low resolution covariance matrix, as discussed in Sect. 4 and in Paradiso et al. (2023). In Fig. 6, we plot the autocorrelation coefficient, defined in a manner analogous to Eq. (29), of low-resolution Q -map pixels, averaged over all chains. With a lag $\Delta_i = 10$ samples, most pixels outside the Galactic Plane show a positive autocorrelation at $\sim 10\%$ level, dropping to a few percent level for $\Delta_i = 50$, with individual pixels showing both positive and negative autocorrelation. The U maps display similar features.

Figure 7 shows a collection of trace plots, namely, parameter values plotted as a function of chain iterations, for various CMB and selected ancillary parameters. The quantities marked with

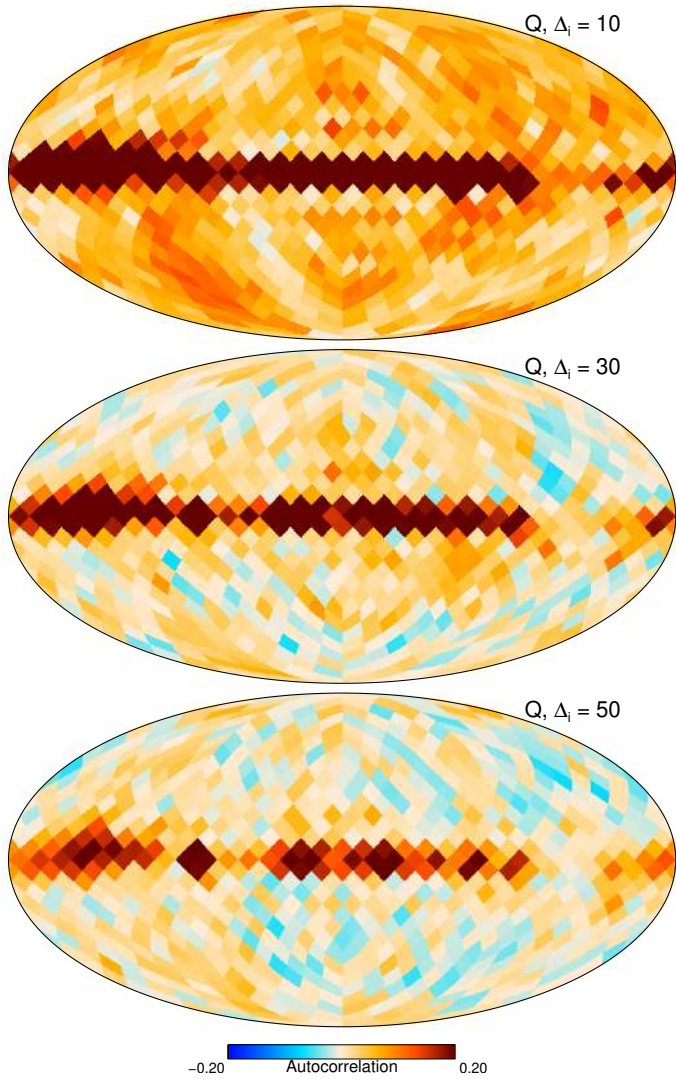


Fig. 6. Main chain CMB Q map pixel amplitude autocorrelation coefficient as a function of sample separation, averaged over all chains. Maps have been downgraded to HEALPix resolution $N_{\text{side}} = 8$, to better highlight correlation structure. *From top to bottom:* the 10, 30, and 50 samples.

subscripts a and b represent sky map pixel values for pixel number 340 and 1960, respectively, in maps downgraded by straight averaging to a HEALPix resolution of $N_{\text{side}} = 16$ with ring ordering. Pixel 340 is located in the top right quadrant at high Galactic latitudes, while pixel 1960 is located near the southern center of the Galactic mask edge. From top to bottom and left to right: plotted quantities are the three Stokes parameters for CMB pixels 340 and 1960; the same for thermal dust emission; the synchrotron spectral index for pixel 1960 in temperature and polarization; the CMB quadrupole spherical harmonic coefficient a_{21} for T , E , and B ; the same for $a_{200,100}$; the three components of the CMB dipole in Cartesian coordinates; the CMB angular temperature power spectrum, D_ℓ , for $\ell = 2, 200$, and 500; and the time-independent radiometer gain fluctuation for the 70 GHz 21M radiometer, Δg_{21M} . Of course, these represent only 26 parameters out of billions, but they still convey some useful intuition regarding the overall behavior of the Gibbs chain as far as the CMB component is concerned.

The first immediate conclusion that can be drawn from these plots at a visual level is that the overall correlation length is

relatively short, and the Markov chain mixing is reasonable, in agreement with the Gelman-Rubin and autocorrelation results discussed above. Furthermore, all chains appear stationary, suggesting that the burn-in samples have been successfully removed. Going slightly deeper into detail, we see that while a_{21}^T appears significantly non-Gaussian, with a pronounced negative tail, I_{CMB} looks more Gaussian and symmetric, although with a longer correlation length. As a result, uncertainties and covariances at low multipoles are generally easier to summarize in pixel space than in harmonic space. Regarding the power spectrum coefficients, D_ℓ , we note that these are not Gaussian-distributed at all, but rather follow an inverse Gamma (or inverse Wishart) distribution, which has a very heavy tail toward positive values at low multipoles. This behavior is clearly seen for D_2^{TT} .

In Fig. 8, we show the corresponding matrix of Pearson's correlation coefficients for each pair of parameters. The lower triangular part shows raw correlations, while the upper triangular part shows correlations after high-pass filtering each Markov chain with a boxcar window of ten samples; the latter highlights white noise correlation structures, while the former includes also long trends.

Overall, most correlations are relatively weak and typically smaller than 5%, while three are very strong. The first is a 60% correlation between the x - and z -components of the CMB Solar dipole. This is caused by the relative orientation of the Galactic plane mask, which directly aligns with the z -component, and the diffuse foregrounds at high latitudes that are anti-symmetric with respect to Galactic longitude $l = 0^\circ$ and are therefore coupled to the x -direction. In contrast, the Galactic plane is symmetric with respect to Galactic longitudes $l = 90^\circ$ and 270° and therefore it is weakly coupled to the y -dipole.

A second strong correlation is between the CMB and dust Stokes Q and U parameters within a single pixel, which reflects the internal degeneracies of our sky model given BEYONDPLANCK data selection. While here we show only the CMB and thermal dust correlations, similar levels of correlations affects also the other sky model components, as discussed in Basyrov et al. (2023) and Andersen et al. (2023).

The third important strong (anti-)correlation in Fig. 8 is seen between the large-scale CMB harmonic $a_{2,1}^T$ (as well as individual temperature pixel values) and the synchrotron spectral index, β_s . While diffuse foregrounds play only a limited role in CMB temperature reconstruction, as measured relative to CMB cosmic variance, and even relatively simple foreground cleaning methods therefore perform very well (e.g., Bennett et al. 2013; Planck Collaboration IV 2020), the same foregrounds still play a very important role as measured relative to the noise level of the experiment and that is what is probed by these correlations. For noise-dominated applications, such as CMB B -mode reconstruction, properly accounting for these foreground uncertainties is therefore key.

The remaining correlations are, as already mentioned, modest – but not negligible, however. For instance, there is a 10% correlation between the two CMB intensity pixels, despite the fact that they are separated by almost 90° on the sky, and not located on the same *Planck* scanning ring. This correlation is therefore due to general global parameters, for instance, the overall instrument calibration and gains, the CMB dipole parameters, and the bandpass corrections.

Moving on from individual pixel values to full sky maps, the two bottom panels in Fig. 9 show pixel-by-pixel cross-correlations between the CMB Stokes Q parameter and the time independent part of the 70 GHz 21M and 21S radiometer gain variations, Δg_{21M} and Δg_{21S} . We note that there is

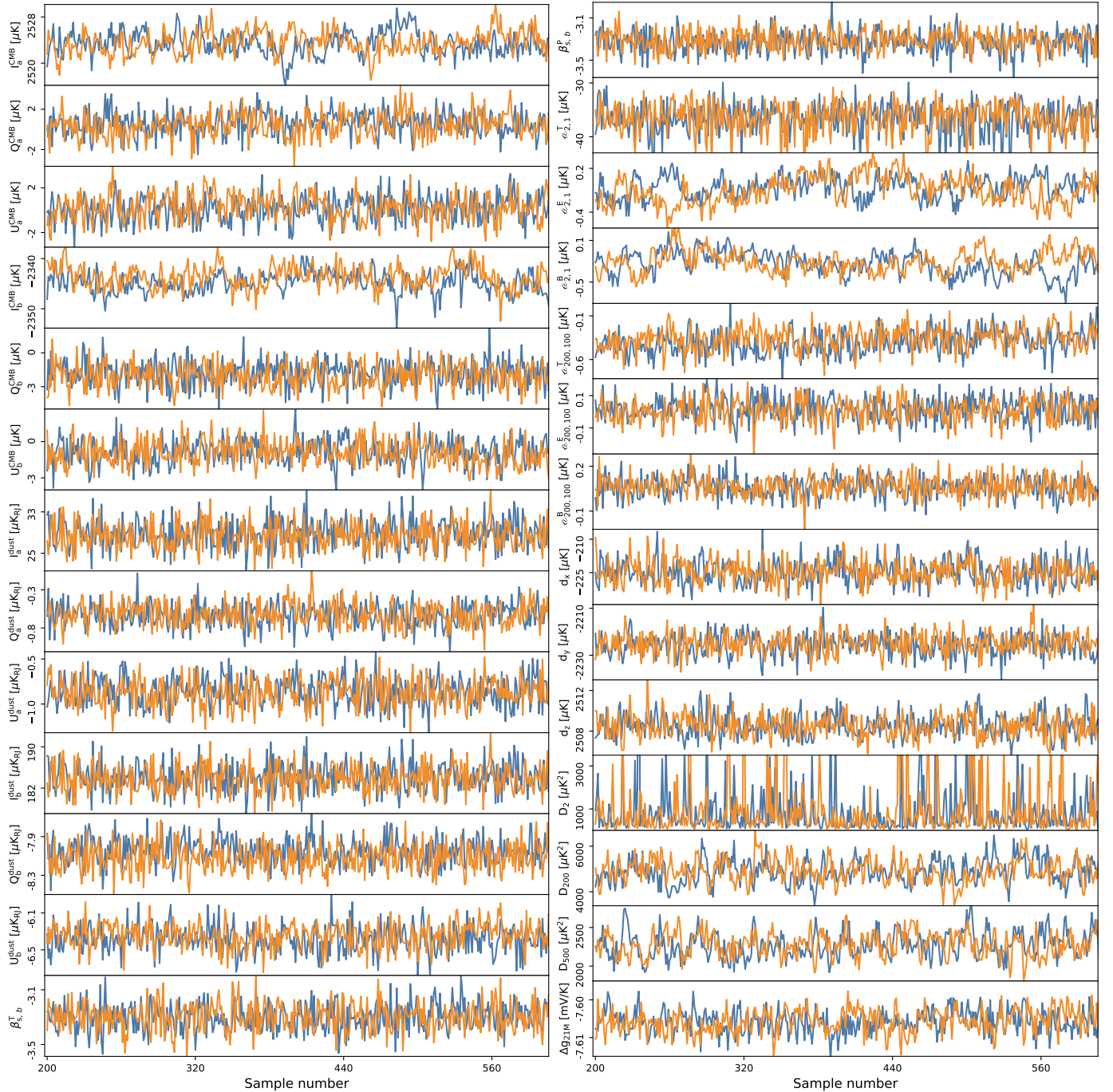


Fig. 7. Trace plots of a set of selected CMB, component separation, and instrument parameters; see main text for full definitions. The different colors indicate independent Gibbs chains, and ‘a’ and ‘b’ subscripts indicate HEALPix pixel numbers ‘340’ and ‘1960’ at resolution $N_{\text{side}} = 16$ in ring ordering, respectively.

nothing special about the 21M and 21S radiometers in this respect, beyond the fact that they are 70 GHz detectors, and the BEYONDPLANCK CMB map is strongly dominated by this frequency channel. In this case, we see coherent large-scale wave patterns at the 5% level, with a wave direction that is loosely aligned with the CMB Solar dipole direction. This pattern is already known, and explained in terms of correlations between inter-detector gain variations and the Solar dipole. An example of this is the LFI gain residual template (Planck Collaboration II 2020) shown in the top panel; the morphology of this template is qualitatively very similar to the correlation structures seen in the BEYONDPLANCK CMB–gain cross-correlations. We also

note that the 21M and 21S correlations are anti-correlated, as expected by the fact that the polarization angles of these two detectors are rotated internally by 90° .

In Fig. 10, we plot the skewness and kurtosis of the CMB Stokes Q and U parameters per HEALPix $N_{\text{side}} = 8$ pixel, which is the same resolution as used by the BEYONDPLANCK low- ℓ likelihood. It is important to note that these estimates do not measure non-Gaussianity of the CMB signal itself, but rather of the uncertainties of the CMB map. Except for a few statistically significant non-Gaussian pixel distributions in the center of the Galactic plane, the skewness and kurtosis maps appear noise dominated, small in amplitude, and statistically isotropic at high Galactic

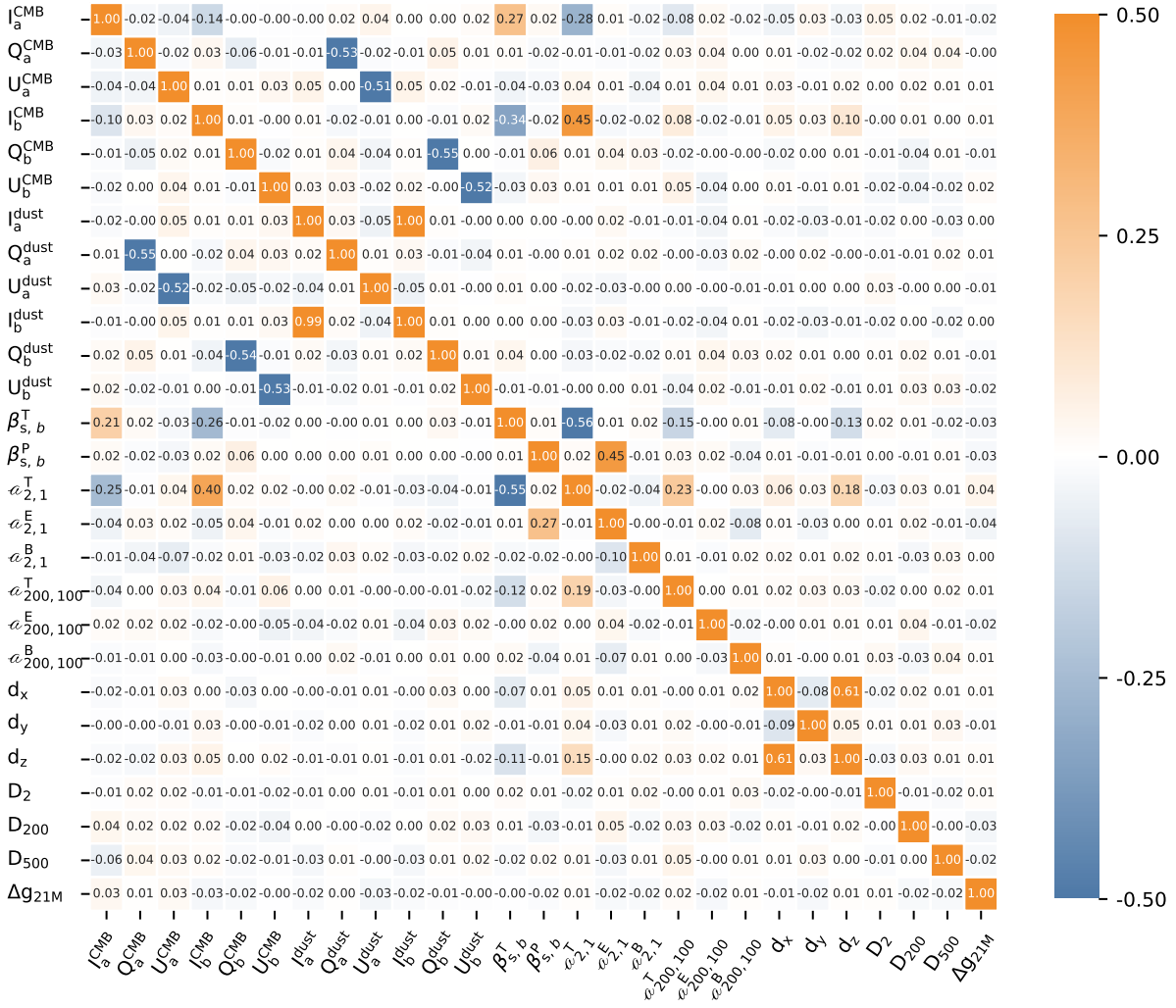


Fig. 8. Correlation coefficients between the same parameters as shown in Fig. 7. The lower triangle shows raw correlations, while the upper triangle shows correlations after high-pass filtering with a running mean with a ten-sample window. For further explanation of and motivation for this filtering, see Andersen et al. (2023).

