

COSMOGLOBE. Optimal joint end-to-end analysis of *WMAP* and *Planck* LFI

D. J. Watts^{1*}, A. Basyrov¹, H. T. Ihle¹, and F. Rahman² et al.

¹ Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo, Norway

² Indian Institute of Astrophysics, Koramangala II Block, Bangalore, 560034, India

January 4, 2023

ABSTRACT

We present the first joint analysis of *WMAP* and *Planck* LFI data, presenting maps that have been generated from a fully consistent joint treatment, including the sampling of sky signals and instrumental properties. The joint analysis approach yields improved *WMAP* data with better treatment of poorly constrained modes, as well as the first fully optimal sampling of all nine years of data. We also improve on the BEYONDPLANCK analysis, by reducing poorly measured modes in LFI polarization. In particular, we find a $\sim 4\mu\text{K}$ change in the 30 GHz channel as a result of including the higher signal-to-noise *WMAP* *K*-band maps. The *WMAP* maps we present are free of previously documented systematic effects, and have an $x\%$ reduction in the white noise level. As the first release of COSMOGLOBE products, the maps from this analysis should be considered both a considerable improvement over previous analyses, as well as the first iteration of future joint analyses with other data, including, e.g., the ground-based QUIET experiment and the DIRBE instrument aboard *COBE*.

Key words. ISM: general – Cosmology: observations, polarization, cosmic microwave background, diffuse radiation – Galaxy: general

Contents

1	Introduction
2	COSMOGLOBE map-making and end-to-end error propagation
2.1	Statistical framework
2.2	<i>Planck</i> LFI data processing
2.3	<i>WMAP</i> data processing
2.4	Calibration
3	Characteristics of frequency maps
4	Systematic error and uncertainties
5	Component separation
6	Conclusions

1. Introduction

A to-do list:

- Find the time it takes for each beam to cross itself.
- Fix AME model (*I'm not sure what motivated this, perhaps not necessary?*)
- Fix noise model (*Explained because of the Bessel filter plus linear trend*)

A table to include

- Spin rate – 0.464 rpm (7.57 mHz), but translations to 2.6 degrees per second in boresight?

* Corresponding author: D. J. Watts; duncanwa@astro.uio.no

- 1**
– Precession – 1 rev/hour (0.3 mHz)
– Signal bandwidth extends from 0.008–8 Hz (?)
– Beam size in degrees – 0.88, 0.66, 0.51, 0.35, 0.22.

2
The cosmic microwave background (CMB) is the most direct probe of the initial state of the Universe. Since the initial discovery of the CMB (?), subsequent experiments have continually refined the measurements, to the extent that the *WMAP* results are generally considered bringing cosmology into the regime of precision science (?). Prior to *WMAP*, it was common for CMB experiments to be superseded by more sensitive successors, with the noteworthy exceptions of *COBE/FIRAS* and *COBE/DIRBE*.
3

3
The *Planck* experiment, rather than superseding *WMAP*, consistently used *WMAP* data in its calibration, component separation, and cosmological analyses. The most direct comparison between *WMAP* and *Planck* is through analysis of the two experiments' frequency maps, as *WMAP*'s *K*, *Ka*, *Q*, *V*, and *W* maps are interleaved by the *Planck* LFI's 30, 44, and 70 GHz bands. Since the initial *Planck* data release, there have been several analyses comparing the two experiments by members of the *WMAP* team (?????) and by the *Planck* team (???).

6
While the *WMAP* low-level analysis has remained stable since ?, there has been continued work on *Planck* time-ordered data processing, notably BEYONDPLANCK for the LFI instrument (?), SRoll2 for the HFI instrument (?), and *Planck* DR4 for both LFI and HFI (NPIPE, ?). The LFI instrument in particular has had several systematics mitigated by improved analysis, particularly a smoothed gain solution and an improved noise model (????). When comparing *WMAP* *K*-band with the *Planck* LFI data, the residuals are mainly characterized by *WMAP*'s poorly measured modes, which can be seen clearly in Figures 50 and 51 of ? and Figures 4 and 7 of ?.

One of the primary outcomes of the BEYONDPLANCK project is that end-to-end analysis of a dataset with poorly mea-

sured modes can be mitigated by a joint analysis with another dataset that measures these modes well. In particular, *Planck* LFI had large scale polarized modes aligned with the instrument's scan strategy, induced by relative errors between different polarization-sensitive radiometers (?). The BEYOND-PLANCK project mitigated this by using *WMAP*'s polarized *Ka*–*V* maps for component separation, where these modes were well-measured. In order to properly combine these datasets, the polarized maps were the $N_{\text{side}} = 16$ HEALPix¹ products with a pixel-pixel covariance matrix that explicitly projected out the poorly measured modes.

In principle, the *Planck* experiment can be used to identify *WMAP*'s poorly measured modes in the same way that *WMAP* removed *Planck*'s poorly measured modes. This was shown in ?, in which *WMAP* data was calibrated against the BEYOND-PLANCK sky model, and the resulting maps differed from the *WMAP9* products mainly through the lack of the poorly measured modes. This work mainly functioned as a demonstration that the Commander3 framework could be applied to the *WMAP* dataset, and was not a true end-to-end analysis.

