

Planck intermediate results LVII. Joint Planck LFI and HFI data processing

Planck Collaboration, Yashar Akrami, K. J. Andersen, M. Ashdown, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, et al.

▶ **To cite this version:**

Planck Collaboration, Yashar Akrami, K. J. Andersen, M. Ashdown, C. Baccigalupi, et al.. Planck intermediate results LVII. Joint Planck LFI and HFI data processing. Astronomy & Astrophysics - A&A, 2020, 643, pp.A42. ⟨10.1051/0004-6361/202038073⟩. ⟨cea-03001891⟩

HAL Id: cea-03001891

https://cea.hal.science/cea-03001891v1

Submitted on 12 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

A&A 643, A42 (2020)
https://doi.org/10.1051/0004-6361/202038073
© ESO 2020

Astronomy & Astrophysics

Planck intermediate results

LVII. Joint *Planck* LFI and HFI data processing

Planck Collaboration: Y. Akrami[11,49,50], K. J. Andersen[50], M. Ashdown[56,3], C. Baccigalupi[65], M. Ballardini[16,33], A. J. Banday[78,6], R. B. Barreiro[52], N. Bartolo[20,53], S. Basak[71], K. Benabed[48,73], J.-P. Bernard[78,6], M. Bersanelli[23,38], P. Bielewicz[64,65], J. R. Bond[5], J. Borrill[9,76], C. Burigana[37,21,40], R. C. Butler[33], E. Calabrese[68], B. Casaponsa[52], H. C. Chiang[18,4], L. P. L. Colombo[23], C. Combet[58], B. P. Crill[54,8], F. Cuttaia[33], P. de Bernardis[22], A. de Rosa[33], G. de Zotti[34], J. Delabrouille[1], E. Di Valentino[55], J. M. Diego[52], O. Doré[54,8], M. Douspis[47], X. Dupac[26], H. K. Eriksen[50], R. Fernandez-Cobos[52], F. Finelli[33,40], M. Frailis[35], A. A. Fraisse[18], E. Franceschi[33], A. Frolov[72], S. Galeotta[35], S. Galli[48,73], K. Ganga[1], M. Gerbino[29], T. Ghosh[67,7], J. González-Nuevo[13], K. M. Górski[54,79], A. Gruppuso[33,40], J. E. Gudmundsson[77,18], W. Handley[56,3], G. Helou[8], D. Herranz[52], S. R. Hildebrandt[54,8], E. Hivon[48,73], Z. Huang[69], A. H. Jaffe[45], W. C. Jones[18], E. Keihänen[17], R. Keskitalo[9,⋆], K. Kiiveri[17,31], J. Kim[62], T. S. Kisner[60], N. Krachmalnicoff[65], M. Kunz[10,47,2], H. Kurki-Suonio[17,31], A. Lasenby[3,56], M. Lattanzi[41,21], C. R. Lawrence[54], M. Le Jeune[1], F. Levrier[74], M. Liguori[20,53], P. B. Lilje[50], M. Lilley[48,73], V. Lindholm[17,31], M. López-Caniego[26], P. M. Lubin[19], J. F. Macías-Pérez[58], D. Maino[23,38,42], N. Mandolesi[33,21], A. Marcos-Caballero[52], M. Maris[35], P. G. Martin[5], E. Martínez-González[52], S. Matarrese[20,53,28], N. Mauri[40], J. D. McEwen[63], P. R. Meinhold[19], A. Mennella[23,38], M. Migliaccio[25,43], S. Mitra[44,54], D. Molinari[21,33,41], L. Montier[78,6], G. Morgante[33], A. Moss[70], P. Natoli[21,75,41], D. Paoletti[33,40], B. Partridge[30], G. Patanchon[1], D. Pearson[54], T. J. Pearson[8,46], F. Perrotta[65], F. Piacentini[22], G. Polenta[75], J. P. Rachen[14], M. Reinecke[62], M. Remazeilles[55], A. Renzi[53], G. Rocha[54,8], C. Rosset[1], G. Roudier[1,74,54], J. A. Rubiño-Martín[51,12], B. Ruiz-Granados[51,12], L. Salvati[32,36], M. Savelainen[17,31,61], D. Scott[15], C. Sirignano[20,53], G. Sirri[40], L. D. Spencer[68], A.-S. Suur-Uski[17,31], L. T. Svalheim[50], J. A. Tauber[27], D. Tavagnacco[35,24], M. Tenti[39], L. Terenzi[33], H. Thommesen[50], L. Toffolatti[13,33], M. Tomasi[23,38], M. Tristram[57], T. Trombetti[37,41], J. Valiviita[17,31], B. Van Tent[59], P. Vielva[52], F. Villa[33], N. Vittorio[25], B. D. Wandelt[48,73], I. K. Wehus[50], A. Zacchei[35], and A. Zonca[66]

(Affiliations can be found after the references)

Received 2 April 2020 / Accepted 1 July 2020

ABSTRACT

We present the NPIPE processing pipeline, which produces calibrated frequency maps in temperature and polarization from data from the *Planck* Low Frequency Instrument (LFI) and High Frequency Instrument (HFI) using high-performance computers. NPIPE represents a natural evolution of previous *Planck* analysis efforts, and combines some of the most powerful features of the separate LFI and HFI analysis pipelines. For example, following the LFI 2018 processing procedure, NPIPE uses foreground polarization priors during the calibration stage in order to break scanning-induced degeneracies. Similarly, NPIPE employs the HFI 2018 time-domain processing methodology to correct for bandpass mismatch at all frequencies. In addition, NPIPE introduces several improvements, including, but not limited to: inclusion of the 8% of data collected during repointing manoeuvres; smoothing of the LFI reference load data streams; in-flight estimation of detector polarization parameters; and construction of maximally independent detector-set split maps. For component-separation purposes, important improvements include: maps that retain the CMB Solar dipole, allowing for high-precision relative calibration in higher-level analyses; well-defined single-detector maps, allowing for robust CO extraction; and HFI temperature maps between 217 and 857 GHz that are binned into 0.′9 pixels ($N_{\rm side} = 4096$), ensuring that the full angular information in the data is represented in the maps even at the highest *Planck* resolutions. The net effect of these improvements is lower levels of noise and systematics in both frequency and component maps at essentially all angular scales, as well as notably improved internal consistency between the various frequency channels. Based on the NPIPE maps, we present the first estimate of the Solar dipole determined through component separation across all nine *Planck* frequencies. The amplitude is $(3366.6 \pm 2.7)\,\mu$K, consistent with, albeit slightly higher than, earlier estimates. From the large-scale polarization data, we derive an updated estimate of the optical depth of reionization of $\tau = 0.051 \pm 0.006$, which appears robust with respect to data and sky cuts. There are 600 complete signal, noise and systematics simulations of the full-frequency and detector-set maps. As a *Planck* first, these simulations include full time-domain processing of the beam-convolved CMB anisotropies. The release of NPIPE maps and simulations is accompanied with a complete suite of raw and processed time-ordered data and the software, scripts, auxiliary data, and parameter files needed to improve further on the analysis and to run matching simulations.

Key words. cosmic background radiation – cosmology: observations – cosmological parameters – Galaxy: general – methods: data analysis

1. Introduction

This paper, the last by the Planck Collaboration, describes a *Planck*[1] data-processing pipeline called NPIPE and the calibrated data, maps, simulations, and other data products that it produces. NPIPE represents a natural evolution of the previous *Planck* analysis efforts made within the *Planck* data processing centres (DPCs) in Paris and Trieste, but is uniquely designed to analyse data from both the Low Frequency Instrument (LFI) and the High Frequency Instrument (HFI) within the same framework. Furthermore, NPIPE is implemented to execute efficiently on massively parallel high-performance computing (HPC) systems. Indeed, it was both developed and run on the HPC facilities hosted by the National Energy Research Scientific Computing Center (NERSC), motivating the "N" in NPIPE. The main technical design criteria for NPIPE are therefore:

⋆ Corresponding author: R. Keskitalo,
e-mail: rtkeskitalo@lbl.gov

[1] *Planck* (http://www.esa.int/Planck) is a project of the European Space Agency (ESA) with instruments provided by two scientific consortia funded by ESA member states and led by Principal Investigators from France and Italy, telescope reflectors provided through a collaboration between ESA and a scientific consortium led and funded by Denmark, and additional contributions from NASA (USA).

– the ability to handle both LFI and HFI data;
– efficient execution on HPC systems;
– adaptability to evolving HPC architectures;
– *Planck* data access based on Exchange File Format files rather than databases;
– minimal I/O during processing.