latitudes. We also note that the standard deviation of excess skewness and kurtosis of a random Gaussian sample with $N_{\text{samp}} \gg 1$ are given by $\sqrt{6/N_{\text{samp}}}$ and $\sqrt{24/N_{\text{samp}}}$, which translate into standard deviations of 0.043 and 0.086, respectively, for $N_{\text{samp}} = 3200$. The observed skewness and kurtosis seen in Fig. 10 are thus both consistent with zero and this suggests that the CMB posterior distribution may be well approximated in terms of a multi-variate Gaussian distribution (Paradiso et al. 2023).

Finally, we conclude this section by measuring the cross-correlation power spectra between the CMB intensity map and each of the four diffuse foregrounds included in the BEYONDPLANCK data model. These cross-correlations are defined by:

$$\rho_\ell^{XY} = \left\langle \frac{C_\ell^{XY}}{\sqrt{C_\ell^{XX} C_\ell^{YY}}} \right\rangle, \quad (30)$$

where $C_\ell^{XY} \equiv \sum_m |a_{\ell m}^X (a_{\ell m}^Y)^*| / (2\ell + 1)$, X denotes CMB and Y any one of the foreground components, and $\langle \dots \rangle$ denotes average over the chain samples. We note that the cross-correlations spectra are computed between the residual maps, $\mathbf{m}_i - \langle \mathbf{m}_i \rangle$, to highlight the impact of model degeneracies, rather than chance

correlations between the components. The results from these calculations are summarized in Fig. 11.

Considering these functions in order from weak to strong correlations, we first note that the thermal dust emission (red curve) is for all practical purposes statistically uncorrelated with the CMB component. It is important to stress that this does not imply that the thermal dust component does not induce foreground modelling errors in the CMB mean map. Rather, it just shows that the thermal dust uncertainty fluctuations are not correlated with the CMB uncertainty fluctuations. The reason for this is that the BEYONDPLANCK analysis (BeyondPlanck Collaboration 2023) relies on the *Planck* PR4 857 GHz sky map as a dust tracer, for which the CMB component is very low, and the only free intensity thermal dust parameter in the entire model is a single full-sky power law spectral index. As such, there is (in practice) no feedback from the CMB to the thermal dust component in the model and we only propagate the thermal dust uncertainties as predicted by the 857 GHz channel to the CMB component, but we do not perform a joint fit. In short, the current analysis effectively assumes that the *Planck* PR4 analysis is accurate as far as thermal dust emission is concerned.

A similar consideration holds true for the synchrotron component (green curve). In this case, the very low frequency of the

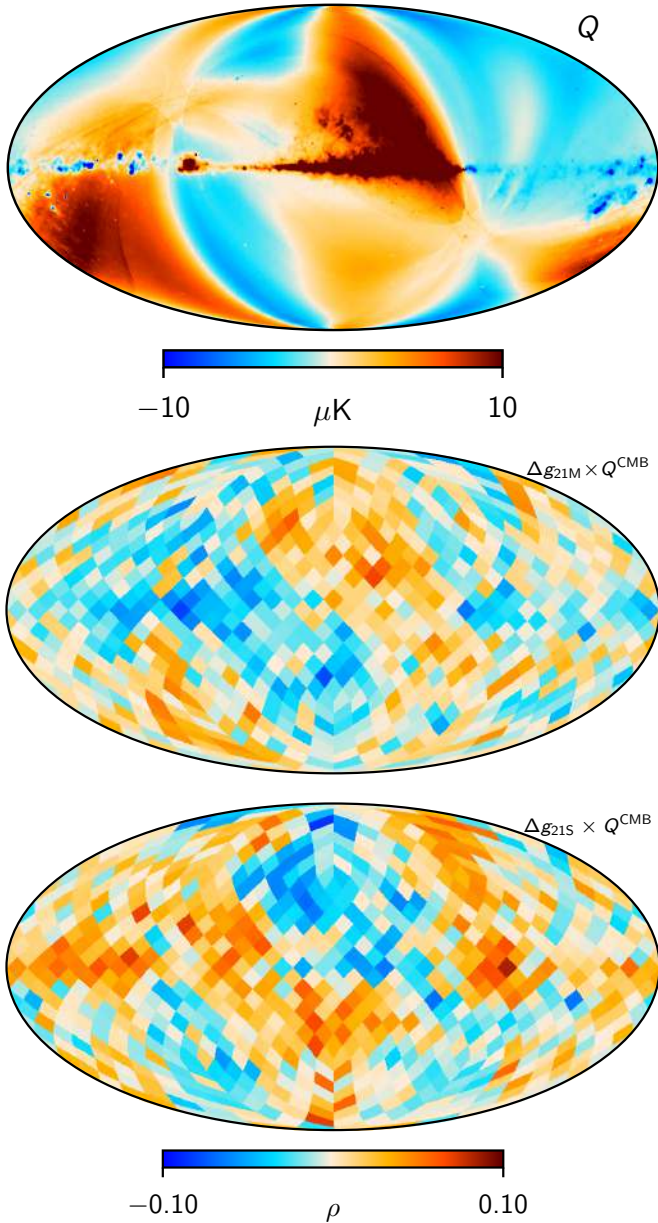


Fig. 9. LFI DPC 30 GHz Stokes Q gain residual template (Planck Collaboration II 2020) shown in the upper panel. Pixel-by-pixel cross-correlation coefficients between CMB Stokes Q and the time-independent absolute gain fluctuation of two 70 GHz radiometers shown in the lower two panels, along with Δg_{21M} (middle panel) and Δg_{21S} (bottom panel).

Haslam 408 MHz effectively decorrelates the synchrotron and CMB components, although not quite as strongly as the 857 GHz map for thermal dust emission. The smooth drop around $\ell \approx 300$ is caused by the algorithmic smoothing prior discussed by Andersen et al. (2023), which suppresses small-scale synchrotron fluctuations.

Significantly higher correlations are seen for the AME component. In this case, there is no single frequency map that gives a clear picture of the component in question, but the spatial structure of AME has to be estimated from the same maps as the CMB itself. The particular data selection adopted for BEYONDPLANCK, which focuses on *Planck* LFI and WMAP measurements between 30 and 70 GHz, leads to correlations at the 15–20% range for AME.

It is important to stress that significant correlations, such as those seen for the AME component, by themselves are no cause for alarm as far as CMB analysis is concerned, as long as the assumed statistical model is correct. In this case, the corresponding uncertainties are fully accounted for in terms of the sample distribution. At the same time, large correlations are nevertheless undesirable, because they make the CMB component susceptible to modeling errors in the correlating component. For BEYONDPLANCK, this is most clearly seen in the free-free component, which, as seen in Fig. 11, is anti-correlated with the CMB component at the 50% level between $\ell \approx 400$ and 600. The reason for this are two-fold. First, the free-free emission scales roughly as ν^{-2} , and therefore falls much more slowly with frequency than both synchrotron emission and AME. Secondly, it is spatially much more localized than either of the other two low-frequency components. The maximum multipole required to model free-free emission without ringing is therefore relatively high ($\ell_{\max} \approx 800$). The only BEYONDPLANCK frequency channels that provide useful information at such small angular scales and high frequencies are, primarily, the LFI 70 GHz channel, and secondarily the LFI 44 GHz and WMAP V-band channel. These are also precisely the same channels that are used to constrain the CMB component – and the two are therefore highly correlated.

As discussed by Andersen et al. (2023), a partial solution to this problem is the introduction of the HFI-dominated spatial free-free prior from Planck Collaboration X (2016). While this is effective at breaking the degeneracy in question, which is necessary for constraining important instrumental parameters such as calibration and bandpasses, it also introduces an uncontrolled level of unmodeled systematic errors and uncertainties. This is both because the LFI and WMAP data were in fact used to generate the free-free prior template in the first place and because of unmodeled systematic and statistical uncertainties in the HFI data. While we consider these unmodeled uncertainties acceptable for instrument modeling, which only depend weakly on the free-free model, they are not acceptable for the CMB component, which is the main scientific product from the entire analysis. This is a main reason for performing the BEYONDPLANCK analysis in a two-step manner, in which the prior-constrained free-free component is included during the main Gibbs analysis, but excluded during the CMB resampling stage, while at the same time applying a large Galactic mask that excludes all regions with statistically significant free-free emission. For further discussion, we refer to Sect. 4.3, Paradiso et al. (2023), and Andersen et al. (2023).

4. CMB maps, covariance matrices, and power spectra

4.1. Posterior mean sky maps

The individual parameter samples discussed in the previous section represent the most fundamental products from the current analysis and we strongly recommend using the set of such individual Gibbs samples for any high-level statistical analysis. That ensemble provides the most convenient approach to fully propagate uncertainties into any given statistic. In order to do so, we must simply analyze all available samples individually, as if they were ideal CMB map estimates, and then report the full distribution as final results. Worked examples of this procedure are given in Sect. 6 for select previously reported low- ℓ anomalies.

Still, for visualization and comparison purposes it is still convenient to consider sample averaged mean and standard deviation

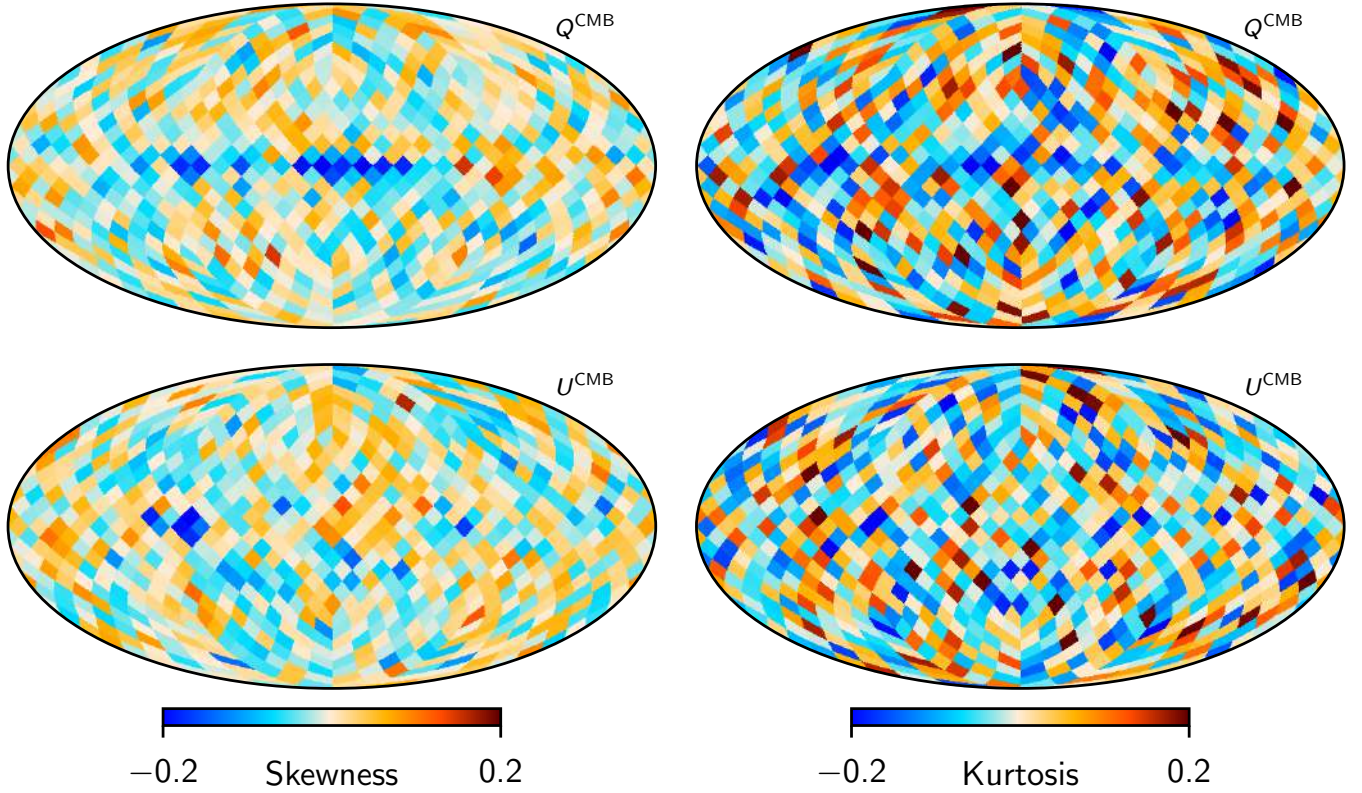


Fig. 10. Skewness (left column) and kurtosis (right column) of the CMB Stokes Q (top row) and U (bottom row) posterior distributions, evaluated pixel-by-pixel at a HEALPix resolution of $N_{\text{side}} = 8$.