In this work, we present the first joint TOD analysis in the COSMOGLOBE² framework, in which we analyze the full *WMAP* dataset along with time-ordered *Planck* LFI data. In Sect. 2, we review the COSMOGLOBE statistical framework and the data processing for *Planck* LFI and *WMAP* in the Commander3 pipeline. In Sect. 3, we present the *Planck* and *WMAP* joint frequency maps, and compare these frequency maps with the fiducial analyses in Sect. ?. We discuss outstanding systematic errors and the propagation of uncertainty in Sect. 4. We summarize our results and lay a path forward in Sect. 6.

2. COSMOGLOBE map-making and end-to-end error propagation

Describe COSMOGLOBE.

2.1. Statistical framework

In the COSMOGLOBE paradigm, it is simplest to characterize the data's goodness-of-fit to a model,

$$\mathbf{d} = \mathbf{GP}[\mathbf{B}^{\text{symm}} \mathbf{Ma} + \mathbf{B}^{4\pi}(\mathbf{s}^{\text{orb}} + \mathbf{s}^{\text{fsl}})] + \mathbf{s}^{\text{inst}} + \mathbf{n}^{\text{corr}} + \mathbf{n}^{\text{w}}, \quad (1)$$

where \mathbf{G} is the time-dependent gain in the form of the matrix $\text{diag}(g_t)$, \mathbf{P} is the $n_p \times n_t$ pointing matrix, \mathbf{B}^{symm} and $\mathbf{B}^{4\pi}$ are the symmetrized and full symmetric beam, respectively, \mathbf{M} is the mixing matrix between a given component c with spectral energy distribution $f_c(\nu/\nu_{0,c})$ and a detector j with bandpass $\tau_j(\nu)$, given by

$$\mathbf{M}_{cj} = \int d\nu \tau_j(\nu) f_c(\nu/\nu_{c,0}). \quad (2)$$

The maps \mathbf{a} are the component amplitudes, \mathbf{s}^{orb} is the orbital dipole induced by the motion of the telescope with respect to the Sun, and \mathbf{s}^{fsl} is the time-dependent far sidelobe signal. In the Commander3 (?) implementation, \mathbf{n}^{corr} is a realization of the correlated noise component whose SED is parametrized by $P(f | \xi_n)$, where ξ_n generally includes f_{knee} , a slope α , and whose amplitude is fixed by the white noise σ_0 . This model is often augmented, as we will discuss in Sects. 2.2 and 2.3. Similarly, each

experiment has particular signals that are specific to the instrument in question, e.g., the 1 Hz spike in *Planck* LFI or the horn imbalance in *WMAP*, which can be modeled by \mathbf{s}^{inst} .

As shown in ?, this allows us to write down a total model for the data, $\mathbf{d} = \mathbf{s}^{\text{tot}}(\omega) + \mathbf{n}^{\text{w}}$, where \mathbf{s}^{tot} encompasses all of the terms in Eq. (1) except for the white noise term. Assuming that all instrumental effects have been modelled, the data should be Gaussian distributed with a mean of $\mathbf{s}^{\text{tot}}(\omega)$ and variance σ_0^2 . Given this model, we can evaluate the likelihood for arbitrary chunks of time-ordered data in the context of the entire model, so that individual chunks of data with poor fits can be more easily identified. In general, the likelihood is written

$$P(\mathbf{d} | \omega) \propto \exp\left(-\frac{1}{2} \sum_t \frac{(d_t - s_t^{\text{tot}}(\omega))^2}{\sigma_0^2}\right). \quad (3)$$

If $\mathbf{d} \sim \mathcal{N}(\mathbf{s}^{\text{tot}}, \sigma_0^2)$ is the correct model for the data, the argument of the exponent is proportional to a χ^2 -distribution with n_{TOD} degrees of freedom. In the limit of large n , a χ^2 distribution is well-approximated by a Gaussian with mean n and variance $2n$. Therefore we define and use the reduced- χ^2 statistic,

$$\chi^2 \equiv \frac{\sum_t ((d_t - s_t^{\text{tot}})^2 / \sigma_0^2) - n_{\text{TOD}}}{\sqrt{2n_{\text{TOD}}}}, \quad (4)$$

which is approximately drawn from the standard normal distribution $\mathcal{N}(0, 1)$.

The COSMOGLOBE Gibbs chain is given by

$$\mathbf{g} \leftarrow P(\mathbf{g} | \xi_n, \mathbf{s}^{\text{inst}}, \boldsymbol{\beta}, \mathbf{a}, C_\ell, \theta) \quad (5)$$

$$\mathbf{n}_{\text{corr}} \leftarrow P(\mathbf{n}_{\text{corr}} | \mathbf{g}, \xi_n, \mathbf{s}^{\text{inst}}, \boldsymbol{\beta}, \mathbf{a}, C_\ell, \theta) \quad (6)$$

$$\xi_n \leftarrow P(\xi_n | \mathbf{g}, \mathbf{n}_{\text{corr}}, \mathbf{s}^{\text{inst}}, \boldsymbol{\beta}, \mathbf{a}, C_\ell, \theta) \quad (7)$$

$$\mathbf{s}^{\text{inst}} \leftarrow P(\mathbf{s}^{\text{inst}} | \mathbf{g}, \mathbf{n}_{\text{corr}}, \xi_n, \boldsymbol{\beta}, \mathbf{a}, C_\ell, \theta) \quad (8)$$

$$\boldsymbol{\beta} \leftarrow P(\boldsymbol{\beta} | \mathbf{g}, \mathbf{n}_{\text{corr}}, \xi_n, \mathbf{s}^{\text{inst}}, C_\ell, \theta) \quad (9)$$