Running in a massively parallel HPC environment, NPIPE overcomes some of the practical difficulties and limitations imposed on the pipelines operated by the *Planck* DPCs. With the above design requirements in mind, NPIPE developed incrementally, with the first step being a prototyping platform for testing new implementations of existing HFI preprocessing modules. Early success in 4 K line removal and inclusion of the data taken during repointing manoeuvres (excluded in the DPC processing) presaged the development of a full data-processing pipeline from raw time-ordered data (TOD) all the way to maps. Since most of the DPC pipeline modules were deeply integrated into the database-driven architecture at the DPC, this led to fresh implementations of a majority of the processing modules. The capability to handle LFI data was added while studying reduction of the high-frequency noise in the LFI TOD through improved decorrelation of the $1/f$ noise.

Having a single data-processing pipeline that could handle both LFI and HFI data led to significant improvements in the calibration procedures, which combine the strengths of both LFI and HFI calibration. Specifically, the NPIPE calibration procedure is a synthesis of the high S/N approach that works well for the HFI 353 GHz channel and the noise-limited approach that is critical to LFI 70 GHz calibration. Numerous smaller improvements to the data-processing modules were motivated by experiences with the nine *Planck* frequencies and their diverse features. The LFI- and HFI-DPC pipelines evolved in the post-launch period, often in response to instrument-specific effects that emerged as the calibration accuracy improved. NPIPE builds on knowledge accumulated through the years, and takes a global approach, aiming at coherent treatment of the full *Planck* data set.

In what follows, we detail NPIPE processing and a full data set of *Planck* maps that result from it. These NPIPE data maps are supported by a suite of simulated maps that capture the relevant noise and systematic errors present in the data. We also compare the NPIPE results to the two previous public releases of *Planck* temperature and polarization maps, namely the second data release in 2015 ("PR2"; Planck Collaboration I 2016) and the third release in 2018 ("PR3"; Planck Collaboration I 2020). Comparison with both earlier releases allows the reader to assess the magnitude of the differences we find.

To characterize the new maps, we consider a few important applications for which component separation is essential. Indeed, robustly supporting component-separation applications ranks among the primary motivations for the NPIPE products, and several features have been added specifically to meet the requirements of future astrophysical analysis. Among these are: maps that retain the cosmic microwave background (CMB) Solar dipole, and thereby allow for joint component separation and high-precision relative calibration; robust single-detector temperature-only sky maps that allow for fine-grained CO extraction, and thereby weaker degeneracies with respect to CMB, thermal dust, and free-free emission; and HFI maps between 217 and 857 GHz pixelized at a HEALPix[2] (Górski et al. 2005) resolution of $N_{\rm side} = 4096$, corresponding to a pixel size of 0.′88, which ensures properly signal-bandwidth-limited (not limited by pixel size) sampling of the highest resolution *Planck* beams of 5′.

Taking advantage of the new NPIPE maps, we present the first *Planck*-only astrophysical sky model that simultaneously constrains the CMB Solar dipole, higher-order CMB temperature fluctuations, low-frequency foregrounds, thermal dust, and CO line emission. This joint analysis has at least three major advantages compared to previous analyses. First, it results in a global LFI+HFI CMB Solar dipole estimate with minimal foreground contamination. Second, it eliminates the need for estimating dipole residuals per frequency band during component separation, and thereby reduces large-scale degeneracies in the astrophysical component maps. Third, it allows for high-precision relative, inter-frequency gain calibration during component separation. The resulting astrophysical components are compared to corresponding previous releases, and are found to be consistent with earlier maps, but exhibiting lower noise and systematics. Similarly, the derived CMB dipole parameters are also consistent with previous estimates, although the new result has a slightly higher amplitude. Properly taking into account uncertainties from higher-order CMB fluctuations (Thommesen et al., in prep.), the NPIPE results are consistent with the *Planck* 2018 constraints at the 1–2 σ level.

Finally, we constrain the optical depth of reionization τ with the large-scale polarization NPIPE maps. Due to the low levels of systematic residuals in these maps, we observe good consistency among results derived from different frequency maps, and the resulting large-scale angular CMB power spectrum and cosmological parameters appear robust with respect to both data and sky cuts.

The release of NPIPE maps and simulations is accompanied with the full software suite and input data so that the results can be reproduced, and even improved, given sufficient computational resources. The rest of the paper is organized as follows. In Sect. 2 we summarize the low-level NPIPE data processing, highlighting in particular where NPIPE differs from the DPC processing. In Sect. 3 we comment on the destriping mapmaking algorithm employed by NPIPE. In Sect. 4 we present and characterize the NPIPE frequency and detector maps, and in Sect. 5 we describe the associated end-to-end simulations. In Sect. 6 we compare NPIPE results with those of previous data releases. In Sect. 7 we discuss the astrophysical component maps derived from the NPIPE-only data set, before considering the CMB Solar dipole in Sect. 8 and the optical depth of reionization in Sect. 9. Conclusions are presented in Sect. 10, and algorithmic details are provided in the appendices.

2. Data processing

NPIPE re-implements most of the Level 2 data processing performed at the *Planck* DPCs (Planck Collaboration II 2019; Planck Collaboration III 2020), while introducing a number of detailed changes that will be described later in this section. We divide our processing into two main steps.

Local *preprocessing* covers all the steps that can take place at the single-detector[3] level without projecting the signal onto

[2] https://healpix.sourceforge.io. The square roots of the uniform pixel areas at $N_{\rm side} = 1024$, 2048, and 4096 are 3.′4, 1.′7, and 0.′88, respectively.

[3] In this paper, a "detector" is understood to be an LFI radiometer or an HFI bolometer. Each LFI horn and each polarized HFI horn feeds two linearly polarized detectors aligned in orthogonal directions. Note that each LFI radiometer actually comprises two detector diodes that are co-added to optimize rejection of coherent fluctuations.

maps. Furthermore, every preprocessing step (except for the mission averaging of the LFI low-pass filter and spike templates) can be performed at the single-pointing-period[4] level (see Sect. 2.3.9) without reference to data beyond it.

Global *reprocessing* encompasses single- and multi-detector operations that require much or all of the data to be in memory. Examples of systematics handled during reprocessing are gain fluctuations and residuals from the low-frequency bolometer transfer function.

The data are written to disk after preprocessing and reprocessing to flexibly allow development of these two major steps. Furthermore, mapmaking of the reprocessed TOD can be separately optimized after the data are processed.

The parts of NPIPE data processing that continue to employ DPC-derived products are the following.

– LFI analogue-to-digital converter (ADC) correction. We apply the DPC-provided correction profiles.

– HFI Initial ADC correction measurement. We start from the DPC-provided correction profiles.

– HFI bolometer transfer-function measurement. We deconvolve the DPC-measured transfer function.

– Beam and focal plane geometry measurements. Such changes can be implemented in NPIPE, but the data set described and released here retains the DPC values. Neither are expected to differ from PR3, since we have not updated the bolometer transfer functions.

2.1. Differences between NPIPE and the PR3 data processing

We list here the most important differences between NPIPE and PR3 data processing, with section numbers for further details.

– We use a hybrid calibration scheme to work around a degeneracy in the *Planck* calibration process. The 30 and 353 GHz data are first calibrated along with the polarized sky much like PR3 HFI is calibrated. We then mimic the PR3 LFI calibration and calibrate the intermediate CMB frequencies (44–217 GHz), using an approximation that the sky polarization can be captured with templates derived from the foreground frequencies (Sect. 2.4.13). The calibration for the CMB frequencies is thus based on temperature-only sky and polarized foreground templates (see Sect. 2.4.13). The use of this polarization prior significantly reduces the uncertainty in gain and the large-scale polarization systematics, but also requires a measurable and correctable transfer function affecting large-scale CMB polarization (described in Sect. 4.3).

– We include the *Planck* repointing manoeuvre data in all of our processing, leading to a roughly 8% increase in total integration. time (2.3.1).

– We retain the Solar dipole in the frequency maps for component separation.

– We include in our pointing solution the latest star-camera distortion corrections that were developed for *Herschel* data processing (Sect. 2.2.2).

– We correct all LFI radiometers for 1 Hz housekeeping spikes, rather than just the 44 GHz ones (Sect. 2.3.5).

– We fit for more 4 K cooler lines, and use the global signal estimate (Sect. 2.3.2) for signal removal (Sect. 2.3.5).

– We modify the LFI sky–load differencing to include a low-pass filter that reduces the injection of uncorrelated noise *into* the differenced signal (Sect. 2.3.6).

– We use our own signal estimate during HFI glitch detection and removal (Sect. 2.3.4).

– We extend the HFI glitch flags during transfer function deconvolution, to avoid leaking constrained-realization power from the gaps into the surrounding data. This leads to less small-scale noise and lower noise correlations between half-rings[5] (Sect. 2.3.9).

– We fit for more HFI transfer-function harmonics, to better address the odd-even survey differences (Sect. 2.4.5).

– We subtract only the seasonally varying part of zodiacal emission, in order not to bias the dust component in the maps (Sect. 2.4.6).