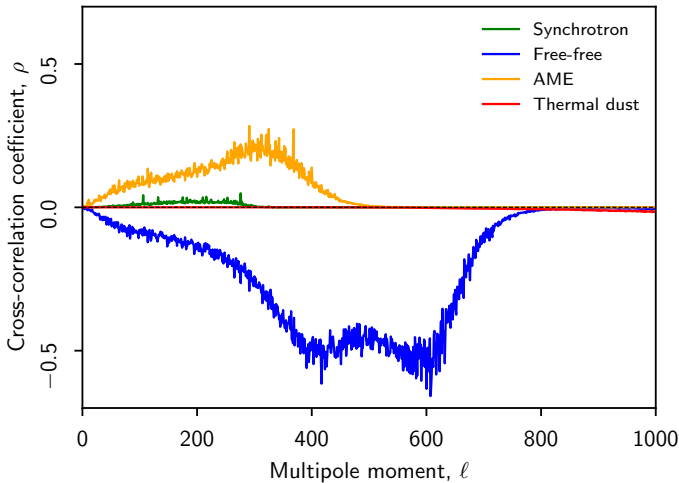


Fig. 11. Cross-correlation spectra between CMB and foreground maps. Colors indicate cross-correlations with synchrotron (green), free-free (blue), AME (orange), and thermal dust emission (red).

maps, which correspond most closely to the best-fit CMB maps derived with traditional pipelines. These are shown in Fig. 12. In this figure, each sample is convolved with a Gaussian azimuthally symmetric beam of $14'$ FWHM for temperature and 1° FWHM for polarization, before projecting the data onto sky maps. The first 200 samples from each chain are conservatively discarded as burn-in, leaving a total of 3200 samples for actual analysis. The mean and standard deviation is then evaluated pixel-by-pixel from these samples. We note that the CMB Solar dipole has been

removed from the temperature maps in these plots (this component is discussed separately in the next section).

Starting with the temperature mean map in the top left panel, we see that this CMB map is visually similar to the Commander CMB map presented by in the *Planck* 2015 analysis (Planck Collaboration X 2016). At high Galactic latitudes, the familiar isotropic CMB fluctuations are visually obvious, while at low Galactic latitudes, there is a clear negative foreground imprint. This is due to over-subtraction of thermal dust and free-free emission and it can be removed through detailed foreground modeling that also includes *Planck* HFI observations; see Planck Collaboration IV (2020) for a Commander-based analysis that successfully eliminates this effect. For the current analysis, which does not include HFI observations and only fit the thermal dust SED with a single spectral index, β_d , across the full sky, this foreground leakage represents the main limiting effect at low Galactic latitudes – and this clearly demonstrates why a large Galactic mask is needed.

The upper right panel shows the corresponding standard deviation map and we see that this is dominated by three main effects. At high latitudes, the dominant feature is the *Planck* scanning strategy, and individual features are associated with the white noise distribution of the 70 GHz LFI channel (Basyrov et al. 2023). We may also see a number of bright dots, corresponding to individual point sources, as described by Eq. (8). At low Galactic latitudes, the uncertainties are dominated by diffuse foregrounds, and the morphology is visually dominated by free-free and thermal dust emission (Andersen et al. 2023).

The two bottom rows show the same for the Stokes Q and U polarization maps. In these cases, the mean maps are visually dominated by white noise over most of the sky, as evidenced

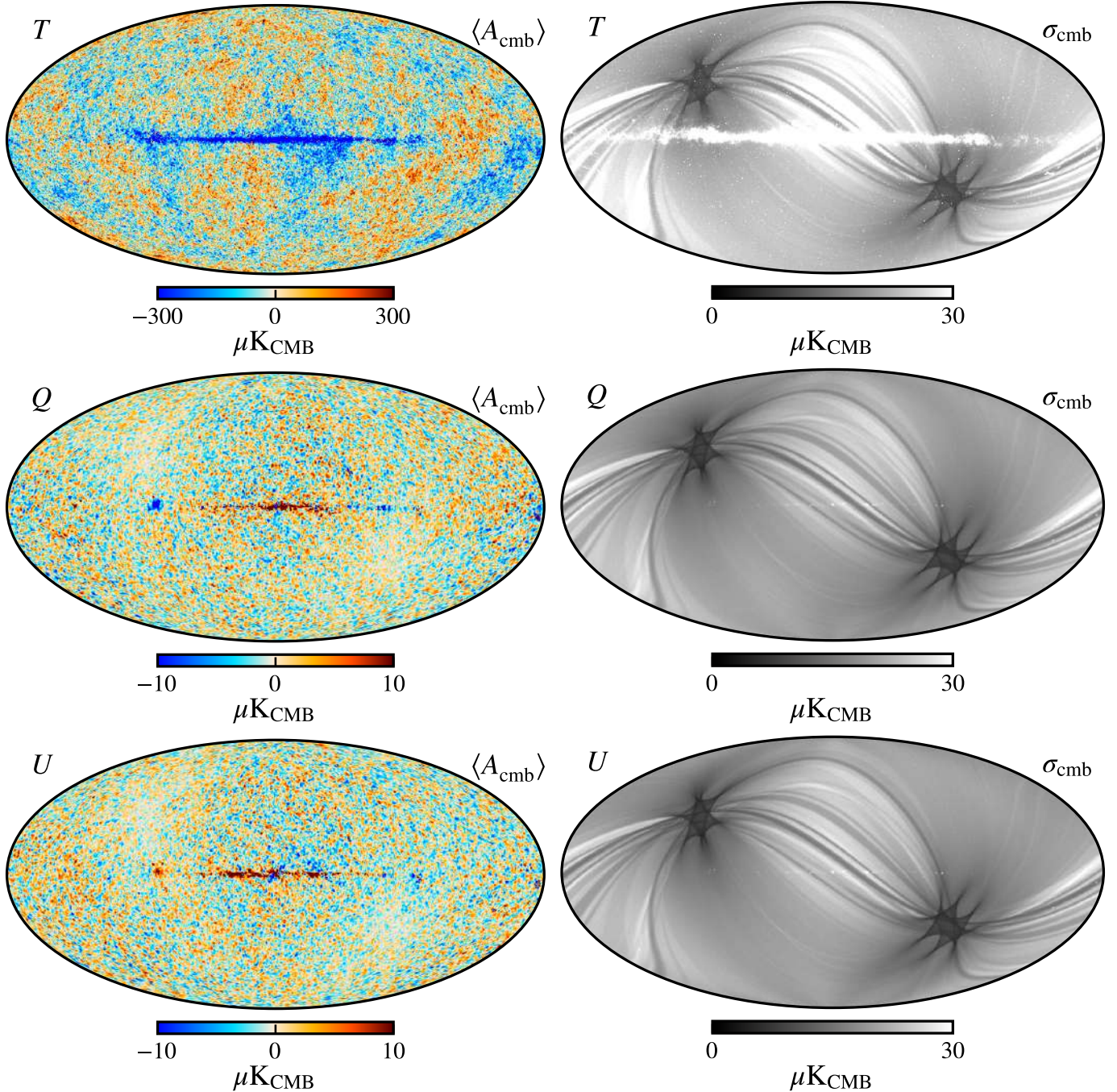


Fig. 12. BEYONDPLANCK posterior mean (left column) and standard deviation (right column) CMB fluctuation maps. Rows show from top to bottom the temperature and Stokes Q and U parameters, respectively. The temperature maps are smoothed to $14'$ FWHM resolution, while the polarization maps are smoothed to 1° FWHM.

by the fact that one may see the Ecliptic pole regions also in the mean maps. Of course, this is fully expected, given that the average standard deviation per pixel is about $20 \mu\text{K}$, while the expected CMB signal for an ideal ΛCDM CMB map smoothed to 1° FWHM is $\lesssim 3 \mu\text{K}$. Thus, the signal-to-noise ratio is less than 0.5 per pixel.

The only obvious visually recognizable features are Galactic plane residuals with an alternating sign, which is a classic signature of temperature-to-polarization leakage from bandpass mismatch (Svalheim et al. 2023a). This is, however, confined to a narrow region of less than 1% of the full sky. The polariza-

tion CMB confidence mask shown in the bottom panel Fig. 1 is more than sufficient to eliminate these residuals from higher-level analysis.

In Fig. 13 we show difference maps between the BEYONDPLANCK posterior mean CMB temperature map and the four foreground-reduced CMB maps presented by Planck Collaboration IV (2020), generated by Commander (Eriksen et al. 2008), NILC (Basak & Delabrouille 2012, 2013), SEVEM (Leach et al. 2008; Fernández-Cobos et al. 2012), and SMICA (Cardoso et al. 2008), respectively. The gray region indicates the BEYONDPLANCK temperature CMB confidence

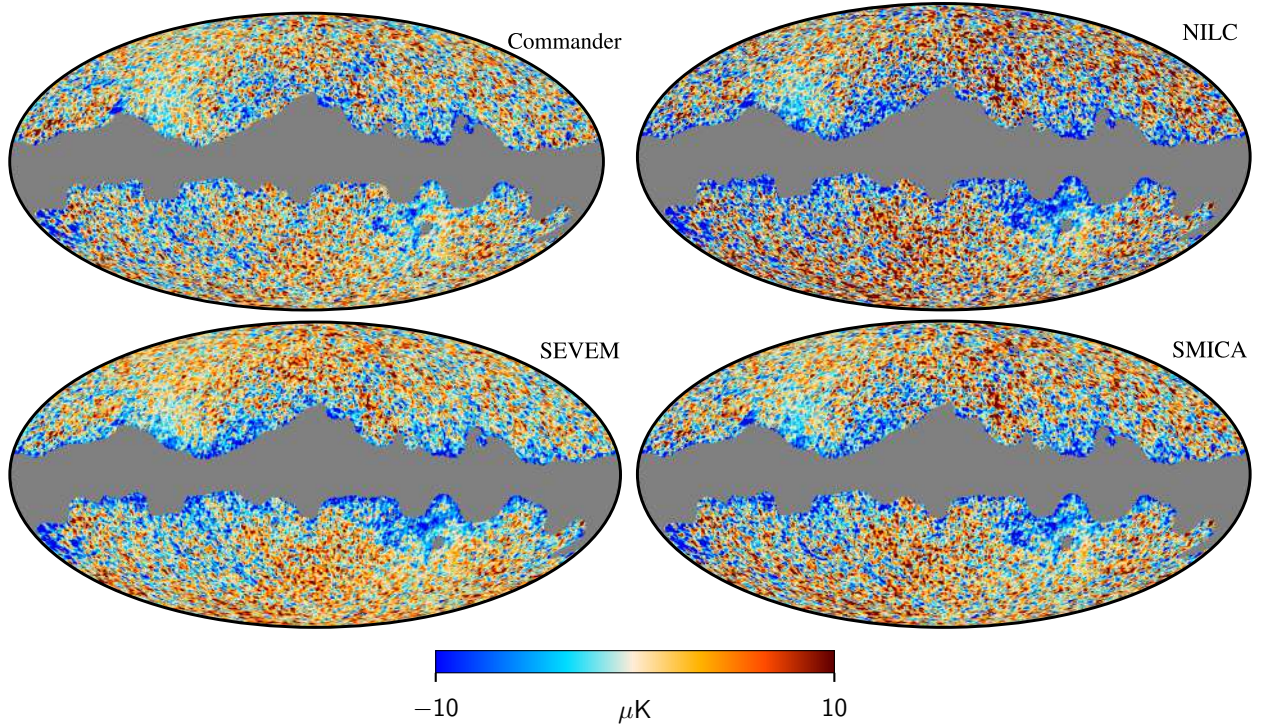


Fig. 13. Difference maps between the BEYONDPLANCK CMB temperature map and those derived from the full *Planck* 2018 data set (Planck Collaboration IV 2020). From left to right and from top to bottom, the various panels show differences with respect to Commander, NILC, SEVEM, and SMICA. All maps are smoothed to a common angular resolution of 1° FWHM.

mask, and a constant offset has been removed from each map outside this mask. First, we note that the color range is $\pm 10 \mu\text{K}$, which is the same range as was used in Figs. 6 and 7 of Planck Collaboration IV (2020) to show differences between the 2015 and 2018 CMB maps and internally among the four *Planck* component separation algorithms. As such, the BEYONDPLANCK CMB map agrees about as well with either of those maps as the *Planck* maps do internally. However, a closer comparison of our Fig. 13 with their Fig. 7 reveals two important differences, namely: a large white noise contribution at high Galactic latitudes and a blue edge around the Galactic plane mask. Both of these effects have fundamentally the same explanation, namely that the current analysis does not involve the CMB-dominated and high-sensitivity HFI frequency channels, and the current map therefore has both higher noise and more free-free and thermal dust contamination. The latter of these effects dictates our larger confidence mask for high-level analysis.

4.2. Low- ℓ polarization power spectrum

We now turn our attention to CMB power spectrum estimation, starting with low- ℓ polarization. For this task, we adopted the well-established machinery of multivariate Gaussian likelihood estimation (e.g., Tegmark et al. 1997; Page et al. 2007; Planck Collaboration V 2020; Gjerløw et al. 2015), and mapped out the following distribution with respect to C_ℓ ,

$$P(C_\ell | \hat{s}_{\text{CMB}}) \propto \frac{\exp\left\{-\frac{1}{2} \hat{s}_{\text{CMB}}^t [\mathbf{S}(C_\ell) + \mathbf{N}]^{-1} \hat{s}_{\text{CMB}}\right\}}{\sqrt{|\mathbf{S}(C_\ell) + \mathbf{N}|}}, \quad (31)$$

where \hat{s}_{CMB} denotes the posterior mean CMB map, \mathbf{N} is the corresponding noise covariance matrix, and $\mathbf{S}(C_\ell)$ is the signal covariance matrix, which is fully defined by the angular power

spectrum. For a detailed review of the implementation used in BEYONDPLANCK, we refer to Paradiso et al. (2023).

The main scientific goal of the entire BEYONDPLANCK framework is precisely to derive estimates of \hat{s}_{CMB} and \mathbf{N} for which astrophysical and instrumental systematic effects are fully marginalized over. Given the Gibbs samples described above, this may be done very conveniently as follows:

$$\hat{s}_{\text{CMB}} = \langle s_{\text{CMB}}^i \rangle, \quad (32)$$

$$\mathbf{N} = \left\langle \left(s_{\text{CMB}}^i - \hat{s}_{\text{CMB}} \right) \left(s_{\text{CMB}}^i - \hat{s}_{\text{CMB}} \right)^t \right\rangle, \quad (33)$$

where i indicates sample number, and brackets denote averages over all available Monte Carlo samples. As described by Paradiso et al. (2023), we evaluate both these quantities at a HEALPix resolution of $N_{\text{side}} = 8$ after smoothing the temperature component to 20° FWHM.

The bottom row of Fig. 14 shows a slice through \mathbf{N} , centered on the Stokes Q pixel marked in gray in the upper right quadrant. This plot effectively summarizes all the various systematic corrections described in Sect. 2 to the extent that they are significant for large-scale polarization reconstruction. For comparison, the top row shows the corresponding CMB covariance matrix slice computed by the *Planck* DPC Planck Collaboration II (2020). We note that this was evaluated at $N_{\text{side}} = 16$ and also that the DPC analysis applied an additional cosine smoothing kernel not used by BEYONDPLANCK.