$$\mathbf{a} \leftarrow P(\mathbf{a} | \mathbf{g}, \mathbf{n}_{\text{corr}}, \xi_n, \mathbf{s}^{\text{inst}}, \boldsymbol{\beta}, C_\ell, \theta) \quad (10)$$

$$C_\ell \leftarrow P(C_\ell | \mathbf{g}, \mathbf{n}_{\text{corr}}, \xi_n, \mathbf{s}^{\text{inst}}, \boldsymbol{\beta}, \mathbf{a}, \theta) \quad (11)$$

$$\theta \leftarrow P(\theta | \mathbf{g}, \mathbf{n}_{\text{corr}}, \xi_n, \mathbf{s}^{\text{inst}}, \boldsymbol{\beta}, \mathbf{a}, C_\ell), \quad (12)$$

with each step requiring its own dedicated sampling algorithm, and in the case of BEYONDPLANCK, its own publication. The Commander3 pipeline is designed so that results of each Gibbs sample can be easily passed to each other, and that the internal calculations of each step do not directly depend on the inner workings of each other. Therefore, in order to add another data set to the Gibbs chain, one must implement Eqs. (5)–(8) for each instrument, as was done in ? and ? for *Planck* LFI and in ? for *WMAP*, or simply pass processed maps with beam, mask, and noise information to Eqs. (9)–(12), as was done for the Haslam 408 MHz map (??) and the *Planck* 353 and 857 GHz maps.

Before we discuss the results of this Gibbs chain as applied to the *Planck* LFI and *WMAP* data, we summarize the TOD processing steps in Sects. 2.2 and 2.3.

2.2. Planck LFI data processing

Summary of ?, ?. Artem was here...

2.3. WMAP data processing

As summarized in ?, the TOD processing for *WMAP* has been implemented in the Commander3 Gibbs sampling framework.

¹ <http://healpix.sourceforge.net> (?)

² cosmoglobe.uio.no

The process of implementing this module in Commander3 generally involved implementing the procedures outlined in ?, ?, ?, and ?, and implementing a sampling algorithm whenever an estimate based off of the data was obtained. ? found excellent agreement between our processed maps and the official *WMAP9* products. The most visually obvious exception is that the ? maps have no trace of the poorly-measured modes induced by transmission imbalance uncertainties. Although ? focused on the *Q*-band maps, we show in Sect. 3 that this is representative of all *WMAP* bands, with the notable exception of *K*-band.

The data model adopted in ? can be written in raw digital units (du) as

$$\mathbf{GPBM} \mathbf{a} + \mathbf{n} + \mathbf{b}, \quad (13)$$

where \mathbf{b} is the instrumental baseline and \mathbf{n} is the total instrumental noise. As detailed in ?, Commander3 divides the noise into $\mathbf{n} = \mathbf{n}^w + \mathbf{n}^{corr}$, a white noise term and a correlated noise term. By definition, the white noise does not have any correlations between adjacent pixels, so that any pixel-pixel covariance should be fully described by realizations of the \mathbf{n}^{corr} timestream.

Three notable deviations from the official *WMAP9* pipeline and the COSMOGLOBE approach are the gain estimation, the correlated noise treatment, and the data fraction used. The *WMAP9* gain solution was a parametric fit to the orbital dipole measurements as a function of the instrumental housekeeping parameters, as detailed in Appendix A of ?. This approach avoids the dependence of a robust Galactic foreground model, and requires both excellent knowledge of the radiometers and a clean orbital dipole measurement which can be compared to the known signal. Commander3's sampling approach, as summarized in ?, fits the average gain over all detectors and TOD samples to the orbital dipole. In addition, Commander3 calibrates deviations from the average gain as a function of detector and time to the total sky model. This approach is highly dependent on using a sky model that accurately reflects the data, but makes fewer assumptions regarding the behavior of the detectors in relation to the housekeeping parameters.

The second main deviation is in the treatment the noise power spectra. As shown in Sect 2.5 of ?, the noise autocorrelation spectrum is fit on a year-by-year basis to a polynomial in $\log(\Delta t)$, where Δt is the time lag between data points. This method is very similar to the Commander3 approach, which fits for the power spectrum in Fourier space using a correlated noise model of the form $\sigma_0^2(f/f_{knee})^\alpha$. Properly parameterized, these two approaches should yield similar results, albeit with different levels of uncertainty and time resolution. However, we have confirmed that in many cases the simple $1/f$ noise model does not fit the signal-subtracted TOD, yielding χ^2 values that are up to 10σ discrepant from their expected values. [Show, discuss figure with the PSDs, residual spectrum, and Bessel filter.]

Deviations from the $1/f$ model consist either of a linear increase or downturn above 10 Hz. This can be partially explained by the use of a two pole Bessel low-pass filter just prior to signal quantization, which introduces a 2.62% correlation between 25.6 ms sample integrations (?; Sect. 5.3). The exact form of the Bessel filter was not used on flight data, but rather the parametric fit as discussed above. However, the filter is designed to reduce the signal by half at 100 Hz, and as such has a negligible effect.