– We provide polarization-corrected, single-detector, and single-horn temperature maps at all frequencies (Sects. 4.1.3 and 4.1.4).

– We correct the HFI polarization angles and efficiencies, to address polarized signal-like residuals in single-detector maps (Sect. 2.4.15).

– We bin our single-detector maps from 217 to 857 GHz at $N_{\rm side} = 4096$, to fully sample the narrow beam.

– We make detector-set maps for cross correlation analysis and systematics-level studies that are fully independently processed (Sect. 4.1.5).

– We provide a consistent, low-resolution data set, with estimates of pixel-pixel noise covariance across all *Planck* frequencies (Sect. 4.2).

– We fit for signal distortion to address second-order ADC nonlinearity in the HFI CMB channels from 100 to 217 GHz (Sect. 2.4.2).

– We extend to LFI data the HFI approach of correcting for bandpass mismatch in the time domain (Sect. 2.4.7).

– We destripe all data with Madam, using extremely short 167 ms baselines (Sect. 3).

– We calibrate with apodized masks that allow us to access most of the sky while down-weighting Galactic regions where even small issues in the signal model can cause problems with calibration (Sect. 2.4.12).

2.2. Inputs

The inputs to NPIPE are the raw, digitized data as they were transmitted from the spacecraft and decompressed by the DPC Level 1 processing (Planck Collaboration II 2014; Planck Collaboration VI 2014). The HFI TOD are initially corrected with the same ADC nonlinearity profiles as PR3 data. The data are stored in *Planck* exchange file format FITS files that we have extended to accommodate raw, undifferenced LFI timelines.

Each LFI horn feeds two polarization-orthogonal radiometers ("detectors" in this paper, see footnote 3). Each radiometer comprises two diodes that alternate between observing the sky and a 4 K reference load. In total, each LFI horn provides four sky timelines and four reference timelines. From the 11 LFI horns there are then 88 discrete timelines that are preprocessed into 22 detector timelines.

For each polarized HFI horn there are two polarization-sensitive bolometers (PSBs labelled "a" and "b") at orthogonal

[4] A pointing period or ring consists of the time to orient the spacecraft spin axis with three precise thruster burns (the *repointing manoeuvre*) and the continuous science scan that follows the repointing. Each repointing manoeuvre lasted about four minutes, with the majority of time spent passively slewing the spacecraft spin axis into the new orientation. The spin axis was adjusted by 2′ during a standard repointing. The subsequent science scan lasted between 35 and 75 min.

[5] Half-ring data sets are built from the first and second halves of the pointing periods.

angles of polarization sensitivity. Some of the HFI horns instead house nearly unpolarized "spiderweb" bolometers (SWBs). There are 36 HFI horns in total, 16 polarized horns and 20 nearly unpolarized. In addition, two dark bolometers measured the thermal baseline on the focal plane. With two of the SWBs irrecoverably compromised by random telegraphic signal, we have 52 bolometer timelines that are preprocessed into 50 optical timelines.

2.2.1. Time span

HFI science operations spanned the period between 12 August 2009 and 15 January 2012, comprising almost five sky surveys[6], 885 operational days (29 months), and 27 000 pointing periods. We follow Planck Collaboration III (2020) and exclude the last 22 days (956 pointing periods) of the HFI science data due to large drifts in the thermal baseline. This leaves the fifth HFI sky survey only 80% complete. The drifts were considered harmful both due to the correlated noise fluctuations they induce and due to the challenges the changes in the baseline introduce to ADC nonlinearity correction.

LFI science operations continued until 10 October 2013, comprising 8.3 sky surveys, 1 513 operational days (49.6 months), and 46 000 pointing periods. The LFI maps in PR3 are based on the eight completed sky surveys (Planck Collaboration II 2019). NPIPE maps include the additional 60 days of the ninth sky survey.

2.2.2. Pointing

The boresight pointing in NPIPE is based on the same raw attitude information as PR3. We processed the raw attitude quaternions to include a star-camera field-of-view distortion correction that was developed for *Herschel* and updates some guide-star positions necessitated by the delay in the *Planck* launch (Tuttlebee 2013). The corrections are only a few arc seconds at most, but are systematic rather than statistical.

Reprocessing the attitude history allowed us to extend the low-pass filtering used to suppress high-frequency noise in the reconstructed pointing. In PR3, the filtering only covered the stable science scans, leaving the repointing manoeuvres considerably noisier. In NPIPE the same low-pass filter is used on both science scans and repointing manoeuvres, better supporting our use of the repointing manoeuvre data in mapmaking.

NPIPE uses the same PTCOR correction (Planck Collaboration I 2016) as PR3 for the thermal deformation of the angle between the boresight and the star camera. For convenience, we apply the correction directly to the attitude history files (AHFs), thus removing the need to apply the correction every time the pointing is read from the AHFs.

2.3. Local preprocessing

We present flow charts of the preprocessing steps for LFI in Fig. 1, dark HFI bolometers in Fig. 2, and optical HFI bolometers in Fig. 3. The flow charts are annotated with the appropriate section numbers.

2.3.1. Repointing manoeuvres

NPIPE was designed to retain the data from the times between fixed pointings of the spin axis that were previously neglected. While these repointing manoeuvres are marked by changes in the rotation of the spacecraft, these changes alter the angular momentum as little as possible to conserve propellant. This means that the scanning rate on the sky does not change and the effective beam is unaffected. The DPC pipelines omitted these data for a number of reasons. Firstly, they lack the repetitive overlapping scans that allow effective compression of the data onto ring maps. Secondly, the pointing solution is less accurate. Lastly, there was some concern that the thermal environment onboard the spacecraft was less stable during and immediately after thruster burns. With the benefit of hindsight, it is now clear that none of these concerns is very serious.

NPIPE deals with the lack of repetitive scans by using a global signal estimate (Sect. 2.3.2), and only using ring-compressed data for calibration and other template fitting. Our pointing solution for the repointings is more stable than the earlier versions, due to our use of a low-pass filter to suppress small-timescale fluctuations (Sect. 2.2.2). The concern for thermal fluctuations is greatly alleviated by our application of horn-symmetric detector weights (Sect. 2.4.11), leading to effective cancellation of thermal common modes within polarized horns. Tests show that the repointing-period data are perfectly usable (Sect. 4.4).

2.3.2. Global signal estimate

Many of the preprocessing modules require an estimate of the detected sky signal. In particular, the HFI-DPC modules make extensive use of a phase-binned[7] signal estimate, where data from a single pointing period and a single detector are used to estimate the periodic signal present in the data. This binning is made possible by the repetitive *Planck* scanning strategy, which repeats the same scanning circle 39–65 times at a fixed rate of one revolution per minute. Binning the data according to the spacecraft spin phase produces an unbiased estimate of the signal, but the phase-binned data contain a significant amount of noise that is made scan-synchronous by the binning. Moreover, the phase-binned estimate is not applicable to the 4 min repointing manoeuvres at the beginnings of each science scan.

Except in glitch removal, NPIPE uses an alternative that we call the "global signal estimate". We sample a full-mission, single-detector temperature map from a previous iteration of NPIPE, together with polarization from the full set of detectors within each frequency channel. Using the last run of the pipeline as input to preprocessing is justified by the fact that the relevant, high S/N modes of the maps converged early in the development process[8]. This map already contains the Solar dipole. We further add an estimate of the orbital dipole based on the known spacecraft velocity. The time-dependent orbital dipole is not included in the map. For HFI, this global estimate of the sky signal is

[6] The scan strategy was such that *Planck* observed about 95% of the sky over the course of six months, but the entire sky exactly twice every 12 months. The approximately six-month periods are called "surveys". See Sect. 4 of Planck Collaboration I (2014), and particularly Table 1, for precise definitions.

[7] Each pointing period consists of tens of repeating scanning circles. Individual detector data for one period form a one-dimensional data set that can be indexed and binned according to the spin phase of the spacecraft.

[8] In fact, the large-scale temperature signal has remained essentially unchanged since Public Release 1 in 2013 (Planck Collaboration I 2014) except for small overall calibration adjustment. One telling measure of the robustness of the temperature map is the small component separation residual discussed in Sect. 7.