Comparing the BEYONDPLANCK and DPC covariance matrices is useful for building intuition regarding these products. First, we note that the BEYONDPLANCK covariance appears noisier than the DPC matrix. This is due to the fact that it is constructed by Monte Carlo sampling as opposed to analytic calculations. A computational disadvantage of the sampling approach, relative to the analytic approach, is that any high-level

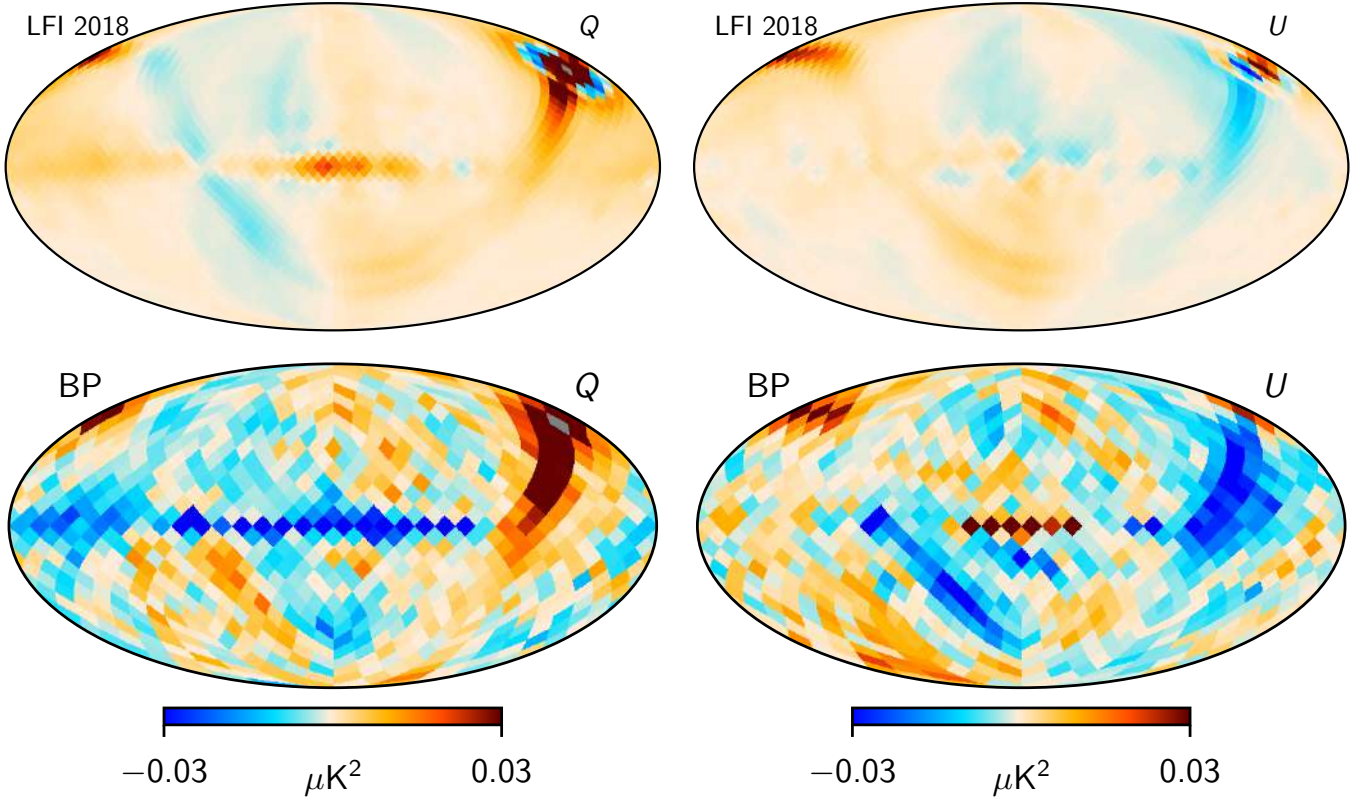


Fig. 14. Single column of the low-resolution CMB noise covariance matrix, as estimated by the LFI DPC (*top row*) and BEYONDPLANCK (*bottom row*). The column corresponds to the Stokes Q pixel marked in gray, which is located in the top right quadrant near the ‘ Q ’ label. We note that the DPC covariance matrix is constructed at $N_{\text{side}} = 16$ and includes a cosine apodization filter, while the BEYONDPLANCK matrix is constructed at $N_{\text{side}} = 8$ with no additional filter.

product derived from the covariance matrix must be accompanied by a corresponding convergence analysis that verifies that the final result is robust with respect to the number of samples; for BEYONDPLANCK, this is done explicitly for the optical depth of reionization, τ , by [Paradiso et al. \(2023\)](#).

However, this minor disadvantage is more than compensated for by the fact that the sampling approach is able to jointly account for many more systematic effects than the analytic approach, and this is clearly seen in [Fig. 14](#): while the DPC matrix only models correlated noise (seen as the ring passing through the gray pixel) and simple template-based foreground corrections, the BEYONDPLANCK matrix additionally accounts for absolute detector calibration differences (seen as large-scale red and blue regions aligned along the Solar CMB dipole direction; [Gjerløw et al. 2023](#); [Basyrov et al. 2023](#)); time-dependent gain fluctuations (seen as additional power along the scanning ring passing through the gray pixel; [Gjerløw et al. 2023](#)); and bandpass leakage (seen as the sharp Galactic plane; [Svalheim et al. 2023a](#)). There are also many other effects that are not as visually obvious, but they still contribute to the final results, such as spatially varying foreground spectral indices ([Andersen et al. 2023](#); [Svalheim et al. 2023b](#)) and time-dependent noise power spectral density parameters ([Ihle et al. 2023](#)). Propagating all these effects analytically into a final joint covariance matrix can be for all practical purposes impossible, while with the novel sampling approach introduced here it is quite straightforward.

[Figure 15](#) shows the corresponding noise-weighted (or “whitened”) posterior mean map, $N_{\text{CMB}}^{-1/2} s_{\text{CMB}}$; when plotted directly in terms of s_{CMB} , the maps are dominated by the *Planck*

scanning pattern and poorly constrained large-scale modes, which complicates the visual interpretation of actually statistically significant features. We note that the color scale ranges over $\pm 3\sigma$. Overall, we see that these maps appear noise dominated, with most pixels having values below 2σ . However, there are a handful of saturated pixels as well, in particular close to the Galactic plane and near the Orion complex (lower right quadrant). Most likely, these are due to unmodeled foreground errors, and should in principle be removed. However, since they are isolated, they only contribute with high- ℓ power, well above $\ell \gtrsim 10$, and they are therefore of minor concern for the current low- ℓ focused analysis; [Paradiso et al. \(2023\)](#) explicitly shows that all main large-scale polarization results are stable with respect to mask variations, from $f_{\text{sky}} \approx 0.25$ to 0.75. We do also see some fainter coherent structures on larger angular scales, but these are all well below 1.5σ . Some of those structures are real CMB signal and some are just coherent large-scale noise fluctuations generated by the same effects as are seen in the covariance matrix slices in [Fig. 14](#). As reported by [Paradiso et al. \(2023\)](#), the total signal-plus-noise χ^2 has a probability-to-exceed of 32% when evaluated for the best-fit Λ CDM power spectrum with $\tau = 0.066 \pm 0.013$, which indicates that the data are fully consistent with the model.

[Figure 16](#) compares the low- ℓ BEYONDPLANCK power spectra with corresponding results reported by *Planck* ([Planck Collaboration V 2020](#)) and WMAP ([Hinshaw et al. 2013](#)). The BEYONDPLANCK constraints shown here is computed by slicing the full probability distribution in [Eq. \(31\)](#) ℓ -by- ℓ , while fixing all other multipoles at their Λ CDM bestfit value ([Planck Collaboration VI 2020](#)); error bars indicate asymmetric

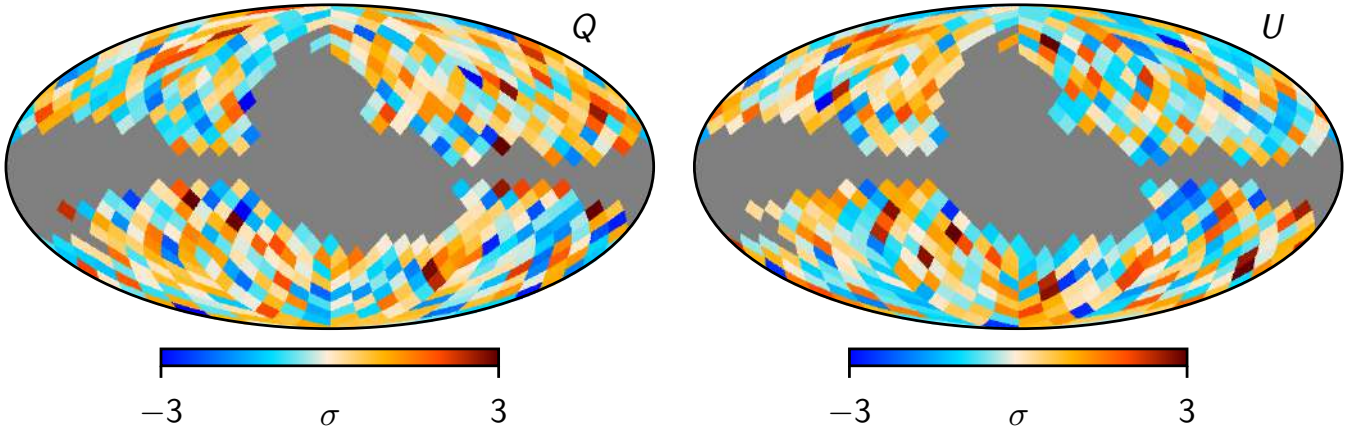


Fig. 15. BEYONDPLANCK low-resolution and “whitened” CMB polarization map, as defined by $N_{\text{CMB}}^{-1/2} s_{\text{CMB}}$ at a HEALPix resolution of $N_{\text{side}} = 8$ and masked with the BEYONDPLANCK analysis mask. *Left and right panels* show Stokes Q and U parameters, respectively, and the color scales span $\pm 3\sigma$.

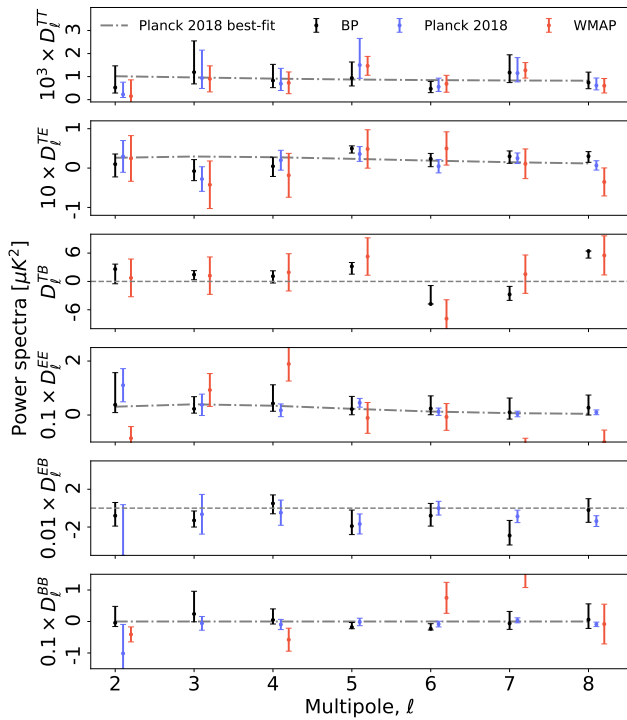


Fig. 16. Comparison between low- ℓ angular CMB power spectra, as derived by the Planck Collaboration using both LFI and HFI data (blue points; Planck Collaboration V 2020), by the WMAP team using just WMAP data (red points; Hinshaw et al. 2013), and by BEYONDPLANCK using both LFI and WMAP data (black points; this work). Thin black lines indicate the Planck 2018 best-fit Λ CDM spectrum (Planck Collaboration VI 2020). The BEYONDPLANCK data points are evaluated by conditionally slicing the posterior distribution ℓ -by- ℓ with respect to the best-fit Λ CDM model, by holding all other multipoles fixed at the reference spectrum while mapping out $P(C_\ell | \mathbf{d})$, to visualize the posterior structure around the peak. For WMAP, the reported BB octopole amplitude is $D_3^{BB} = 1.12 \pm 0.03$, which is outside the plotted range.

68% confidence ranges. Overall, we find good agreement between BEYONDPLANCK, Planck, and WMAP. For TT , we see that the BEYONDPLANCK uncertainties are generally somewhat larger than either of the other two, and that is due to the larger analysis mask. The most notable multipole in this spectrum is

$\ell = 2$, with a peak value of $526 \mu\text{K}^2$, which is substantially higher than the typical values of about $200 \mu\text{K}^2$ reported previously. However, the reason for this is algorithmic in nature, and driven by our conditioning on Λ CDM TE and EE in this particular plot; when marginalizing over polarization, we do recover a quadrupole amplitude of $181 \mu\text{K}^2$, fully consistent with previous results; for further discussion of this multipole, see Sect. 6.1.

For both TE and EE , the most notable feature is that our uncertainties fall between Planck and WMAP in magnitude, which is expected given that the current analysis include both WMAP and LFI data, but not HFI. For BEYONDPLANCK, the most significant outlier in the full set of results is at $\ell = 8$ in TB . Planck has not publicly released TB measurements for the default HFI cross-spectrum based pipeline, while the LFI pixel-base results show a qualitatively similar outlier at that multipole, but with a lower statistical significance. We note, however, that the full probability distribution for this multipole is highly asymmetric, and a full inspection shows that for BEYONDPLANCK this is discrepant with respect to Λ CDM at the 3σ level, with a PTE of 0.2%. The probability of having one such outlier among 49 measurements by random chance is 9%. This multipole may thus provide some slight evidence for residual systematics, for instance associated with the saturated pixels in Fig. 15, but the statistical significance is low. In the case of EB , WMAP does not report any results, while both BEYONDPLANCK and Planck are consistent with zero, as they are for BB . In the latter case, WMAP do report results, but with large error bars; note that $\ell = 3$ and $\ell = 5$ fall far outside the plotted range.

4.3. High- ℓ temperature power spectrum

Next, we consider the high- ℓ temperature power spectrum, and in this case we employ the Gaussianized Blackwell-Rao (GBR) estimator (Chu et al. 2005; Rudjord et al. 2009; Planck Collaboration V 2020) to map out the posterior distribution; for specific details on the BEYONDPLANCK implementation of this estimator, see Paradiso et al. (2023). In short, this estimator is defined by averaging the inverse Gamma distribution, which is the appropriate distribution for the ideal CMB sky:

$$P(C_\ell | s^{\text{CMB}}) = \sum_{i=1}^{n_{\text{samp}}} \frac{\exp\left(-\frac{2\ell+1}{2} \frac{\sigma_i^2}{C_\ell}\right)}{|C_\ell|^{\frac{2\ell+1}{2}}}, \quad (34)$$

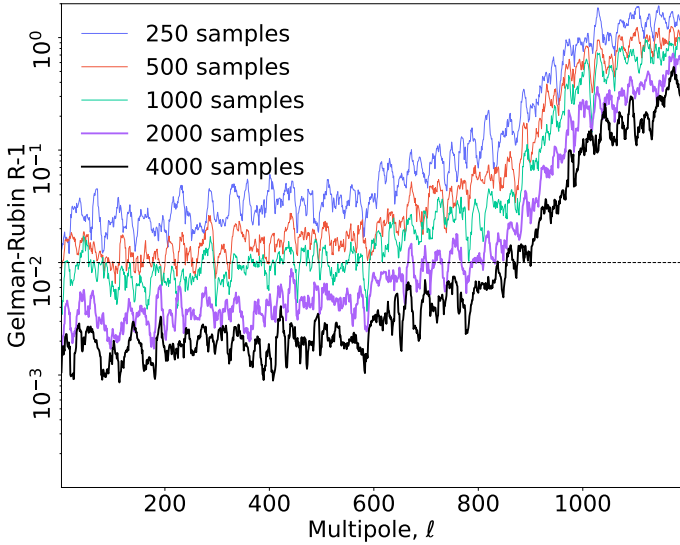


Fig. 17. Gelman-Rubin convergence statistic for the BEYONDPLANCK TT angular power spectrum, as evaluated from four independent σ_ℓ chains. The various curves show results for different total number of samples included in the analysis. A value lower than 1.01 (dotted line) typically indicates good convergence. Accordingly, BEYONDPLANCK multipoles above $\ell \sim 800$ should be acceptable for parameter estimation. However, in the current paper, we conservatively include only modes $\ell \leq 600$ in the cosmological analysis, see text.

over all available Monte Carlo samples, where σ_ℓ^i is the measured full-sky power spectrum of sample i . The resulting marginal distribution, $P(C_\ell | \mathbf{d}) \approx \langle P(C_\ell | s^{\text{CMB}}) \rangle$, is then Gaussianized through a nonlinear mapping, $x_\ell(C_\ell)$, by matching percentiles to a standard normal distribution,

$$\int_0^{C_\ell} P(C'_\ell | \mathbf{d}) dC'_\ell = \frac{1}{2} \left(1 + \text{erf} \left(\frac{x_\ell}{\sqrt{2}} \right) \right), \quad (35)$$

and the final likelihood expression takes the following form,

$$P(C_\ell | \mathbf{d}) \approx \left(\prod_\ell \frac{\partial C_\ell}{\partial x_\ell} \right)^{-1} e^{-\frac{1}{2} (x-\mu)^T C^{-1} (x-\mu)}, \quad (36)$$

where the first factor denotes the Jacobian resulting from the change-of-variables.