The third main difference in the COSMOGLOBE treatment is the amount of data used, and the size of individual scans used. As shown in ?, a large fraction of Commander's computational time is spent performing FFTs on individual scans. Rather than truncating datastreams to have lengths equal to "magic numbers" for which FFTW (?) is fastest, as in ?, we split the data into scans

Table 1. Flagging statistics

Band	Flagged (%)	Discarded (%)	Used (%)
<i>K</i>	1.72	0.87	97.4
<i>Ka</i>	1.64	0.88	97.5
<i>Q1</i>	1.84	0.84	96.5
<i>Q2</i>	1.62	0.81	97.6
<i>V1</i>	1.62	1.10	97.3
<i>V2</i>	1.61	1.01	97.4
<i>W1</i>	1.76	1.03	97.2
<i>W2</i>	1.60	0.81	97.6
<i>W3</i>	1.61	0.87	97.5
<i>W4</i>	1.60	0.81	97.6

of length 2^N , where $N = 22$ for *K-Q*, $N = 23$ for *V-W*. This yields scans with lengths of 6.21 days for *K*- and *Ka*-band, 4.97 days for *Q*-band, 7.46 days for *V*-band, and 4.97 days for *W*-band. These datastream lengths are short enough to be processed quickly and distributed efficiently across multiple processors, while being long enough to properly characterize the noise properties of the timestreams, whose f_{knee} 's are on the order 1 mHz.

? lists a series of flagged events, which we isolated to their own TOD segments. When we encountered these events, TOD segments that were interrupted by the event were appended to the previous TOD, in most cases creating TODs with lengths $> 2^N$. We found that events of length $< 2^N$ were too short to accurately estimate the noise PSD parameters. This criterion led to unflagged data to be discarded. A similar problem occurred when $> 10\%$ of the TOD was flagged. Together, these two effects led to $\simeq 1\%$ of the data to be discarded despite on their face acceptable. We present the full statistics for our maps in Table 1. In total, the COSMOGLOBE maps use slightly less data than the *WMAP9* official products, which had a total efficiency of $\simeq 98.4\%$ (?). The total difference in data can be entirely accounted for by the cuts described in this paragraph.