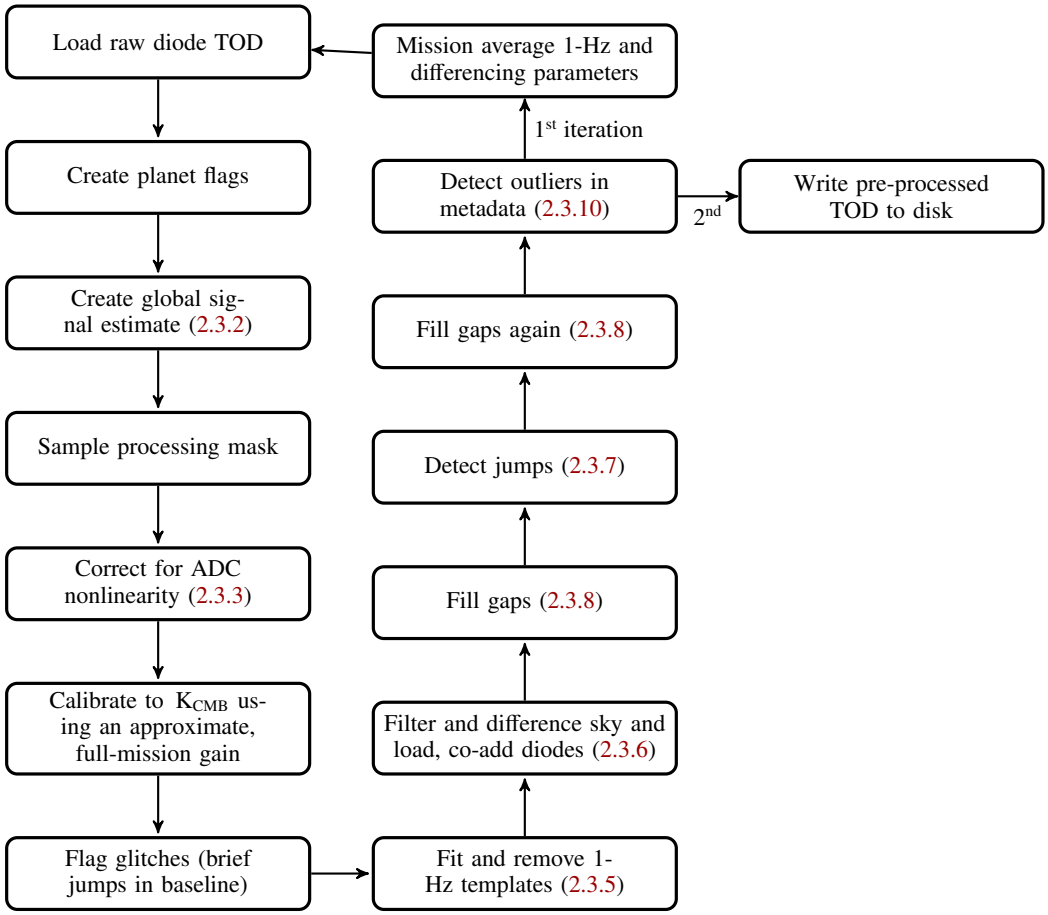


Fig. 1. NPIPE preprocessing flow chart for LFI, indicating the sections in which major steps are discussed. The main loop runs twice, first fitting the 1 Hz spike templates and low-pass filter parameters independently on each pointing period, then mission-averaging the template amplitudes and filter parameters and running the entire processing pipeline again with these fixed values.

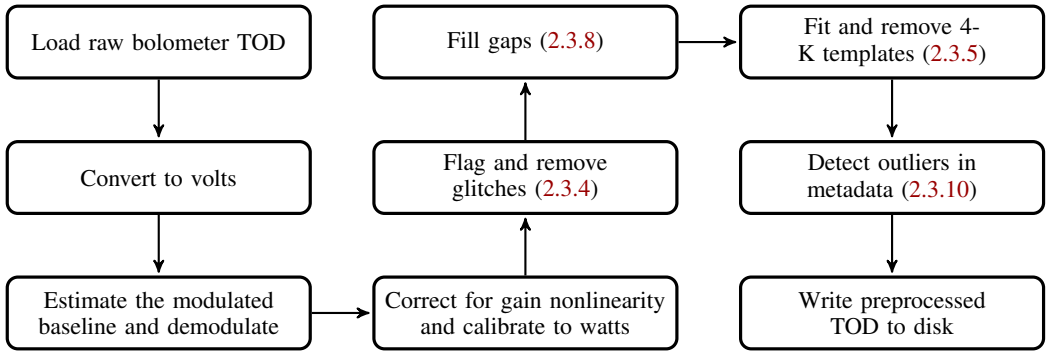


Fig. 2. NPIPE preprocessing flow chart for dark HFI bolometers.

then convolved with the estimated bolometer transfer function in order to produce the required estimate of the sky signal.

2.3.3. ADC nonlinearity correction

Both *Planck* instruments experience detectable levels of nonlinearity in their analogue-to-digital conversion (ADC). The level of nonlinearity in LFI data is much less than in HFI, and is straightforward to correct. NPIPE applies the LFI DPC correction profiles as part of our preprocessing. Details of the LFI nonlinearity and the measured correction can be found in Planck Collaboration II (2019).

For HFI nonlinearity in the digitization was found to be the most important systematic error affecting large-scale polarization (see discussion in Planck Collaboration III 2020; Delouis et al. 2019). The first-order manifestation of the nonlinearity is apparent gain changes that correlate with changes in the level of signal in the detector. Secondarily we observe distortions of the signal itself.

After the end of HFI observations, a two-year campaign was carried out during the LFI extended mission to gather statistical information about the ADC effects. The problem is not fully tractable, since the digitized data transmitted to Earth are downsampled onboard the spacecraft. Nevertheless, the derived ADC

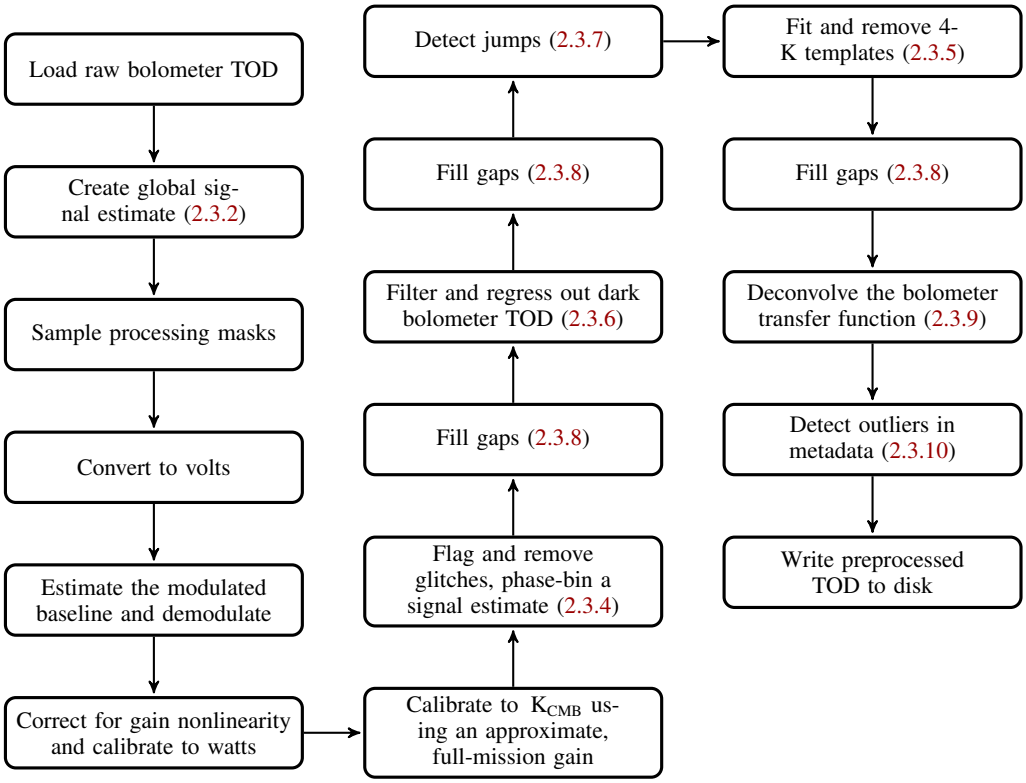


Fig. 3. NPIPE preprocessing flow chart for optical HFI bolometers.

correction profiles greatly improve the effective gain stability in the data, as well as consistency between the odd- and even-parity data sets split from the square-wave-modulated data (Planck Collaboration VII 2016). Even after applying these ADC correction profiles, however, residual ADC nonlinearity (ADCNL) remains the primary systematic error in the large-scale HFI polarization data (Planck Collaboration Int. XLVI 2016). NPIPE begins with the same initial ADCNL-corrected TOD as PR3, but uses a different model to measure and remove the residual ADCNL (see Sect. 2.4.2).

2.3.4. Glitch removal

Cosmic-ray glitches in the HFI data (see Planck Collaboration X 2014) are detected, fitted, and removed in NPIPE using the same despike software that runs in the HFI-DPC pipeline Planck HFI Core Team (2011), Planck Collaboration VI (2014). However, NPIPE does its own signal removal and runs despike in the "dark" mode developed for processing the dark bolometer data.