This expression formed the basis of the default low- ℓ temperature likelihood in both the *Planck* 2015 and 2018 data releases (Planck Collaboration XI 2016; Planck Collaboration V 2020) for $\ell \leq 30$, and was in the latter also used as an experimental likelihood up to $\ell \leq 250$. The main limitation from extending it to even higher multipoles stemmed from the fact that the samples that defined the *Planck* GBR estimator were computed from foreground-cleaned CMB maps, and those have effectively smoothed white noise contributions which are difficult to describe accurately at high multipoles. In contrast, the novel BEYONDPLANCK approach generates the samples from full-sky foreground-subtracted frequency maps, which do have unsmoothed white noise contributions. On the one hand, this implies that these power spectrum estimates are conceptually different from the pseudo- C_ℓ based one of main chain maps (discussed in Sect. 3), which include a noise contribution and whose evaluation required the application of a confidence mask to exclude foreground residuals. On the other hand, there is no noise modeling limitation associated with the new implementation, and the GBR estimator can therefore in principle be used to arbitrary high multipoles.

In practice, however, the effective range of the GBR estimator is limited by Monte Carlo convergence. This is illustrated in Fig. 17, which shows the Gelman-Rubin statistic for each power spectrum multipole for different numbers of Monte Carlo samples. In this figure, we see that R increases rapidly above $\ell \approx 800$, where the CMB signal-to-noise ratio of the BEYONDPLANCK dataset falls below unity. This behavior is theoretically well understood (e.g., Eriksen et al. 2004a), and may be solved by introducing additional sampling steps (Jewell et al. 2009; Racine et al. 2016); implementing this in the latest version of Commander is currently ongoing. For now, we conservatively restrict the range for which this estimator is used to $\ell \leq 600$.

The final BEYONDPLANCK temperature power spectrum is shown in Fig. 18, together with *Planck* 2018 and WMAP. The top panel shows the full power spectrum; the middle panel shows the difference with respect to the best-fit *Planck* 2018 ΛCDM spectrum in units of each estimate's respective σ_ℓ , and the bottom panel shows the fractional difference with respect to ΛCDM in units of percent. Overall, we see that all three analyses agree very well. For BEYONDPLANCK, the most significant outliers is $\ell = 416$, which is anomalous at the 4σ level; we note that this multipole is also low in the HFI-dominated *Planck* 2018 spectrum, although at a slightly lower significance of about 3σ . The probability of having one such outlier among 599 trials by random chance is about 8%.

Before concluding this section, we return to the issues of strong free-free correlations and resampling discussed in Sect. 2.5. Specifically, Fig. 19 compares the angular power spectra derived from chains that samples free-free emission per pixel (red curve) with the baseline approach that excludes this component (for plotting purposes, spectra have been smoothed with a Gaussian kernel to suppress the cosmic variance scatter). Here, we see a highly statistically significant excess between $\ell = 300$ and 600, with a general behavior that overall mirrors the CMB-versus-free-free cross-correlation shown in Fig. 11. The explanation for this behavior is quite simple: taking into account the beam sizes and white noise levels of the data involved in the BEYONDPLANCK analysis, by far most of the constraining power for $\ell \gtrsim 300$ comes from the LFI 70 GHz channel alone, with only slight additional support from the LFI 44 GHz WMAP Q - and V -bands. This leaves the free-free and CMB components highly degenerate. At the same time, accurate modeling of free-free emission on larger scales is key for obtaining a robust calibration and foreground model. As a temporary solution to this problem, the current main analysis adopts the (HFI-dominated) *Planck* 2015 free-free map as a spatial template prior (Andersen et al. 2023). While this prior does regularize the foreground fit as such, it also biases the CMB component at intermediate angular scales. For this reason, we only include the prior-constrained free-free component while estimating the instrumental and astrophysical parameters, but exclude it when estimating the final CMB parameters. This issue will of course be resolved in a future Bayesian end-to-end analysis that jointly analyzes both LFI and HFI data from scratch.

5. CMB Solar dipole

We now turn our attention to the CMB Solar dipole. In the BEYONDPLANCK framework, this component is in principle estimated on completely the same footing as any other mode in the CMB sky, and is represented in terms of three spherical harmonic coefficients in s_{CMB} . No special-purpose component separation algorithms were applied to derive the CMB dipole, nor did any individual frequency play a more important role than

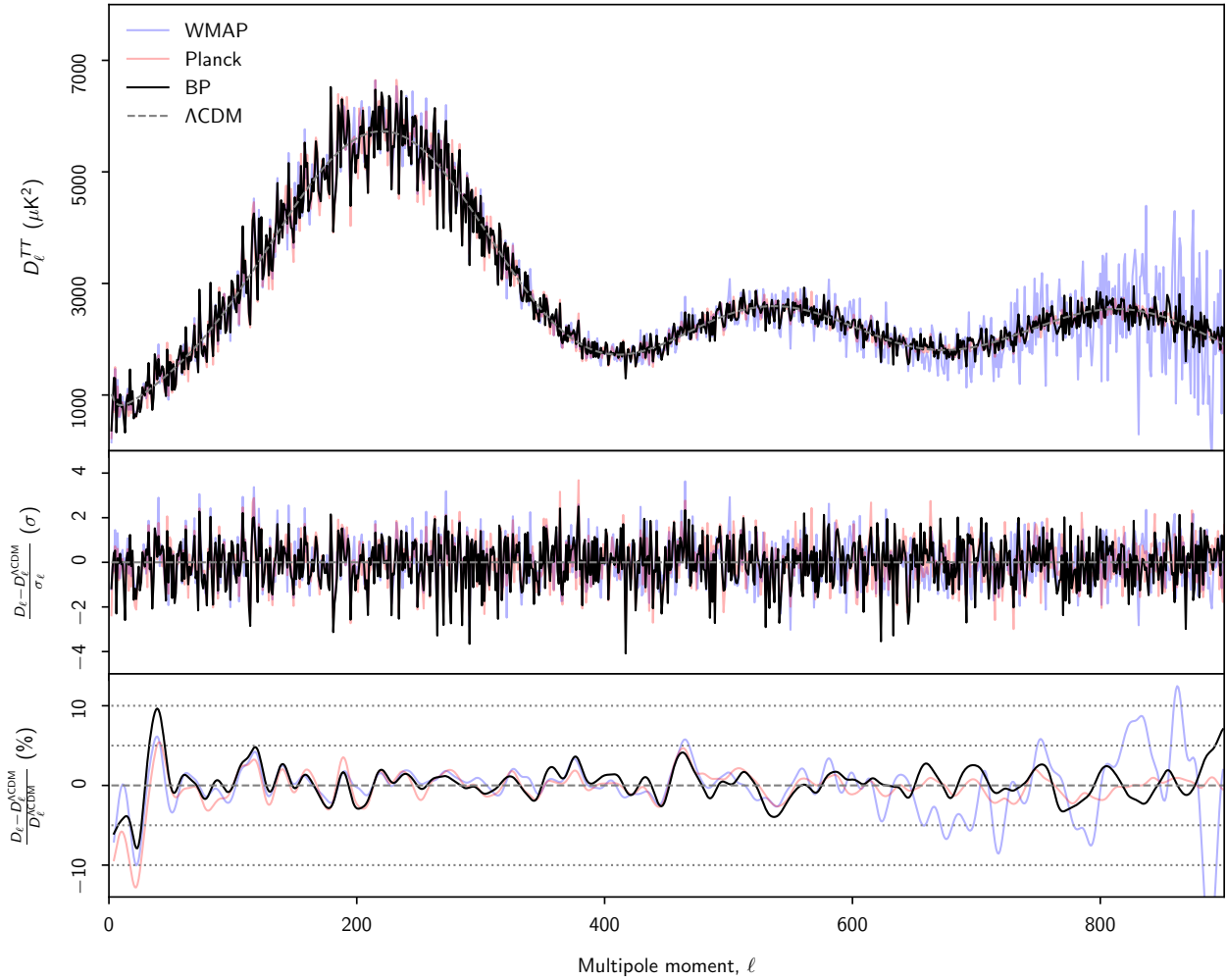


Fig. 18. Angular CMB temperature power spectrum, D_ℓ^{TT} , as derived by BEYONDPLANCK (black), *Planck* (red), and WMAP (blue), shown at the *top*. The best-fit *Planck* 2018 Λ CDM power spectrum is shown in dashed gray. Residual power spectrum relative to Λ CDM, measured relative to full quoted error bars, $r_\ell \equiv (D_\ell - D_\ell^{\Lambda\text{CDM}})/\sigma_\ell$, shown in the *middle*. For pipelines that report asymmetric error bars, σ_ℓ is taken to be the average of the upper and lower error bar. Fractional difference with respect to the *Planck* Λ CDM spectrum, shown at the *bottom*. In this panel, each curve has been boxcar averaged with a window of $\Delta\ell = 100$ to suppress random fluctuations.

others, except as dictated by the relative instrumental noise level in each channel.

However, as discussed by [Ihle et al. \(2023\)](#), [Gjerløw et al. \(2023\)](#), and [Basyrov et al. \(2023\)](#), this apparent algorithmic simplicity does not, by any means, imply that robust CMB dipole estimation is easy in the BEYONDPLANCK procedure. Indeed, the CMB dipole is quite possibly the single most difficult parameter to estimate in the entire model, simply because it both affects, and relies on, a wide range of partially degenerate parameters. The first and foremost of these is the absolute calibration, g_0 . This parameter directly scales the amplitude of the entire CMB map, including the Solar dipole. This parameter is itself constrained from the orbital dipole, which is both weaker in terms of absolute amplitude, and for significant parts of the mission it is nearly aligned with, and thereby obscured by, the Galactic plane ([Gjerløw et al. 2023](#)).

Secondly, astrophysical foregrounds have in general both a non-zero dipole moment, as well as higher-order moments with unknown parameters, and these must be estimated jointly with the CMB dipole. Considering that the current data set includes five astrophysical components, each with a free value in each pixel, and there are only eight significantly independent frequency channels, the full system is rather poorly con-

strained. It is therefore possible to add a significant dipole to the CMB map and subtract appropriately scaled dipoles from each of the foreground maps, with only a minimal penalty in terms of the overall χ^2 . In practice, we observe particularly strong degeneracies between the CMB, AME and free-free components, when exploring the full system without priors ([Andersen et al. 2023](#)).

Thirdly, correlated noise, n_{corr} , is only weakly constrained through its $1/f$ -style PSD parameters, and this component is therefore able to account for a wide range of modeling errors, including calibration errors ([Ihle et al. 2023](#); [Watts et al. 2023](#)). In particular, incorrectly estimated gains leave a spurious dipole-like residual in the time-ordered data. Since this residual is detector-dependent, it will typically be interpreted by the algorithm as correlated noise, thereby exciting a dipolar structure in n_{corr} . Coherent large-scale patterns aligned with the Solar CMB dipole in n_{corr} is one of the most typical signs of overall calibration errors.

Finally, the coupling between the large-scale CMB quadrupole, foreground, and bandpass corrections all affect the Solar CMB dipole. While the CMB E -mode polarization quadrupole by itself is predicted by current Λ CDM models to have a very small quadrupole, with a variance of typically

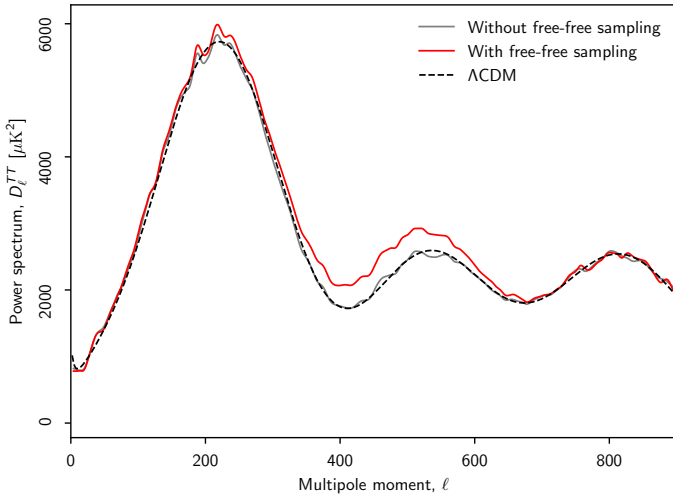


Fig. 19. Comparison of power spectra derived with (red) and without (gray) including prior-constrained free-free emission in the TT resampling procedure. Spectra have been smoothed with a Gaussian kernel to highlight the discrepancies between the two runs.

less than $0.05 \mu\text{K}^2$, there is nothing in the current parametric BEYONDPLANCK model that explicitly enforces this. This particular mode therefore opens up a particularly problematic degeneracy for *Planck* through coupling with the gain and bandpass shift as follows: an error in the absolute gain will lead to an apparently wrong orbital dipole. However, this can be countered by adding a polarized CMB quadrupole, which has the same SED and nearly the same spin harmonics as the orbital dipole, due to the *Planck* scanning strategy that observes along nearly perfect great circles (see Fig. 1 in Gjerløw et al. 2023)⁷. Second-order residuals in the total polarized sky signal as observed at each frequency can then finally be countered by adjusting the combination of relative gains, polarized foreground signals, and bandpass corrections between radiometers, leaving the total χ^2 nearly unchanged. To break this degeneracy, we actually did impose a ΛCDM power spectrum prior on the $E \ell = 2$ mode during gain estimation alone and marginalized over its amplitude; this prevents the polarization quadrupole from taking on obviously pathological values. In addition, we note that we included the large-scale WMAP polarization in the CMB fit and this also helps in regularizing the large-scale polarization signal. For comparison, we note that both the *Planck* LFI DPC and PR4 pipelines set the entire CMB polarization signal to zero during the gain estimation process; this is a far more aggressive approach to resolving this degeneracy, and for *Planck* PR4 it results in a non-negligible transfer function on large angular scales (see Planck Collaboration Int. LVII 2020 for full details). For the *Planck* LFI detectors alone, the signal-to-noise ratio is too low to make any measurable difference during gain calibration, resulting in an effectively unbiased algorithm for *Planck* LFI, even with this strong prior (Planck Collaboration III 2016).