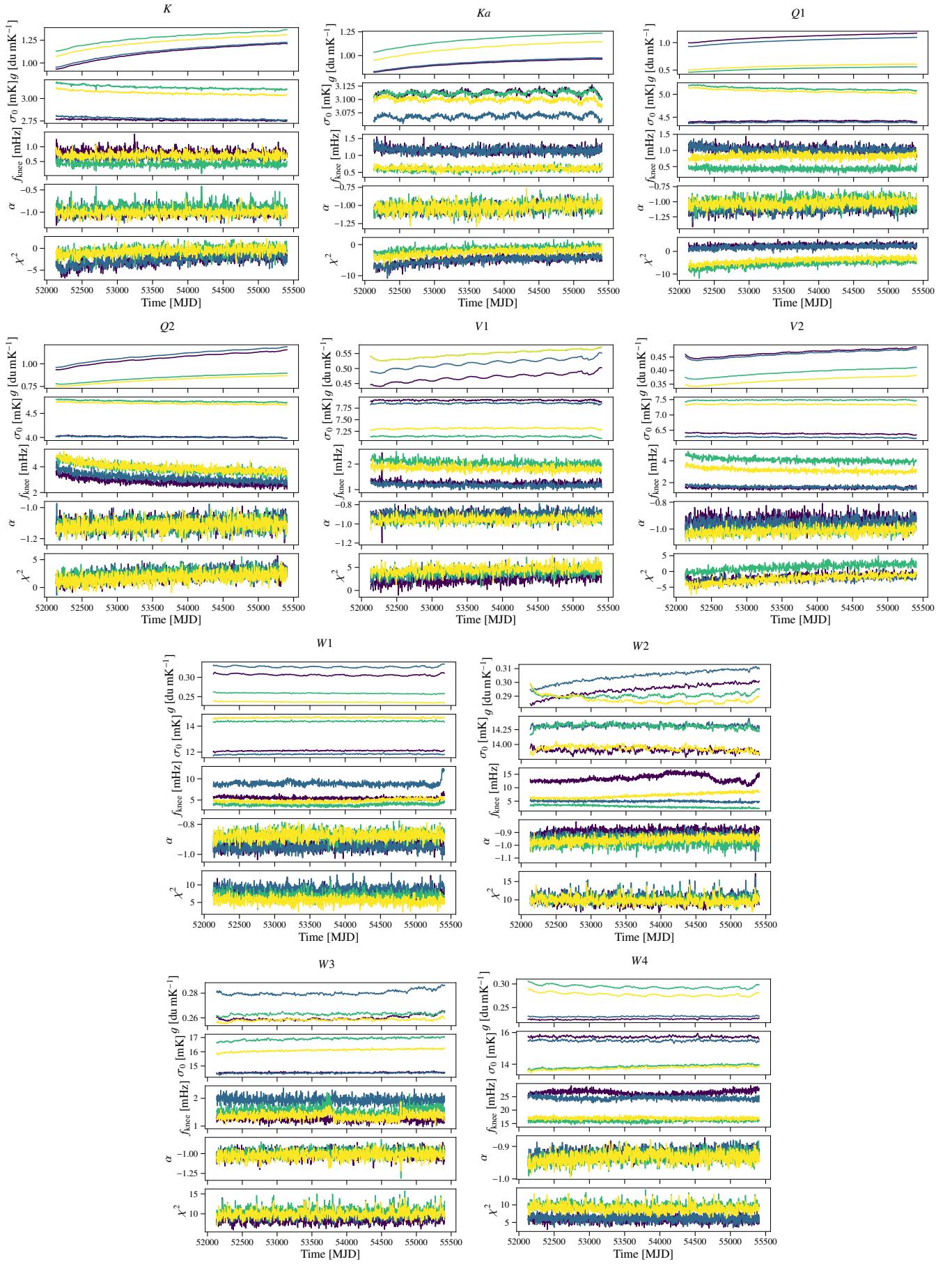
2.4. Calibration

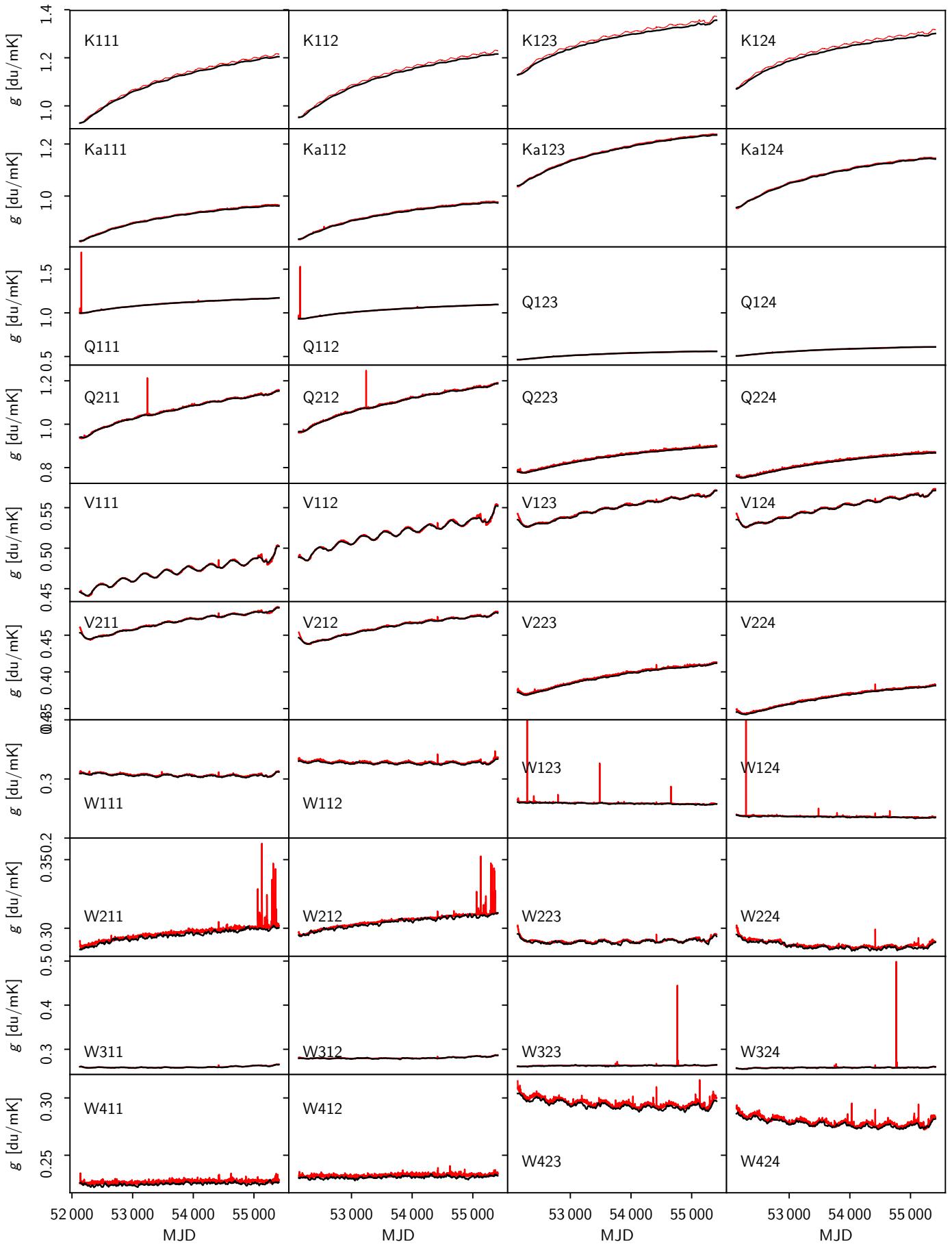
In order to get the raw data into the sky units mK_{CMB}, it is necessary to make a quantitative assessment of the data's goodness-of-fit. To do this, we use mapped time-domain residuals of the timestream based on an initial sky model. Following the procedure of ?, we create a smoothed map of the absolute deviation of the model from the sky signal. The residual maps were first smoothed to 3° , the absolute value was taken, and then the map was smoothed again with a 15° beam. This procedure gave us an indication of the regions of the sky that are not fully characterized by the sky model. We then evaluated the Cosmoglobe Sky model³ associated with the BEYONDPLANCK data release⁴ to create a map of the expected thermal dust, AME, free-free, and radio emission at each *WMAP* frequency. We thresholded using the 80th percentile of this map to match the residual map. The combination of this sky model-informed mask and the residual-informed mask formed the main processing mask for the calibration and correlated noise analysis. Outside of these mask, the data are fit will by the sky model, so the calibration and the correlated noise estimates should both be unbiased.

Emphasize, we use the orbital dipole for the absolute calibration, but use the full sky model from the rest of the Gibbs chain for the rest of the gain and the imbalance parameters.

³ cosmoglobe.readthedocs.io

⁴ beyondplanck.science

**Fig. 1.** Instrument parameters.

**Fig. 2.** Gain.

3. Characteristics of frequency maps

4. Systematic error and uncertainties

5. Component separation

Want to compare the QU correlation in WMAP and Planck LFI, get a quantitative number. Point out that the polarization solution itself is much better, but the covariance between pixels themselves is much higher. This wasn't an issue for LFI, so we had to take that into account here.