For the signal estimate, we reorder the signal by the satellite spin phase and fit a sixth-order polynomial in a sliding window of three bin widths at a time, keeping the polynomial fit in the centre bin. Each bin is as wide as the average full width at half maximum (FWHM) of the beam at that frequency. The DPC signal estimate uses 1.′5 bins regardless of the optical beam width.

We also experimented with using the global signal estimate in place of the phase-binned signal estimate. This has the advantage that the signal estimate does not include noise from the pointing period that is being analysed. However, since the signal estimate and the resulting glitch residuals touch virtually every datum in HFI, it was safer to use the local signal estimate instead of potentially biasing the entire data set with errors in the global signal estimate.

Using a phase-binned signal estimate introduces a noise bias in the glitch fits as the fitted glitch templates attempt to accommodate some of the noise in the signal estimate. This bias was evident in the noise estimates derived from half-ring difference maps, which were about a percent short of the full-data noise due to correlated noise cancellation (Planck Collaboration XII 2016). Our FWHM-wide binning scheme and the extra smoothing from the polynomial fits limits the amount of noise in the signal estimate and the amount of half-ring bias. Nevertheless, users of NPIPE products are advised to be cognizant of the fact that the small-scale noise in the half-ring maps is correlated at a low level due to correlated errors in the glitch residuals.

We have raised the glitch detection threshold to reflect improvements in the signal estimate and to reduce the rate of false positives. Our short-baseline destriping (see Sect. 3) is also better equipped to remove residual glitch tails than the ring offset model applied in DPC processing. The net effect of all the changes in glitch detection is to reduce the overall glitch detection rates to an average of 8% in each detector, about half that of the DPC pipeline. While this may seem like a large difference, the energy distribution of the glitches is such that the affected glitch population is mostly right at the detection limit. In Sect. 6 we show that the resulting NPIPE maps have lower noise than their 2018 counterparts, despite allowing more glitches through, and even after correcting for the added integration time.

2.3.5. Removal of frequency spikes

The drive electronics for the 4 K cooler cause interference in the HFI data at many harmonics. We follow the DPC approach in fitting time-domain sine and cosine templates at the harmonic frequencies. The DPC pipeline fitted for nine harmonics, namely, 10, 16.7, 20, 30, 40, 50, 60, 70, and 80 Hz. We have

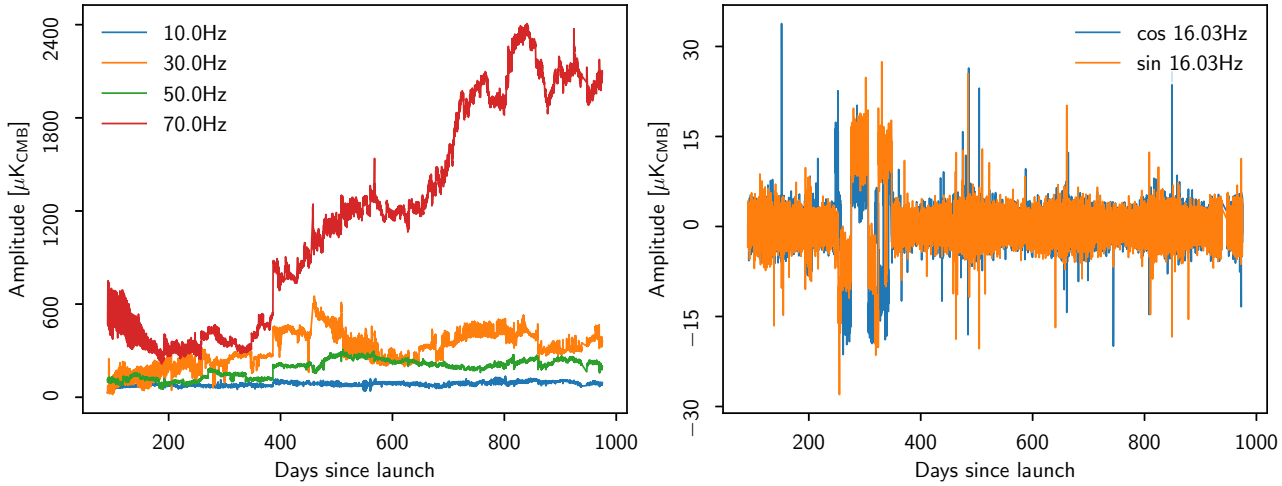


Fig. 4. Interference from the drive electronics of the 4 K cooler in the HFI detectors. *Left*: measured 4 K line amplitudes for the most intense lines in bolometer 143-3a. *Right*: new 16.03 Hz line template amplitudes for bolometer 100-1b. For unknown reasons, the line is only detected intermittently around day 300. For reference, the CMB temperature fluctuations typically fit within $\pm 250\,\mu$K and polarization fluctuations within $\pm 2\,\mu$K; it is therefore clear that the cooler lines must be modelled and removed to high precision. With the *Planck* scan rate being $6^\circ\,\mathrm{s}^{-1}$, the 10 Hz line corresponds to an angular scale of $36'$ and multipole $\ell \approx 600$. For higher-order harmonics of 10 Hz, divide the angular scale and multiply the multipole with the appropriate factor.

added a further six harmonics that are transiently detected in some of the bolometers, namely 16.0, 25.0, 43.4, 46.2, 47.6, and 56.8 Hz.

The line phases and amplitudes in the data are observed to change over timescales of minutes, although the source of the variation is not understood. Fitting for the lines at the pointing period level leaves detectable residuals. We have designated six harmonics as "intense", meaning that their S/N allows for shorter fitting period (5 min). These intense lines are (in approximate order of intensity) at 70, 30, 10, 50, 16.7, and 20 Hz. The 5 min fitting period is a compromise between leaving line residuals and applying a notch filter to the instrumental noise. Figure 4 shows the measured line amplitude for some of the lines in a particularly badly-affected bolometer, 143-3a.

Our pipeline uses the global signal estimate rather than a phase-binned signal estimate for spike removal. The use of this approach avoids a resonant ring issue where the spin rate in some of the pointing periods causes the harmonic templates to be synchronous with the sky signal.

LFI data show frequency spikes at 1 Hz that are understood to originate in the housekeeping electronics. To deal with these, we employ the same methods as with HFI, constructing time-domain sine and cosine templates at 1 Hz and all of its harmonics (and their aliased counterparts) up to the Nyquist frequency. With the LFI sampling frequencies, this amounts to 47, 68, and 116 individual fitting frequencies at 30, 44, and 70 GHz respectively. These template amplitudes were found to be stable (no model exists to predict their amplitude), so we averaged their amplitudes across the mission, then subtracted them from the TOD. While the LFI-DPC processing only corrected for the 1 Hz spikes in the most-affected 44 GHz data, NPIPE fits and corrects for the electronic interference in all of the LFI detectors. Expanding the correction to all of the detectors is particularly important because the low-pass-filtered load signal does not cancel the interference as efficiently as an unfiltered signal. We show co-added line templates for all LFI diodes in Fig. 5.

The LFI-DPC approach to building the 1 Hz spike template differs from that of NPIPE. The DPC template was built by binning diode data into a $1s$-long template according to the phase within the $1s$ period. NPIPE instead considers harmonic templates[9] to model the parasitic oscillation. In the limit of excellent S/N, the two approaches are equivalent, since the DPC template can be binned arbitrarily finely. For shorter fitting periods and lower S/N, the harmonic templates are a more economical decomposition of the parasitic effect, and allow for a more robust fit.

2.3.6. Sky–load differencing and thermal decorrelation

Both LFI and HFI perform decorrelation to reduce time-correlated $(1/f)$ noise fluctuations in the data. The LFI $1/f$ templates are the reference-load timestreams, one for each diode. The HFI thermal templates come from the two dark bolometers.

Traditionally the LFI sky–load differencing has relied on a scaling factor, R:

$$\text{difference} = \text{sky} - R \times \text{load}. \tag{1}$$

For PR3, the LFI DPC calculated the scaling factor for each operational day. These least squares estimates of R balance two opposing effects:

1. low-frequency $1/f$ fluctuations in the load timestreams are highly correlated with the fluctuations in the sky timestream;
2. high-frequency noise in the load timestream is essentially uncorrelated with the sky timestream.

A value of R calculated in this way suppresses the $1/f$ noise, while actually slightly *increasing* the high-frequency noise in the differenced signal.

In NPIPE we have attempted to suppress the injection of uncorrelated high frequency noise into the differenced signal. In

[9] "Harmonic" in this context means sine and cosine templates needed to represent $1s$ periodic signal, accounting for potential aliasing from downsampling the signal. The frequencies of the templates are of the form $nf + kf_\mathrm{r}$, where n is a non-negative integer, f is the parasitic frequency (1 Hz), $k \in \{-1, 0, 1\}$ and f_r is the floating-point modulo of sampling frequency, f_s, and f.