During the initial test phase of the BEYONDPLANCK pipeline, the Markov chain was allowed to explore all the above degeneracies freely, without any informative or algorithmic priors. These early runs resulted in a full marginal uncertainty on the dipole amplitude of more than $40 \mu\text{K}$, as compared to $3 \mu\text{K}$ reported by *Planck* LFI for the 70 GHz channel alone (Planck Collaboration II 2020), or $1 \mu\text{K}$ as reported by

HFI (Planck Collaboration III 2020). Although this value by itself could be considered acceptable, given the limited cosmological importance of the CMB dipole, it was also strikingly obvious that all component maps were compromised by the poorly constrained calibration, ultimately leading to non-physical Galactic component maps with large dipolar residuals. With the introduction of the spatial free-free and AME priors discussed by Andersen et al. (2023) and the ΛCDM -based E -mode quadrupole prior discussed by Gjerløw et al. (2023), these degeneracies can effectively be broken.

Figure 20 shows the marginal CMB temperature fluctuation posterior mean map as derived in BEYONDPLANCK, given both the data, model, and priors described above. This map is massively dominated by the CMB Solar dipole, with only a small imprint of the Galactic plane being visible in the very center. At high latitudes, CMB temperature fluctuations may be seen as tiny ripples superimposed on the dipole.

Because of the small but non-negligible Galactic plane residuals, we must impose an analysis mask before estimating final CMB Solar dipole parameters. For this purpose, we used the Wiener filter estimator described by Thommesen et al. (2020), which inpaints the Galactic mask with a constrained realization prior to parameter estimation; this is necessary in order to account for and marginalize over the coupling to higher order CMB fluctuations. This method was also adopted for the dipole estimates presented in Planck Collaboration Int. LVII (2020), although we introduce one significant difference to that analysis: In the current analysis, we estimate the magnitude of systematic uncertainties directly from the BEYONDPLANCK Gibbs samples, as opposed to putting it in by hand. Specifically, instead of producing 9000 constrained realizations from a single maximum likelihood map, as done by Thommesen et al. (2020) and Planck Collaboration Int. LVII (2020), we now produce 100 constrained realizations from each of the 3200 available end-to-end Gibbs samples. Since each of these realizations have different gain, correlated noise, and foreground residuals, the full ensemble accounts seamlessly for all relevant systematic uncertainties. The only additional term we put by hand into the error budget is a contribution of $0.7 \mu\text{K}$ from the CMB monopole uncertainty (Fixsen 2009).

Using this methodology, we estimate the CMB dipole parameters over a series of Galactic masks, ranging in sky fractions from 20 to 95%. The results from these calculations are shown in Fig. 21. Overall, we see that the posterior distributions for dipole amplitude and latitude are quite stable with respect to sky fraction, while longitude estimates are stable up to $\sim 80\%$ of the sky. Furthermore, we note that the uncertainties do not decrease after $f_{\text{sky}} \approx 0.75$, as they would if the full error budget could be described in terms of white noise and sky fraction. Rather, the weight of the additional sky coverage is effectively reduced when marginalizing over the various systematic contributions, as desired. We conservatively adopt a sky fraction of $f_{\text{sky}} = 0.68$ to define our final dipole estimates, corresponding to the sky fraction close to that used for the main CMB temperature analysis. The resulting values are plotted as black points in Fig. 21 and tabulated together with previous estimates in Table 2.

Several points are worth noting regarding these results. First, we see that the reported best-fit BEYONDPLANCK dipole amplitude is $3362.7 \pm 1.4 \mu\text{K}$, which is slightly lower than the latest LFI 2018 estimate of $3364.4 \pm 3.1 \mu\text{K}$, which in turn is lower than the NPIPE estimate of $3366.6 \pm 2.6 \mu\text{K}$. On the other hand, it is very close to the latest HFI estimate of $3362.08 \pm 0.99 \mu\text{K}$, which is derived from an almost completely independent data

⁷ This particular degeneracy does not exist for WMAP, because of its more complex scanning strategy.

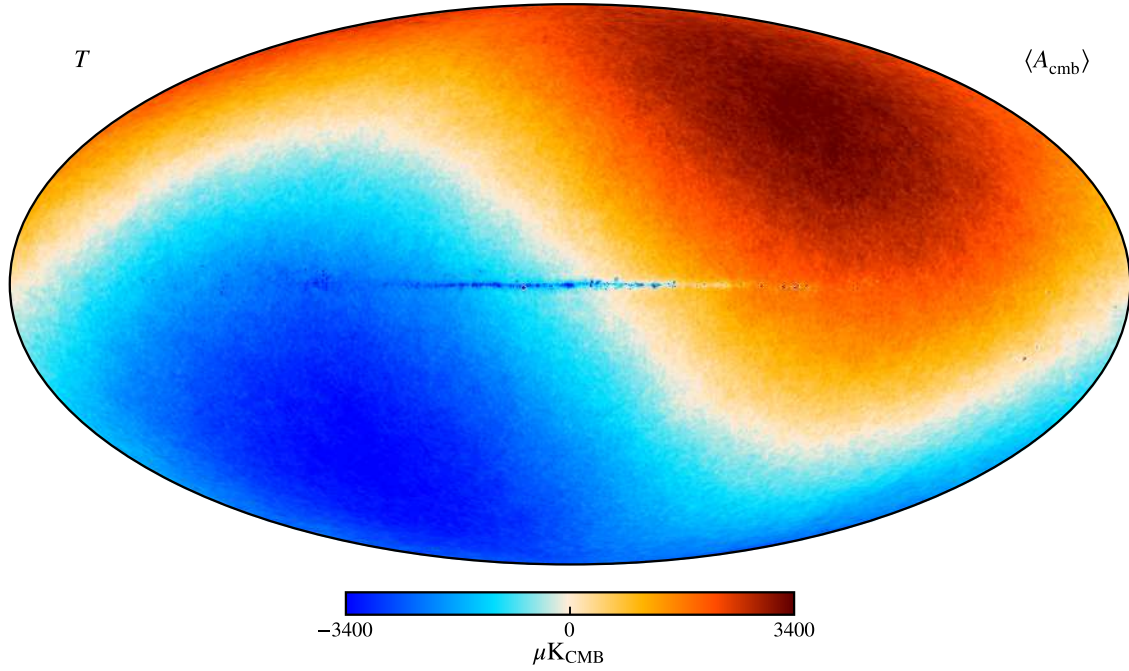


Fig. 20. Posterior mean CMB BEYONDPLANCK temperature map, smoothed to an angular resolution of 14' FWHM.

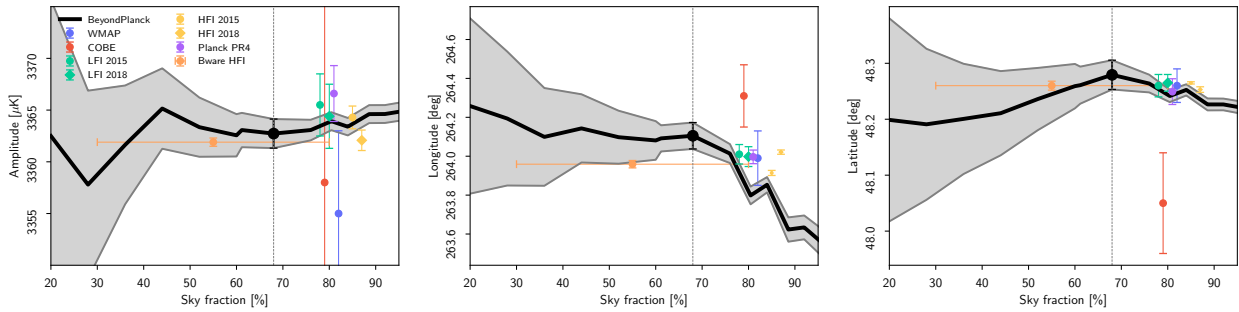


Fig. 21. CMB dipole parameters as a function of sky fraction. *From left to right*, the panels show the dipole amplitude, longitude, and latitude. Gray bands indicate 68% posterior confidence regions.

set. Overall, the agreement between these various data sets and methods is excellent.

Regarding the directional parameters, two observations are worth pointing out. First, we see that the BEYONDPLANCK uncertainties are larger than any of the previous *Planck*-dominated results. Here, it is worth recalling again that no additional systematic error contributions are added by hand to the BEYONDPLANCK directional uncertainties and the reported values are thus the direct result of degeneracies within the model itself. Perhaps the biggest algorithmic difference in this respect is the fact that the current implementation explicitly marginalizes over the full foreground and calibration model, while most other approaches condition on external constraints. The second observation is that the BEYONDPLANCK latitude is very slightly higher than any of the previous results, except COBE. The statistical significance of this difference is low, only about 1.5σ , but compared with the remarkable internal agreement between *Planck* and WMAP, it is still noteworthy. In this respect, we once again recall that we are currently using the *Planck* 2015 free-free map as an informative prior in the current processing, and CMB and free-free emission are known to be strongly correlated for the current data set; see Andersen et al. (2023). Performing a joint

analysis of LFI, HFI, and WMAP without an external free-free prior might be informative regarding this point.

6. Low- l CMB anomalies

The posterior CMB sky map samples generated by the Gibbs sampler discussed in Sect. 2 may be used for any scientific analysis to which standard foreground-reduced CMB maps are subjected. The main practical difference between these maps and traditional maximum-likelihood maps is simply that in the Bayesian case one must analyze an entire ensemble of different CMB maps, rather than just one, and the resulting answer is typically defined in terms of a histogram, rather than a single value.

The main advantage of this approach is full propagation of all modeled systematic effects, some of which are very difficult to account for with traditional approaches. One important example of this is time- and detector-dependent gain variations. As already noted, calibration uncertainties modulate the large CMB Solar dipole, and can consequently also excite other large-scale modes through a coupling from the satellite scanning strategy, noise weighting, and confidence mask. This issue is particularly pertinent to the question of large-scale CMB anomalies, several of which were reported after the release

Table 2. Comparison of Solar dipole measurements from COBE, WMAP, and *Planck*.

Experiment	Amplitude [μK_{CMB}]	Galactic coordinates		Reference
		l [deg]	b [deg]	
COBE ^{(a),(b)}	3358 ± 23	264.31 ± 0.16	48.05 ± 0.09	Lineweaver et al. (1996)
WMAP ^(c)	3355 ± 8	263.99 ± 0.14	48.26 ± 0.03	Hinshaw et al. (2009)
LFI 2015 ^(b)	3365.5 ± 3.0	264.01 ± 0.05	48.26 ± 0.02	Planck Collaboration II (2016)
HFI 2015 ^(d)	3364.29 ± 1.1	263.914 ± 0.013	48.265 ± 0.002	Planck Collaboration VIII (2016)
LFI 2018 ^(b)	3364.4 ± 3.1	263.998 ± 0.051	8.265 ± 0.015	Planck Collaboration II (2020)
HFI 2018 ^(d)	3362.08 ± 0.99	264.021 ± 0.011	48.253 ± 0.005	Planck Collaboration III (2020)
Bware	3361.90 ± 0.40	263.959 ± 0.019	48.260 ± 0.008	Delouis et al. (2021)
<i>Planck</i> PR4 ^{(a),(c)}	3366.6 ± 2.6	263.986 ± 0.035	48.247 ± 0.023	Planck Collaboration Int. LVII (2020)
BEYONDPLANCK ^(e)	3362.7 ± 1.4	264.11 ± 0.07	48.279 ± 0.026	This paper

Notes. ^(a)Statistical and systematic uncertainty estimates are added in quadrature. ^(b)Computed with a naive dipole estimator that does not account for higher-order CMB fluctuations. ^(c)Computed with a Wiener-filter estimator that estimates, and marginalizes over higher order CMB fluctuations jointly with the dipole. ^(d)Higher order fluctuations as estimated by subtracting a dipole-adjusted CMB-fluctuation map from frequency maps prior to dipole evaluation. ^(e)Estimated with a sky fraction of 68%. Error bars indicate 1σ uncertainties, and include statistical uncertainties as defined by the global BEYONDPLANCK posterior framework; they thus account for instrumental noise, gain fluctuations, parametric foreground variations etc., but not for potential modeling error uncertainties.

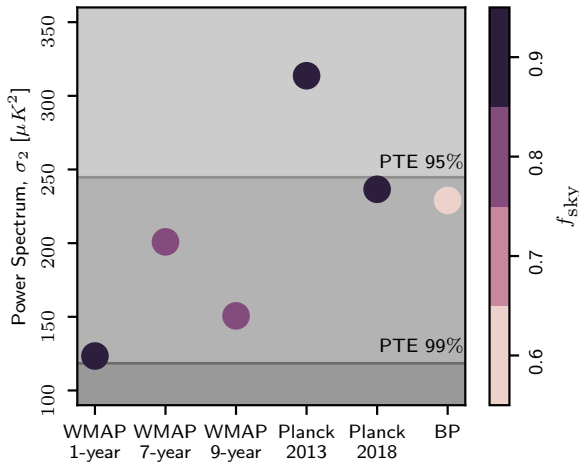


Fig. 22. Estimates of the realization-specific quadrupole amplitude of our universe, σ_2 , from WMAP, *Planck*, and BEYONDPLANCK. The gray background indicate PTEs relative to the best-fit *Planck* 2018 Λ CDM model, and the color of the dots indicate sky fraction used by the respective analysis. The radius of the dots correspond to the diagonal Fisher matrix uncertainty reported by Hinshaw et al. (2013), and provides a very naive noise-only estimate.

of the first WMAP sky maps, including a low quadrupole amplitude (Bennett et al. 1992), lack of large-scale correlations (Spergel et al. 2003), quadrupole-octopole alignment and octopole planarity (de Oliveira-Costa et al. 2004), hemispherical power asymmetry (Eriksen et al. 2004b), a large non-Gaussian cold spot (Vielva et al. 2004), and a low low- ℓ TT power spectrum (Planck Collaboration XV 2014).

Most of these effects are, however, typically only statistically significant at the 3σ level, and unmodeled systematic errors could therefore often be relevant in ways that are difficult to quantify with traditional CMB maps. As such, the new CMB posterior samples presented in this paper offer a unique opportunity to more fully assess the significance of these anomalies. In the following, we consider four examples that are implementationally straightforward to evaluate, namely: (1) the low quadrupole amplitude, (2) quadrupole-octopole alignment, (3) octopole planarity, and (4)

low low- ℓ TT spectrum. We encourage other research groups to revisit the remaining anomalies using their own tools on the new posterior products provided here.

6.1. Low quadrupole amplitude

As already mentioned, the TT quadrupole amplitude has been measured to be relatively weak compared to the Λ CDM predictions ever since COBE-DMR (Bennett et al. 1992) and this observation has been confirmed both by WMAP (Hinshaw et al. 2003) and *Planck* (Planck Collaboration XV 2014). However, it is interesting to note that the various experiments and analyses report quite different values when it comes to the precise value for the quadrupole, as illustrated in Fig. 22. Here we show the reported quadrupole amplitudes⁸, σ_2 , for WMAP, *Planck*, and BEYONDPLANCK; the gray background indicates the PTE relative to the best-fit *Planck* 2018 Λ CDM model, while the color of the dots indicate sky fraction. For reference, Hinshaw et al. (2013) reports a diagonal Fisher uncertainty for this mode of $9 \mu\text{K}^2$, which is comparable to the dot radius.