I also want to put a bit here on why the low- ℓ approach needed to be done separately, how correlated noise sampling addresses it, to what extent it's mitigated, etc.

Note that LFI's 30 and 70 GHz QU correlation is ~ 0.1 , whereas 44 GHz is much larger, ~ 0.5 . This discrepancy is due to the number of horns with differing polarization orientation. Both 30 and 70 GHz have an even number of horns, allowing for pairs of datastreams to be combined to give independent polarization measurements. Conversely, 44 GHz has one horn pair and an unpaired horn, the latter of which induces more correlation in the QU observation matrix. An example for Ka and 30 GHz is shown in Fig. 3. Aside from the obvious morphological changes due to the two experiments' different observing strategies, the magnitude of WMAP's correlation is much larger than Planck's.

The BEYONDPLANCK project took this covariance structure into account using the dense $N_{\text{side}} = 16$ noise covariance matrix provided by the WMAP team.⁵ Properly sampled correlated noise only leaves white noise in the maps, so the noise properties of each map's sample do not require a dense pixel-pixel covariance, even at low resolution (??). The WMAP9 inverse noise covariance matrices were computed using the full time-space noise matrix $N = N^w + N^{\text{corr}}$, so the full pixel-pixel covariance matrix $\Sigma^{-1} = P^T N^{-1} P$ took into account the correlation between neighboring samples. The COSMOGLOBE maps, by subtracting a realization of correlated noise before mapmaking, estimates an inverse noise covariance matrix

$$\Sigma_{pp'}^{-1} = \sum_{t_1, t_2} P_{t_1, p_1}^T N_{t_1, t_2}^{-1} P_{t_2, p_2} = \sum_t P_{t, p_1}^T N_{t, t}^{-1} P_{t, p_2} \quad (14)$$

How much off-diagonal pixel covariance is there here?

dense noise covariance matrix also explicitly projected out the poorly-measured imbalance modes, but because we find no trace of these modes in our sky maps or residual maps, this treatment is not necessary in our approach. However, the correlation between Stokes Q and U was not taken into account in the BEYONDPLANCK LFI analysis. This was not a significant oversight in the LFI analysis because the 30 and 70 GHz maps only had a 10% correlation, and 44 GHz's 50% correlation was subdominant to other systematic effects. We have updated Commander3 to take QU correlation into account for LFI.