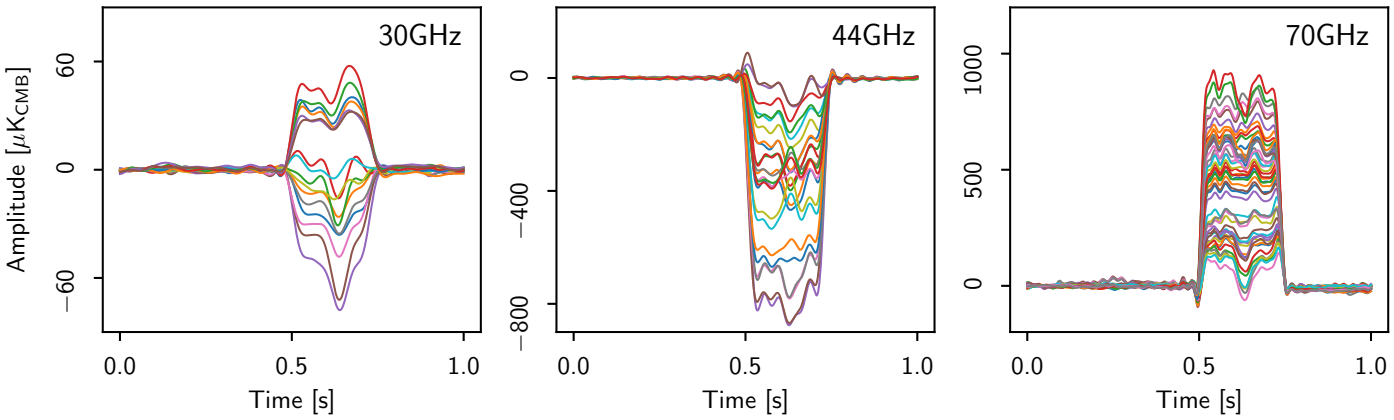


Fig. 5. LFI spike templates built from the harmonic templates fitted to each of the 44 timelines. Each of the coloured curves represents a different diode. The CMB temperature fluctuations have a maximum amplitude of about $250\,\mu K_{CMB}$ and the polarization fluctuations about $2\,\mu K_{CMB}$. The 1 Hz features would be seen at approximately $\ell = 60$.

a somewhat simplified model, the sky and load each have an independent white noise component, σ_{sky} and σ_{load}, and share a fully correlated $1/f$ component with power $P_{corr}(f)$. The corresponding power spectral densities (PSD) are

$$P_{sky}(f) = \sigma_{sky}^2 + R_0^2\, P_{corr}(f), \tag{2}$$

$$P_{load}(f) = \sigma_{load}^2 + P_{corr}(f), \tag{3}$$

$$P_{corr}(f) = \sigma^2 \left[1 + \left(\frac{f}{f_{knee}}\right)^{\alpha}\right], \tag{4}$$

where R_0 is a scaling factor that eliminates the correlated component in the sky–load difference in Eq. (1). Note that we are adopting a notation where σ^2 has PSD units. Using an arbitrary scaling factor, the PSD of the differenced timestream gives

$$P_{diff}(f) = \sigma_{sky}^2 + R^2\sigma_{load}^2 + (R_0 - R)^2 P_{corr}(f). \tag{5}$$

We can thus derive a frequency-dependent scaling factor, $R(f)$, that minimizes the noise power in the differenced signal:

$$R(f) = R_0 \frac{P_{corr}(f)}{P_{corr}(f) + \sigma_{load}^2}. \tag{6}$$

In this simplified model, Eq. (6) defines an optimal low-pass filter that minimizes noise power in the differenced signal. We show examples of the sky and load PSDs and the differenced PSD in Fig. 6, and compare the low-pass-filtered approach to a direct scaling of the load signal. The two examples show that when the uncorrelated instrumental noise is comparable to the correlated $1/f$ noise, low-pass filtering can substantially improve both the low and high frequency noise in the differenced signal.

As a refinement to the model, we consider that a radiometer comprises two diodes, each with sky and load, and the two diodes are also correlated with each other, complicating the situation. We use Eq. (6) to inform us of a parametric representation of a low-pass filter with only three free parameters:

$$F^{lp}(f; R, \sigma, \alpha) = R^2 \frac{f^{\alpha}}{f^{\alpha} + \sigma^2} \tag{7}$$

and we use nonlinear minimization to optimize the low-pass filter for both diodes and their relative weights that define the co-added radiometer signal, $\boldsymbol{d}$:

$$\boldsymbol{d} = w \cdot \left(\boldsymbol{s}_0 - F^{lp}(\boldsymbol{l}_0; R_0, \sigma_0, \alpha_0)\right) + (1 - w) \cdot \left(\boldsymbol{s}_1 - F^{lp}(\boldsymbol{l}_1; R_1, \sigma_1, \alpha_1)\right). \tag{8}$$

Here "$\boldsymbol{s}$" and "$\boldsymbol{l}$" represent the sky and the load, respectively, and indices "0" and "1" identify the two diodes in the radiometer. The fit is performed while marginalizing over a sky-signal estimate.

The best-fit parameters $(w, R_0, \sigma_0, \alpha_0, R_1, \sigma_1, \alpha_1)$ display a fair amount of scatter from pointing period to pointing period. However, their sample medians are very stable, suggesting that the scatter is driven by noise. We decided to run the main LFI preprocessing loop twice. The first iteration is performed to measure the best-fit, low-pass filters (Eq. (7)), as well as to construct 1 Hz spike templates (Sect. 2.3.5) for every pointing period. Then the parameters are averaged across the entire mission and the processing is repeated, keeping the filter and spike parameters fixed.

2.3.7. Jump correction

The *Planck* data also contain "jumps", which are sharp changes in the baseline level that persist (unlike spikes). The `NPIPE` jump-correction module uses a matched filter to detect baseline changes in the TOD. We set the width of the filter to approximately 4 min of data. The width was chosen as a compromise between S/N (wider is better) and $1/f$ noise fluctuations (narrower is better).

We use the global signal estimate for signal removal, avoiding obvious issues that arise from building a signal estimate from data that contain jumps we are trying to detect. Once a jump is detected, the filtered signal provides an estimate of the jump size, which is then corrected for. Since both the exact position and size of the jump are subject to uncertainty, we flag one filter width (4 min) of data around the detected and corrected jump. This approach differs from the implementation used in the HFI-DPC pipeline, which identified pointing periods with jumps by examining the cumulative sum of samples.

The total number of detected jumps varies by detector, but on the average we identify, correct, and flag about a thousand jumps in each detector, once pointing periods with outlier statistics (see Sect. 2.3.10) have been discarded.

2.3.8. Gap filling

As discussed in Sects. 2.3.4 and 2.3.7, the two instruments are subject to glitches and jumps of different origin, and both