All analyses report a generally low amplitude compared to Λ CDM, with all PTEs except one being higher than 0.95. At the same time, it is also striking to note that even very similar analyses that rely on highly correlated datasets tend to use almost identical techniques and are performed by the same research group, find results that vary internally by many sigmas: The seven-year WMAP analysis reports a best-fit value of $201 \mu\text{K}^2$ (Larson et al. 2011), while the corresponding nine-year analysis reports $151 \mu\text{K}^2$ (Hinshaw et al. 2013), which are formally different by more than 5σ . Furthermore, the confidence sky mask used in these two analyses are identical, and sample variance does therefore not contribute at all to this difference. Likewise, *Planck* 2013 and 2018 reports values of 299 and $226 \mu\text{K}^2$, respectively, discrepant at more than 8σ , as measured by naive Fisher uncertainties, taking into account *Planck*'s higher signal-to-noise ratio.

⁸ Recall that σ_2 denotes the realization-specific quadrupole amplitude of our Universe, while C_2 (or $D_2 = C_2 \cdot 6/2\pi$) denotes the ensemble-averaged quadrupole amplitude.

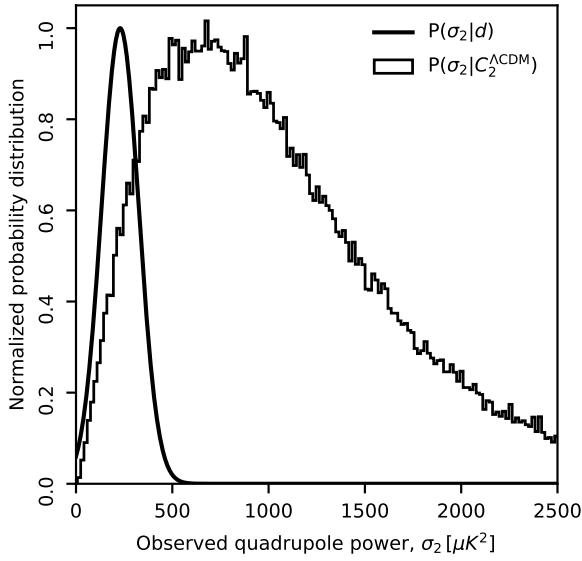


Fig. 23. Comparison of the realization-specific quadrupole amplitude distribution derived from BEYONDPLANCK (smooth black curve) and the predicted distribution from the *Planck* 2018 best-fit Λ CDM power spectrum (histogram).

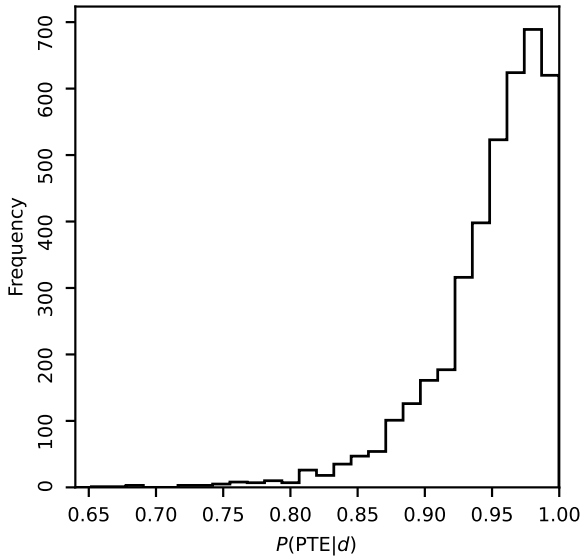


Fig. 24. PTE probability distribution, $P(\text{PTE} | \mathbf{d})$, for the realization-specific quadrupole amplitude, σ_2 to exceed the Λ CDM prediction after marginalizing over all modeled uncertainties.

What these results clearly show is simply that white noise uncertainties only account for a small fraction of the total CMB temperature quadrupole uncertainty. With the BEYONDPLANCK posterior samples, we are finally in a position where this statement may be quantified more precisely in terms of the full marginal posterior distribution $P(\sigma_2 | \mathbf{d})$, which now accounts for important contributions from calibration and foreground uncertainties, and their correlations. This distribution is shown in Fig. 23 as a solid smooth curve, while the histogram is derived from 10^5 ideal realizations of σ_2 drawn from the *Planck* 2018 best-fit Λ CDM prediction, $C_2^{\Lambda\text{CDM}} = 1064.7 \mu\text{K}^2$. The posterior distribution may be summarized as a Gaussian with $\sigma_2 = 229 \pm 97$, and the full marginal uncertainty, including contributions from the instrument, astrophysics and confidence mask, is thus more than ten times larger than the naive diagonal

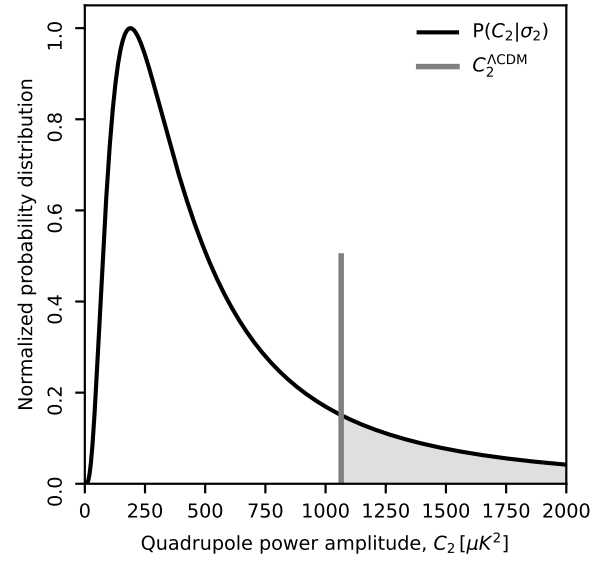


Fig. 25. Marginal probability distribution of the ensemble-averaged quadrupole power spectrum, $P(C_\ell | \mathbf{d})$, estimated in BEYONDPLANCK (solid line). The vertical line at $C_2 = 1064.7$ indicates the value predicted by the *Planck* 2018 best-fit Λ CDM model; 21.7% of the marginal distribution exceeds this value.

Fisher estimate quoted above, despite the fact that more data are included in the current analysis.

We are now interested in deriving a total significance for the low quadrupole amplitude. In a classic frequentist simulation-based analysis this would be done simply by counting how many of the simulated realizations in the histogram Fig. 23 have a lower value than the observed peak posterior value. However, in our case the PTE of the peak position carries no particular statistical significance and, instead, the full distribution of possible σ_2 values must be considered. Accordingly, Fig. 24 shows the probability distribution of PTEs, $P(\text{PTE} | \mathbf{d})$ and here we see that the 95% confidence limit on this PTE is 0.85. Thus, the observed quadrupole value is certainly on the low side compared to the Λ CDM prediction, but the effect is not highly significant with the current dataset. A smaller confidence mask and a better constrained instrument model are required to shed further light on the effect.

To conclude this section, we also turn the question around to consider what the probability distribution for C_2 would be given the measured values of σ_2 . To answer this, we evaluated the GBR estimator discussed in Sect. 4.3 as a function of D_ℓ , as shown in Fig. 25. The PTE for C_2 relative to $C_2^{\Lambda\text{CDM}} = 1064.7 \mu\text{K}^2$ is 21.7%.

6.2. Quadrupole-octopole alignment

A second anomaly regarding the very largest angular scales in the CMB map was first reported by Tegmark et al. (2003) and de Oliveira-Costa et al. (2004), who found that the quadrupole and octopoles appeared morphologically aligned. This was quantified by first defining a preferred direction, $\hat{\mathbf{n}}_\ell$, for each mode separately by maximizing the angular momentum dispersion of the wave function,

$$\langle \psi | (\hat{\mathbf{n}}_\ell \cdot \mathbf{L})^2 | \psi \rangle = \sum_m m^2 |a_{\ell m}(\hat{\mathbf{n}})|^2, \quad (37)$$

and then computing the angular separation between $\hat{\mathbf{n}}_2$ for the quadrupole and $\hat{\mathbf{n}}_3$. de Oliveira-Costa et al. (2004) found that

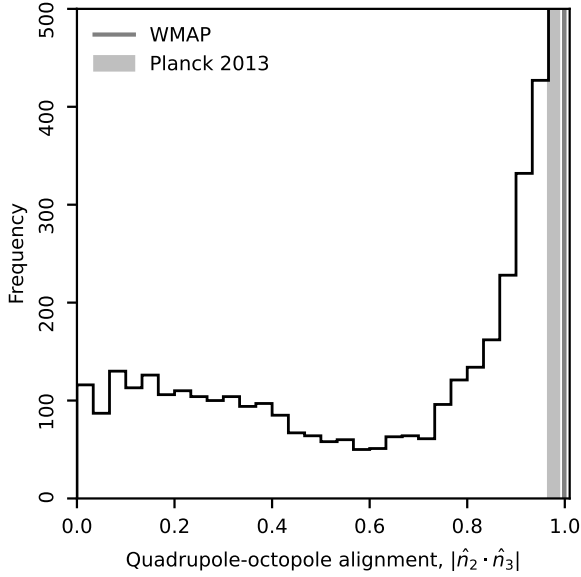


Fig. 26. Histogram of the quadrupole-octopole alignment, $|\hat{n}_2 \cdot \hat{n}_3|$. The official WMAP nine-year result reported the value $|\hat{n}_2 \cdot \hat{n}_3| = 0.9986$ (Bennett et al. 2013) from a misalignment of 3° , while *Planck* (Planck Collaboration XXIII 2014) reported the interval $[0.9663, 0.9877]$ corresponding to 9 and 13° , respectively, depending on the component separation method.

this angle was smaller than the isotropic expectation with a PTE of 0.984 for the first-year WMAP data. This observation was qualitatively later confirmed by the WMAP team (Bennett et al. 2013), who found the angle to be around 3° in the nine year data, corresponding to a probability of 0.14% for such an alignment or stronger to occur assuming isotropy. The WMAP team did however note that the foreground removal procedure was a limiting factor in the measurement of the misalignment. Likewise, Planck Collaboration XXIII (2014) reported an alignment in the interval 9° and 13° depending on the component separation method, corresponding to PTEs in between 1.2 and 3.4% , respectively.

Given the fact that instrumental systematic uncertainties affect the absolute quadrupole amplitude by a large factor, as discussed in the previous section, it is reasonable to assume that also the preferred quadrupole direction is affected by the same uncertainties. In this section, we therefore apply the methodology of de Oliveira-Costa et al. (2004) to the same set of constrained CMB realizations discussed above, and derive the full distribution of alignment PTEs after full systematics marginalization. This is summarized in the form of a histogram in Fig. 26, with the nine-year WMAP and *Planck* 2013 results shown as gray vertical bars. The width of the *Planck* bar indicates the uncertainty derived among the four *Planck* component separation codes.

In this figure, we see that the agreement between the BEYONDPLANCK results and previous results is excellent in terms of single-point maximum posterior values. However, we also see that the full BEYONDPLANCK posterior distribution is very broad, to the extent that all possible angles are in fact allowed by the data, from 0 to 90° . Part of this larger uncertainty does come from the somewhat more conservative analysis mask with $f_{\text{sky}} = 0.64$ employed in the current analysis, as compared to 0.72 for Planck Collaboration XXIII (2014). At the same time, we also note that foreground modeling details appear to have only a small impact of the final results, as very different

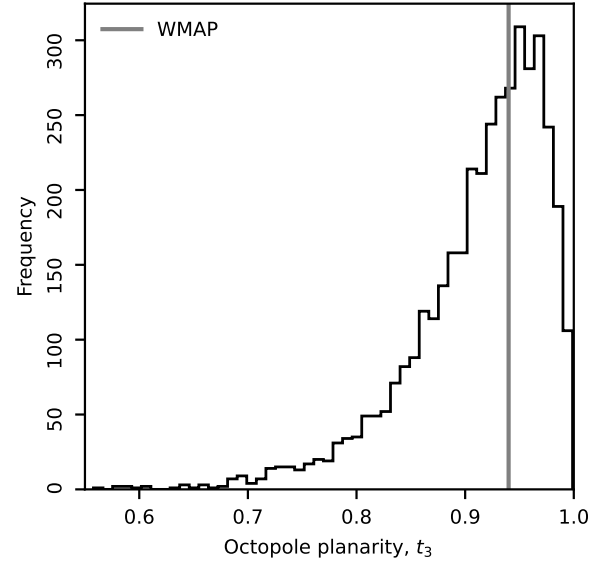


Fig. 27. Histogram of the planarity test statistics, t_3 , derived from the set of BEYONDPLANCK constrained CMB realizations. We also mark the value of $t_3 = 94\%$ originally reported by de Oliveira-Costa et al. (2004) from the first-year WMAP ILC map.

methods reach quite similar conclusions: the WMAP result was derived from an Internal Linear Combination (ILC) map with low-resolution foreground eigenmode error propagation, while the *Planck* results were derived using four qualitatively different methods coupled with end-to-end simulations. All these methods agree internally qualitatively very well and they also agree with the maximum-posterior BEYONDPLANCK result.

The fundamental difference between the BEYONDPLANCK and previous analyses does not lie in different foreground modeling, but rather in the instrument modelling and the general statistical treatment and error propagation. Most importantly, while previous analyses only accounted for relatively simple foreground and noise uncertainties, the BEYONDPLANCK processing additionally accounts for full gain uncertainties and their coupling to the CMB Solar dipole and foregrounds. When doing so, the statistical evidence for a quadrupole-octopole alignment diminishes significantly. Of course, it is also important to emphasize that a substantial contributor to this additional variance is the exclusion of the *Planck* HFI measurements, which would allow both for better CMB constraints (and thereby indirectly also stronger LFI calibration constraints), as well as a smaller Galactic plane by properly fitting free-free and thermal dust emission. Future works that also include HFI data will need to revisit this question.

6.3. Planar octopole

Tegmark et al. (2003) and de Oliveira-Costa et al. (2004) also noted that not only is the plane of the temperature octopole closely aligned with the quadrupole, but the octopole is also intrinsically highly planar. de Oliveira-Costa et al. (2004) quantified this through the test statistic, t_3 ,

$$t_3 \equiv \max_{\hat{n}} \frac{|a_{3-3}(\hat{n})|^2 + |a_{33}(\hat{n})|^2}{\sum_{m=-3}^3 |a_{3m}(\hat{n})|^2}, \quad (38)$$

which measures the ratio of the total octopole power that may be contributed to $a_{3\pm 3}$, maximized over all coordinate systems. This is shown in terms of a histogram for the BEYONDPLANCK CMB

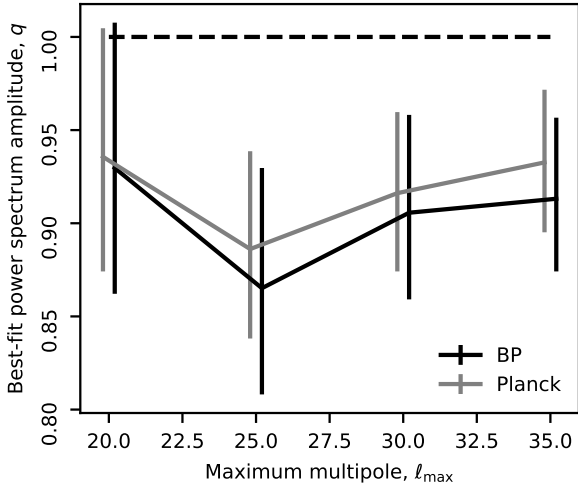


Fig. 28. Best-fit amplitude, q , as a function of maximum multipole, l_{\max} measured relative to the best-fit *Planck* Λ CDM power spectrum, as derived for both *Planck* 2015 (gray) and BEYONDPLANCK (black).

samples in Fig. 27 together with the original WMAP measurement of $t_3 = 94\%$ by de Oliveira-Costa et al. (2004). Once again, we see that the agreement is very good in terms of the single-point peak value – but we also see that the distribution is quite broad when marginalizing over the full set of uncertainties. This distribution is in qualitatively good agreement with the results of Rassat et al. (2014), who measured the octopole planarity for six different foreground-reduced *Planck* 2013 CMB maps and found values ranging between 0.84 and 0.95 (corresponding to PTEs between 7 and 37%) depending on foreground cleaning and mask details. When additionally marginalizing over instrumental systematic effects in BEYONDPLANCK, we see that the range broadens further.