6. Conclusions

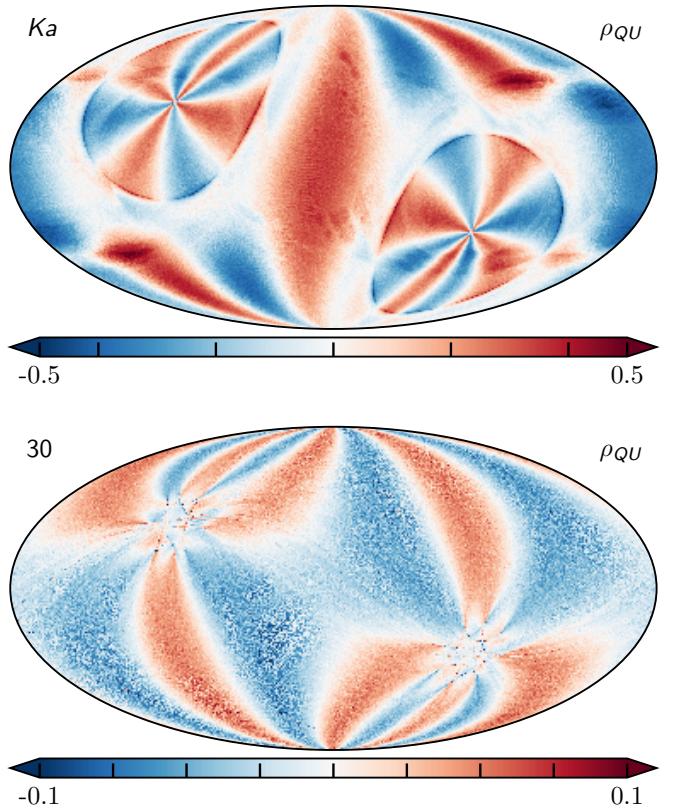
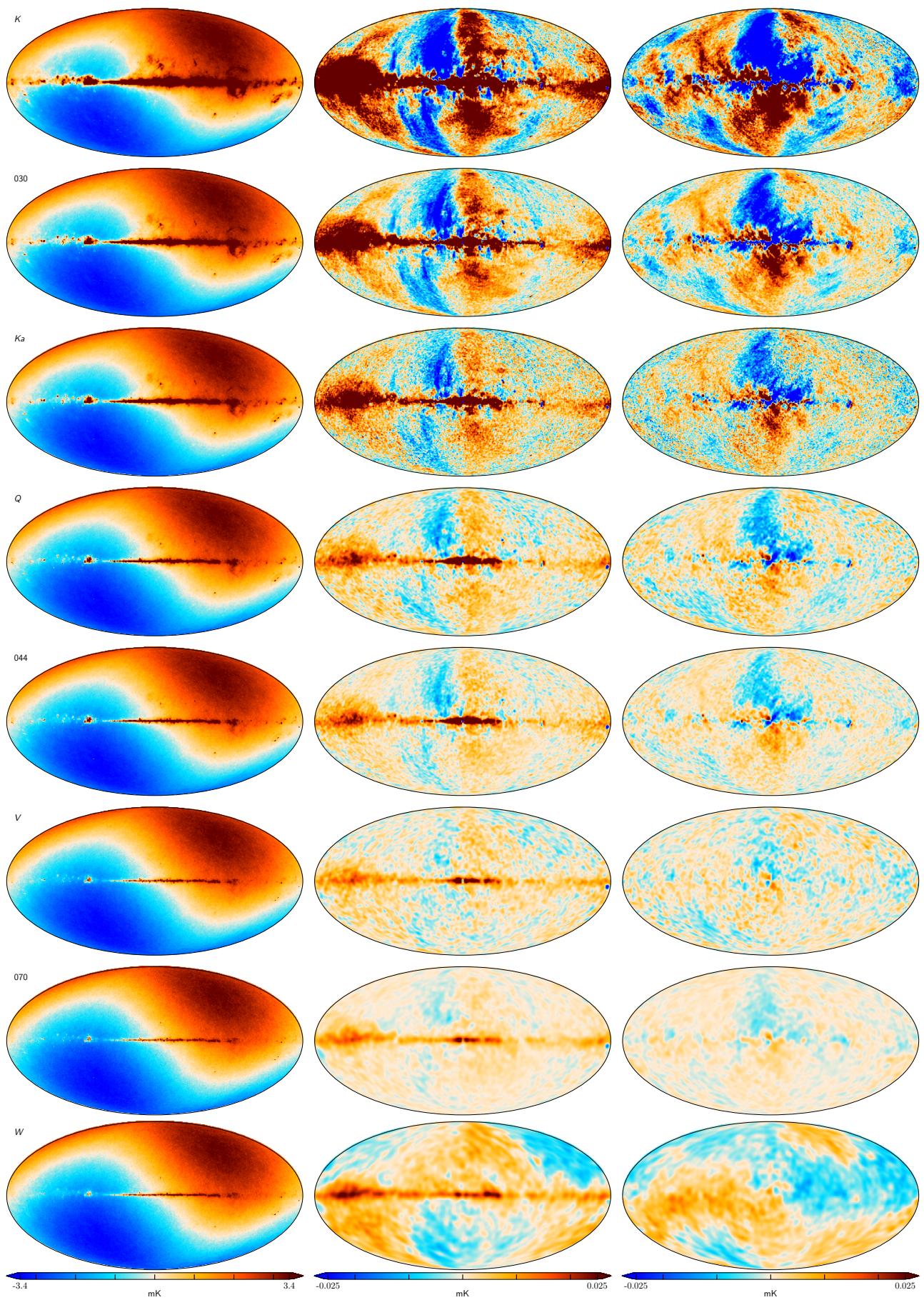
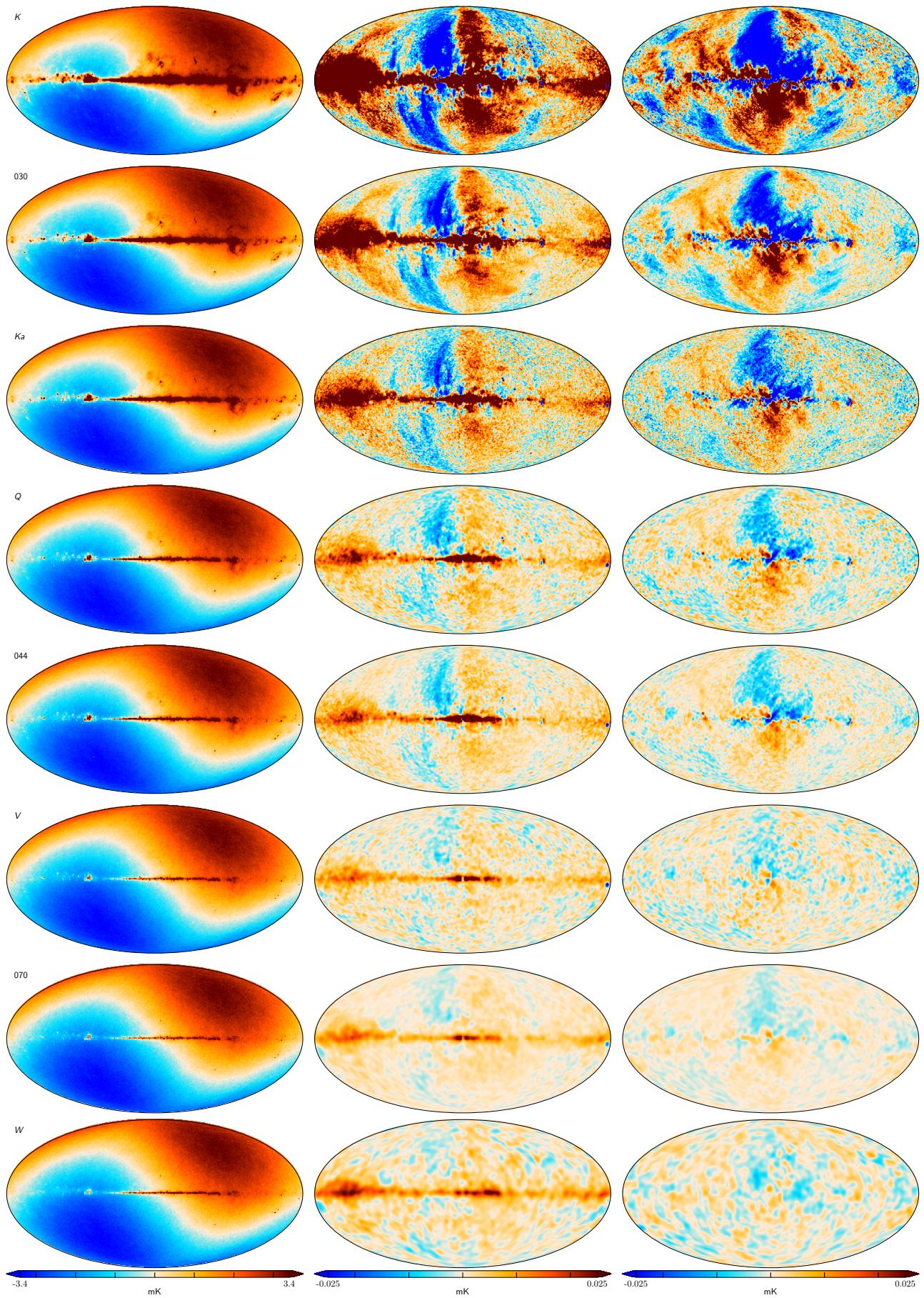