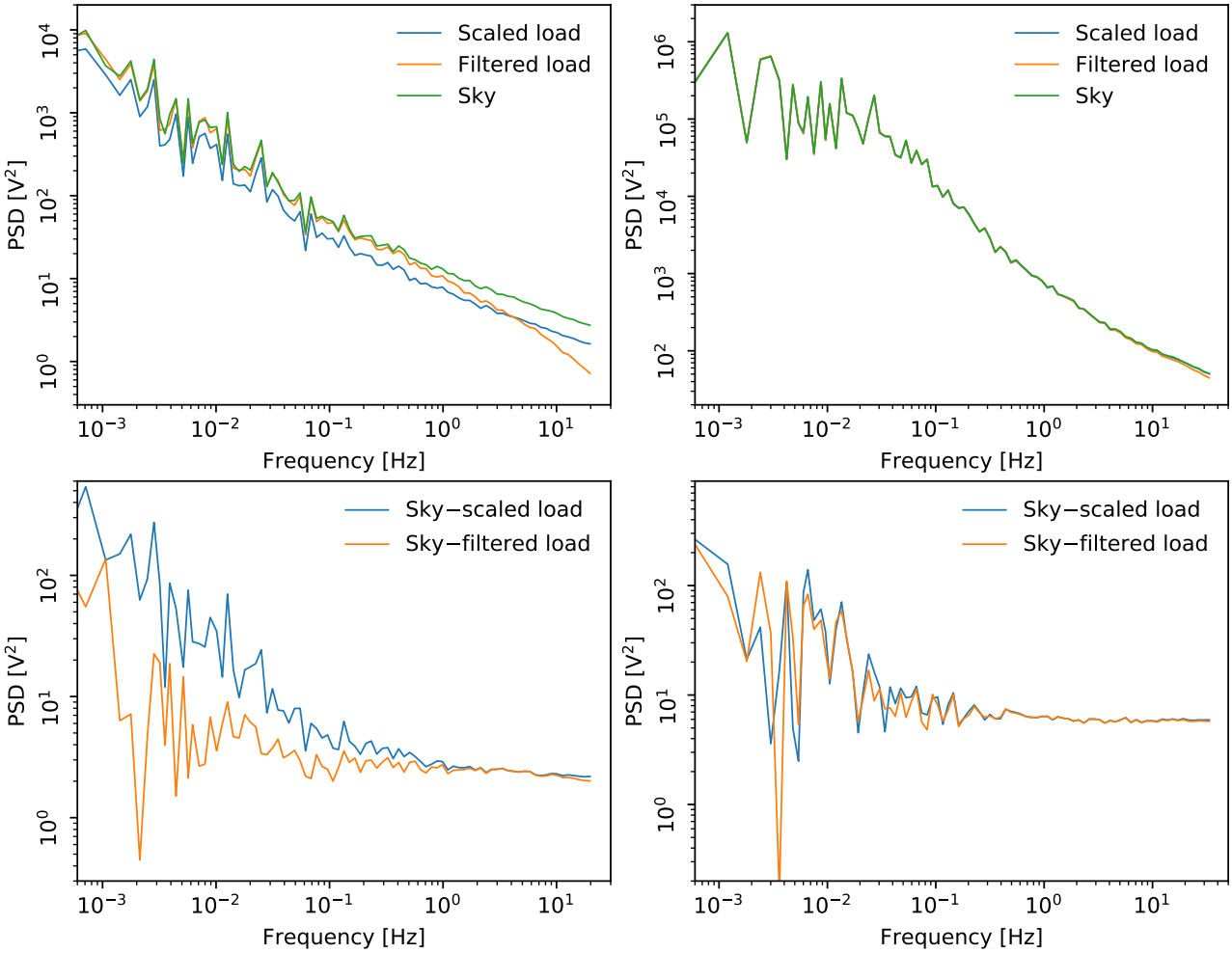


Fig. 6. NPIPE applies a low-pass filter to the LFI load signal before decorrelation. In PR3, a single scaling coefficient was fitted each operational day. We demonstrate the two approaches for one pointing period on two diodes: LFI2500 (*left*), with a relatively low amount of $1/f$ noise; and LFI1800 (*right*), which is dominated by $1/f$ at all frequencies. *Top*: power spectral densities of sky and best-fit load signals. *Bottom*: power spectral densities of sky−load differenced signals. The PSDs were binned into 300 logarithmically-spaced bins for clarity. The scaled load template leaves more noise at both high and low frequencies than a filtered version, where the uncorrelated power is suppressed. When the $1/f$ fluctuations dominate across the spectral domain, the scaling and filtering approaches achieve effectively the same result.

require removal of sections of data. To facilitate the use of Fourier techniques (e.g., bolometer transfer function deconvolution, Sect. 2.3.9), the gaps need to be filled with a constrained realization.

NPIPE deals with gap filling using the same basic approach for all *Planck* channels. For LFI, since it has relatively few gaps, the signal part of the constrained realization comes from the global signal estimate (Sect. 2.3.2), while for HFI it comes from the phase-binned signal estimate. The decision to use the (noisy) phase-binned signal estimate instead of the global signal estimate was made out of an abundance of caution: even small but systematic errors in the signal estimate used to fill 10–20% of the data before applying a filter could potentially lead to detectable bias. The $1/f$ noise fluctuations are matched using a 5th order polynomial fit to the signal-subtracted data, and the gap-filling procedure is completed with a white-noise realization.

2.3.9. Deconvolution of the bolometer time response

The time response of the HFI bolometers gives the relation between the optical signal incident on the bolometers and the output of the readout electronics, characterized by a gain, and a time shift, dependent on the temporal frequency of the incoming signal (Planck Collaboration VII 2016). It is described by a linear complex transfer function in the frequency domain, called the "time transfer function", which must be deconvolved from the data. NPIPE performs this deconvolution at the end of preprocessing. The gap-filled data are Fourier transformed, divided by the measured complex transfer function, and transformed signal back into the real domain. The complex transfer functions are identical to those used in PR3, and described in Planck Collaboration VII (2016).

We discovered that the application of the transfer function to glitch-removed and gap-filled data was picking up a considerable amount of power from the gap-filled samples and mixing it with the science data in the unflagged samples. To suppress this effect, our deconvolution module transforms a delta function signal and measures a window over which more than 10% of the deconvolved power can be attributed to the constrained realization present in the gaps. These samples are flagged in addition to the usual quality flags. The issue is illustrated in Fig. 7. The effect was not corrected for in PR3, contributing a small amount of sky-synchronous noise (especially at 100 GHz), responsible for several percent of the small-scale noise variance in the maps.

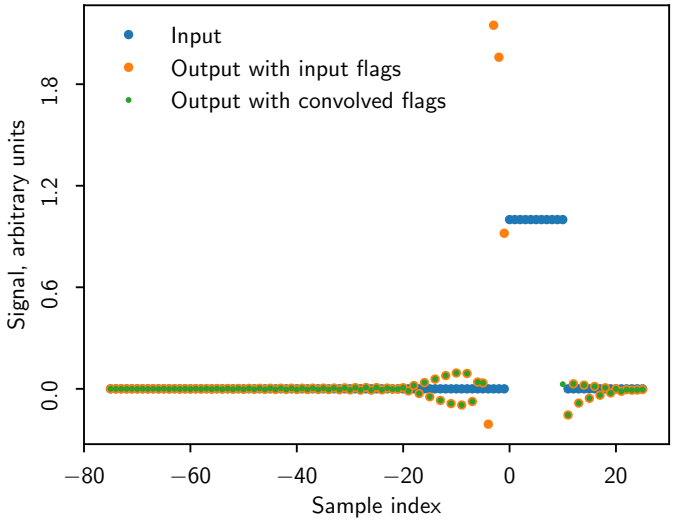


Fig. 7. Demonstration of the leakage of flagged signal onto science data. We created a zero (null) input signal (heavy blue line) with a gap between sample indices 0 and 10. We then filled the gap with an assumed signal = 1.0, then deconvolved using the 100-2b transfer function. Three samples immediately preceding the gap are significantly compromised by the signal in the gap. Such samples are dropped in the NPIPE analysis.

2.3.10. Outlier pointing-period detection

NPIPE accumulates statistics of the data for each pointing period during the preprocessing. Once all of the pointing periods are processed, the metadata are analysed for outlier periods, which are subsequently flagged completely. We used glitch rate, apparent gain, and noise rms to test for outliers by looking for $>5\sigma$ deviations in the statistics after removing a running average. Rings with more than three such deviations were flagged as outliers. Table 1 shows the fraction of outlier pointing periods at each *Planck* frequency. It also shows the total fraction of pointing periods discarded after we also reject unusable data from the sorption-cooler switchover at the end of first year of operations, and data acquired during a 10 day elevated spin-rate campaign.

2.4. Global reprocessing

NPIPE applies a number of templates to correct the data using a generalized destriper, much like the SRoll method described in Planck Collaboration Int. XLVI (2016) for HFI large-scale polarization analysis. Our data model can be cast in the standard destriping form:

$$\boldsymbol{d} = \mathsf{P}\,\boldsymbol{m} + \mathsf{F}\,\boldsymbol{a} + \boldsymbol{n}, \tag{9}$$

where (multi-) detector time-ordered data $\boldsymbol{d}$ are a linear combination of the sought-for sky map $\boldsymbol{m}$, sampled into the time domain with the pointing matrix P, and an arbitrary number of time-domain templates presented as columns of the sparse template matrix F and their amplitudes $\boldsymbol{a}$. When $1/f$ noise offsets are included in F, the noise term, $\boldsymbol{n}$, is approximated as white noise.

If we choose not to impose a prior on the distribution of the template amplitudes, $\boldsymbol{a}$, we can marginalize over the sky map, $\boldsymbol{m}$, and solve for the template amplitudes (Keihänen et al. 2004):

$$\boldsymbol{a} = \left(\mathsf{F}^{\mathsf{T}}\mathsf{N}^{-1}\mathsf{Z}\mathsf{F}\right)^{-1}\mathsf{F}^{\mathsf{T}}\mathsf{N}^{-1}\mathsf{Z}\,\boldsymbol{d}, \tag{10}$$

where

$$\mathsf{Z} = \mathsf{I} - \mathsf{P}\left(\mathsf{P}^{\mathsf{T}}\mathsf{N}^{-1}\mathsf{P}\right)^{-1}\mathsf{P}^{\mathsf{T}}\mathsf{N}^{-1}. \tag{11}$$

Traditionally, F has comprised the baseline offset templates of a step function model of the $1/f$ noise[10] – one column for each baseline period (anywhere between 1 s and 1 h) and every detector. Here, we add additional columns to F for:

- linearized gain fluctuations (discussed in Sect. 2.4.1);
- HFI signal distortion (Sect. 2.4.2);
- orbital dipole (Sect. 2.4.3);
- far-sidelobe pickup (Sect. 2.4.4);
- HFI transfer-function residuals (Sect. 2.4.5);
- 6-component zodiacal light model (Sect. 2.4.6);
- bandpass mismatch (Sect. 2.4.7); and
- foreground polarization (Sect. 2.4.13).