6.4. Low- ℓ temperature amplitude

Finally, we revisit the “low low- ℓ temperature power spectrum anomaly”, first reported by Planck Collaboration XV (2014)⁹. In this case, the *Planck* team fitted the low- ℓ temperature power spectrum with an amplitude-scaled model $C_\ell(q) = qC_\ell^{\Lambda\text{CDM}}$ between $2 \leq \ell \leq l_{\max}$, and varied l_{\max} . In doing so, they found best-fit amplitudes typically ranging between 0.87 and 0.95, which are low with statistical significance typically at the 1.5–2.5 σ level.

Figure 28 shows corresponding results for both BEYONDPLANCK and *Planck* 2015. Once again, we see that the results generally agree well. In this case, we see on the one hand that the BEYONDPLANCK mean results are in fact slightly lower than the *Planck* results, possibly hinting towards a stronger anomaly. On the other hand, the uncertainties are also larger due to the larger confidence mask and more complete instrument error marginalization and the overall significance of the effect is therefore essentially unchanged.

7. Discussion and conclusions

In this paper, we present the CMB results from the Bayesian BEYONDPLANCK analysis (BeyondPlanck Collaboration 2023).

⁹ An analogous effect on the angular correlation function had been previously observed in WMAP data, albeit at lower significance level, e.g. Bennett et al. (2011).

This represents the first example of an end-to-end posterior sample-based CMB analysis for which the inputs are defined in terms of raw time-ordered data and the final outputs are CMB sky maps and power spectra. This method was first suggested in a CMB setting by Jewell et al. (2004) and Wandelt et al. (2004), taking almost twenty years of computer hardware and algorithm development to realize this in practice.

Two of the most fundamental advantages of integrated end-to-end CMB analysis are full joint exploration of all free parameters – instrumental, astrophysical, and cosmological – and true end-to-end error propagation. In principle, this algorithm has a similar statistical foundation as the traditional low- ℓ brute-force CMB likelihood approach used by both WMAP (Hinshaw et al. 2013) and *Planck* (Planck Collaboration V 2020), but with a few key differences: rather than just accounting for correlated noise and template-based foreground residuals at low angular resolution, this method can account for all degrees of uncertainty at full angular resolution. It achieves this through Monte Carlo sampling, as opposed to analytic construction of dense covariance matrices and neither angular resolution nor model complexity therefore carry a similar prohibitive computational cost as the traditional method.

It is important to note the Bayesian method, whether implemented analytically or through sampling, is fundamentally different from the frequentist forward simulation-based method that is commonly used in the CMB field for error propagation. Intuitively, the main difference lies in that, while forward simulations describe some random instrument and universe, the Bayesian approach describes our instrument and Universe. Because of this difference, the two methods are naturally geared toward answering different types of statistical questions. For the Bayesian approach, it is easy to address such questions as what the most likely Λ CDM parameters for our Universe would be. However, it is technically difficult (due to the large dimensionality of the parameter space) to answer the question of whether our data set is consistent with the Λ CDM model. For the frequentist simulation approach, the opposite holds true. It is also interesting to note that the simulation-based approach becomes indistinguishable from the Bayesian approach if constrained realizations are used to generate the instrument and sky model, as opposed to statistically independent realizations, as is typically done. Indeed, the current BEYONDPLANCK implementation may in many respects simply be considered as a constrained realization-based simulator.

In this paper, we use this sampling framework to address several classical problems in CMB analysis. We study cross-correlations between the CMB component and instrumental and astrophysical parameters, and we identify and mitigate a particularly strong degeneracy with free-free emission. We compare the resulting posterior mean CMB maps and power spectra with previously published results and found a good agreement. We have also derived a CMB Solar dipole amplitude of $3362.7 \pm 1.4 \mu\text{K}$, which is in excellent agreement with previous results. However, it is important to note that the quoted uncertainty is derived directly from the global statistical model and not associated with any additional *Planck*-specific systematic error.

Given that all of the above are in good agreement with previous results, we might wonder what the point of this approach is, as we have shown that the traditional method works just as well. The main answer to this question may be formulated in terms of signal-to-noise ratio: as long as the statistic or quantity in question is signal-dominated, such as the *Planck* TT spectrum on large and intermediate scales, the current method provides little or no obvious advantage. However, when the statistic

in question is either systematics- or noise-dominated, then these methods become very powerful through their end-to-end error propagation capabilities. This was explicitly demonstrated in this paper by revisiting a number of previously reported large-scale anomalies in the CMB temperature anisotropies. In many cases, we found that the significance of these anomalies was significantly reduced after accounting for both low-level instrumental parameters and the full non-Gaussian shape of the posterior distribution. Such effects are very difficult to model accurately by non-sampling methods.

We posit that the same will also hold true for any next-generation high-sensitivity CMB B -mode experiment that aims to detect primordial gravitational waves. These experiments are looking for a signal that is five or more orders of magnitude weaker than the CMB dipole and at least a few orders of magnitude weaker than the Galactic diffuse foregrounds. As such, accurate and joint error propagation of both instrumental and astrophysical uncertainties will be key to claiming a robust detection. Indeed, developing methods applicable to this task was the main motivation behind the BEYONDPLANCK project in general. The current work has shown in practice that it is now possible to implement a computationally viable end-to-end Bayesian CMB analysis pipeline.

Acknowledgements. We thank Prof. Pedro Ferreira and Dr. Charles Lawrence for useful suggestions, comments and discussions. We also thank the entire *Planck* and WMAP teams for invaluable support and discussions, and for their dedicated efforts through several decades without which this work would not be possible. The current work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement numbers 776282 (COMPET-4; BEYONDPLANCK), 772253 (ERC; BITS2COSMOLOGY), and 819478 (ERC; COSMOGLOBE). In addition, the collaboration acknowledges support from ESA; ASI and INAF (Italy); NASA and DoE (USA); Tekes, Academy of Finland (grant no. 295113), CSC, and Magnus Ehrnrooth foundation (Finland); RCN (Norway; grant nos. 263011, 274990); and PRACE (EU).

References

- Andersen, K. J., Herman, D., Aurlien, R., et al. 2023, *A&A*, 675, A13 (BeyondPlanck SI)
- Basak, S., & Delabrouille, J. 2012, *MNRAS*, 419, 1163
- Basak, S., & Delabrouille, J. 2013, *MNRAS*, 435, 18
- Basyrov, A., Suur-Uski, A.-S., Colombo, L. P. L., et al. 2023, *A&A*, 675, A10 (BeyondPlanck SI)
- Bennett, C. L., Smoot, G. F., Hinshaw, G., et al. 1992, *ApJ*, 396, L7
- Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2003, *ApJS*, 148, 97
- Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2011, *ApJS*, 192, 17
- Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, *ApJS*, 208, 20
- BeyondPlanck Collaboration (Andersen, K. J., et al.) 2023, *A&A*, 675, A1 (BeyondPlanck SI)
- Cardoso, J., Martin, M., Delabrouille, J., Betoule, M., & Patanchon, G. 2008, *IEEE J. Sel. Top. Signal Process.*, 2, 735
- Chu, M., Eriksen, H. K., Knox, L., et al. 2005, *Phys. Rev. D*, 71, 103002
- Delouis, J. M., Pagano, L., Mottet, S., Puget, J. L., & Vibert, L. 2019, *A&A*, 629, A38
- Delouis, J. M., Puget, J. L., & Vibert, L. 2021, *A&A*, 650, A82
- de Oliveira-Costa, A., Tegmark, M., Zaldarriaga, M., & Hamilton, A. 2004, *Phys. Rev. D*, 69, 063516
- Eriksen, H. K., O'Dwyer, I. J., Jewell, J. B., et al. 2004a, *ApJS*, 155, 227
- Eriksen, H. K., Hansen, F. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2004b, *ApJ*, 605, 14
- Eriksen, H. K., Jewell, J. B., Dickinson, C., et al. 2008, *ApJ*, 676, 10
- Fernández-Cobos, R., Vielva, P., Barreiro, R. B., & Martínez-González, E. 2012, *MNRAS*, 420, 2162
- Fixsen, D. J. 2009, *ApJ*, 707, 916
- Galloway, M., Reinecke, M., Andersen, K. J., et al. 2023a, *A&A*, 675, A8 (BeyondPlanck SI)
- Galloway, M., Andersen, K. J., Aurlien, R., et al. 2023b, *A&A*, 675, A3 (BeyondPlanck SI)
- Gelman, A., & Rubin, D. B. 1992, *Stat. Sci.*, 7, 457
- Geman, S., & Geman, D. 1984, *IEEE Trans. Pattern Anal. Mach. Intell.*, 6, 721
- Gerakakis, S., Brilenkov, M., Ieronymaki, M., et al. 2023, *Open J. Astrophys.*, 6, 10
- Gerbino, M., Lattanzi, M., Migliaccio, M., et al. 2020, *Front. Phys.*, 8, 15
- Gjerløw, E., Colombo, L. P. L., Eriksen, H. K., et al. 2015, *ApJS*, 221, 5
- Gjerløw, E., Ihle, H. T., Galeotta, S., et al. 2023, *A&A*, 675, A7 (BeyondPlanck SI)
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759
- Herman, D., Watson, R. A., Andersen, K. J., et al. 2023a, *A&A*, 675, A5 (BeyondPlanck SI)
- Herman, D., Hensley, B., Andersen, K. J., et al. 2023b, *A&A*, 675, A15 (BeyondPlanck SI)
- Hinshaw, G., Spergel, D. N., Verde, L., et al. 2003, *ApJS*, 148, 135
- Hinshaw, G., Weiland, J. L., Hill, R. S., et al. 2009, *ApJS*, 180, 225
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJS*, 208, 19
- Ihle, H. T., Bersanelli, M., Franceschet, C., et al. 2023, *A&A*, 675, A6 (BeyondPlanck SI)
- Ijjas, A., Steinhardt, P. J., & Loeb, A. 2014, *Phys. Lett. B*, 736, 142
- Jewell, J., Levin, S., & Anderson, C. H. 2004, *ApJ*, 609, 1
- Jewell, J. B., Eriksen, H. K., Wandelt, B. D., et al. 2009, *ApJ*, 697, 258
- Kamionkowski, M., & Kovetz, E. D. 2016, *ARA&A*, 54, 227
- Larson, D. L., Eriksen, H. K., Wandelt, B. D., et al. 2007, *ApJ*, 656, 653
- Larson, D., Dunkley, J., Hinshaw, G., et al. 2011, *ApJS*, 192, 16
- Leach, S. M., Cardoso, J., Baccigalupi, C., et al. 2008, *A&A*, 491, 597
- Lewis, A., & Bridle, S. 2002, *Phys. Rev. D*, 66, 103511
- Lineweaver, C. H., Tenorio, L., Smoot, G. F., et al. 1996, *ApJ*, 470, 38
- Page, L., Hinshaw, G., Komatsu, E., et al. 2007, *ApJS*, 170, 335
- Paradiso, S., Colombo, L. P. L., Andersen, K. J., et al. 2023, *A&A*, 675, A12 (BeyondPlanck SI)
- Penrose, R. 1989, *Ann. N. Y. Acad. Sci.*, 571, 249
- Planck Collaboration I. 2014, *A&A*, 571, A1
- Planck Collaboration XII. 2014, *A&A*, 571, A12
- Planck Collaboration XV. 2014, *A&A*, 571, A15
- Planck Collaboration XXIII. 2014, *A&A*, 571, A23
- Planck Collaboration II. 2016, *A&A*, 594, A2
- Planck Collaboration III. 2016, *A&A*, 594, A3
- Planck Collaboration VII. 2016, *A&A*, 594, A7
- Planck Collaboration VIII. 2016, *A&A*, 594, A8
- Planck Collaboration X. 2016, *A&A*, 594, A10
- Planck Collaboration XI. 2016, *A&A*, 594, A11
- Planck Collaboration XII. 2016, *A&A*, 594, A12
- Planck Collaboration I. 2020, *A&A*, 641, A1
- Planck Collaboration II. 2020, *A&A*, 641, A2
- Planck Collaboration III. 2020, *A&A*, 641, A3
- Planck Collaboration IV. 2020, *A&A*, 641, A4
- Planck Collaboration V. 2020, *A&A*, 641, A5
- Planck Collaboration VI. 2020, *A&A*, 641, A6
- Planck Collaboration Int. XLVI. 2016, *A&A*, 596, A107
- Planck Collaboration Int. LVII. 2020, *A&A*, 643, A42
- Racine, B., Jewell, J. B., Eriksen, H. K., & Wehus, I. K. 2016, *ApJ*, 820, 31
- Rassat, A., Starck, J. L., Paykari, P., Sureau, F., & Bobin, J. 2014, *JCAP*, 2014, 006
- Rudjord, Ø., Groeneboom, N. E., Eriksen, H. K., et al. 2009, *ApJ*, 692, 1669
- Seljebotn, D. S., Bærland, T., Eriksen, H. K., Mardal, K. A., & Wehus, I. K. 2019, *A&A*, 627, A98
- Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, *ApJS*, 148, 175
- Svalheim, T. L., Zonca, A., Andersen, K. J., et al. 2023a, *A&A*, 675, A9 (BeyondPlanck SI)
- Svalheim, T. L., Andersen, K. J., Aurlien, R., et al. 2023b, *A&A*, 675, A14 (BeyondPlanck SI)
- Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, *ApJ*, 480, 22
- Tegmark, M., de Oliveira-Costa, A., & Hamilton, A. 2003, *Phys. Rev. D*, 68, 123523
- Thommesen, H., Andersen, K. J., Aurlien, R., et al. 2020, *A&A*, 643, A179
- Tristram, M., Banday, A. J., Górski, K. M., et al. 2022, *Phys. Rev. D*, 105, 083524
- Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., & Cayón, L. 2004, *ApJ*, 609, 22
- Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A. 2004, *Phys. Rev. D*, 70, 083511
- Watts, D. J., Galloway, M., Ihle, H. T., et al. 2023, *A&A*, 675, A16 (BeyondPlanck SI)
- Zaldarriaga, M., & Seljak, U. 1997, *Phys. Rev. D*, 55, 1830