Fig. 3. QU correlation for WMAP Ka -band (top) and Planck 30 GHz (bottom). Note the differing dynamic range in the two maps.

⁵ https://lambda.gsfc.nasa.gov/product/wmap/dr5/ninv_info.html

**Fig. 4.** Published WMAP+LFI bands (WMAP9 and BP10)

**Fig. 5.** Mean of WMAP+LFI bands

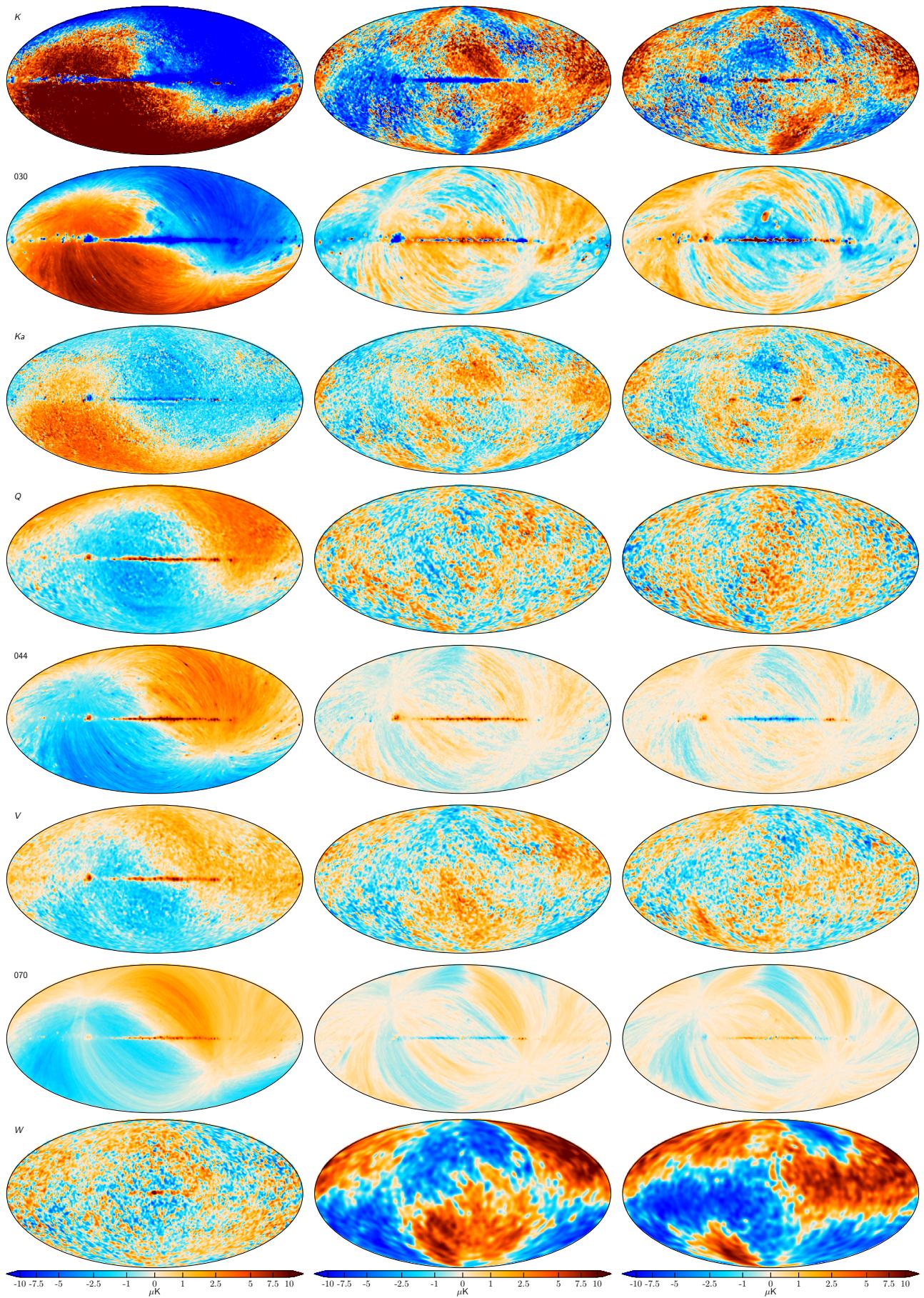


Fig. 6. Differences between Commander-processed chains and official maps (BP10 and WMAP9)