Notice that there is no separate template for the Solar dipole; we consider it as an integral part of the sky, $\boldsymbol{m}$. For this reason, NPIPE calibration is not impacted by uncertainty in the Solar dipole estimate.

Polarization plays a key part in the degeneracy between the template amplitudes, $\boldsymbol{a}$, and the sky map $\boldsymbol{m}$ in Eq. (9). Especially at the CMB frequencies, where the foreground polarization is fainter, *Planck* scanning strategy allows for apparent gain fluctuations that are compensated with scan-synchronous polarization residuals. LFI calibration in the 2015 and 2018 releases addressed this degeneracy by fixing the polarized part of $\boldsymbol{m}$ to a polarized sky model derived using Commander (Eriksen et al. 2004). No such prior was used in the HFI 2015 or 2018 processing. We adopt the LFI approach here and also use a polarization prior, but rather than requiring repeated iterations between NPIPE and Commander, we sample smoothed NPIPE 30, 217, and 353 GHz maps as polarization templates in F. When fitting for these polarization templates, the sky map $\boldsymbol{m}$ is unpolarized, effectively preventing the generation of spurious gain fluctuations that would be compensated by polarization errors. This approach leads to significant improvement in the large-scale polarization systematics in the HFI maps, particularly at 100 GHz. Once time-dependent gain fluctuations and other time-varying signals are corrected for, the last reprocessing iteration freezes the gains and then solves for bandpass mismatch over a polarized sky, without fitting for polarization templates. This final fitting step also fits over a larger fraction of the sky than the previous iterations (Sect. 2.4.12).

NPIPE reprocessing solves for the template amplitudes, $\boldsymbol{a}$, in Eq. (9) and produces a template-cleaned TOD: $\boldsymbol{d}_{\rm clean} = \boldsymbol{d} - \mathsf{F}\,\boldsymbol{a}$. Solving for the amplitudes is made nonlinear by the fact that the gain fluctuation template in F is derived from the sky map, $\boldsymbol{m}$. Notionally, one could explicitly write out $\boldsymbol{m}$ in F in Eq. (9) and solve for the amplitudes using nonlinear minimization techniques. In practice, this is unfeasible owing to the dimensionality of the problem. Instead, we iterate between solving for the amplitudes, $\boldsymbol{a}$, cleaning the TOD and updating $\boldsymbol{m}$ and F. The iterations also enable speeding up the template generation and fitting by compressing the data onto phase-binned HEALPix rings[11],

[10] We use the term $1/f$-noise for instrument noise that has a power spectral density (PSD) that can be approximated with a power law: $P(f) \propto f^{\alpha}$, where f is the frequency and α is the slope of the spectrum. The approximation is valid up to a so-called knee frequency, where f^{α} transitions into a high frequency noise plateau and other features. See Fig. 27 for examples.

[11] Each pointing period is binned onto a subset of HEALPix pixels at the destriping resolution. We must exclude the repointing manoeuvres in doing so, since they break the repetitive scanning pattern we are leveraging.

Table 1. Discarded and quality-flagged data.

Frequency [GHz]	Outlier rings [a] [%]	Discarded rings [b] [%]	Discarded data [c] [%]	Relative total samples [d]	Relative masked samples [e]
30	0.71	1.88	2.62	1.119	1.095
44	0.57	1.76	2.19	1.120	1.098
70	0.71	1.89	2.84	1.114	1.091
100	0.97	2.17	20.00	1.167	1.163
143	1.16	2.36	21.53	1.167	1.160
217	0.95	2.16	20.98	1.221	1.209
353	0.90	2.08	21.78	1.179	1.174
545	0.89	2.06	15.93	1.068	1.066
857	0.93	2.11	17.76	1.019	1.015

Notes. [a]Rings with anomalous glitch rate, noise level, gain, or numerous jumps. A percentage of rings is not necessarily the same as a percentage of integration time, the ring length varies seasonally by almost a factor of 2. [b]Total fraction of rings discarded because of outlier statistics and spacecraft events such as the sorption-cooler switchover, elevated spin-rate campaign and LFI reboot on day 1085 since launch. [c]Total fraction of samples flagged because of pointing, jumps, cosmic-ray glitches, outlier pointing periods and spacecraft events. [d]Ratio of total integration time in the full-frequency maps, in the sense (NPIPE /2018). [e]Ratio of total integration time in the full-frequency maps, in the sense (NPIPE /2018), over the 50% of the sky with the least foregrounds. The distinction matters for LFI, because the ninth survey – included only in NPIPE – lies almost entirely in the Galactic plane.

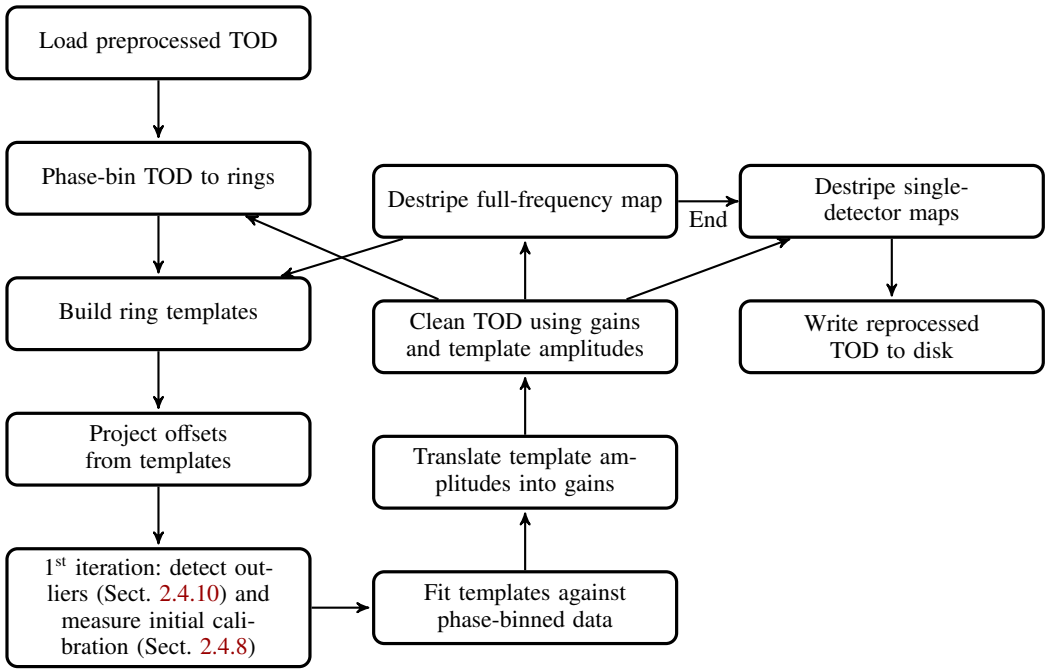


Fig. 8. NPIPE reprocessing flow chart. The *first two columns* contain the main iteration loop where the most recent frequency map is sampled into a new calibration template and the templates are iteratively fitted and subtracted from the TOD. Reprocessing finishes with the *third column* where single-detector data are de-polarized using the full-frequency map.

while retaining the fully sampled data for cleaning and mapmaking. NPIPE reprocessing iterations follow these steps:

1. compress TOD onto HEALPix rings;
2. solve for maximum likelihood template amplitudes $\boldsymbol{a}$ in Eq. (9) *using compressed data*;
3. clean TOD using the amplitudes in $\boldsymbol{a}$; and
4. update the gain template by running Madam on the cleaned TOD.

The recovered template amplitudes rapidly converge to zero as the cleaned data become consistent with the frequency estimate of the sky, $\boldsymbol{m}$. In practice we only need three iterations for the residual errors to become negligible.

Figure 8 shows the flow chart of NPIPE reprocessing. We will now discuss how we build each of the time-domain templates in F.

2.4.1. Gain fluctuations

Calibrating the detector data is a fundamentally nonlinear problem. We linearize the calibration problem by fitting a gain fluctuation template that measures deviations from the average gain for a given frequency. The amplitude of the template, δg, provides us with a gain correction that we may apply to the detector in question:

$$\boldsymbol{d}' = \frac{1}{1+\delta g}\boldsymbol{d}. \tag{12}$$

Our gain template is the most recent iteration map downgraded to low resolution to improve S/N. The sky map already includes the large Solar dipole. Once the template is sampled, we add to it all of the time-domain templates that we have corrected for

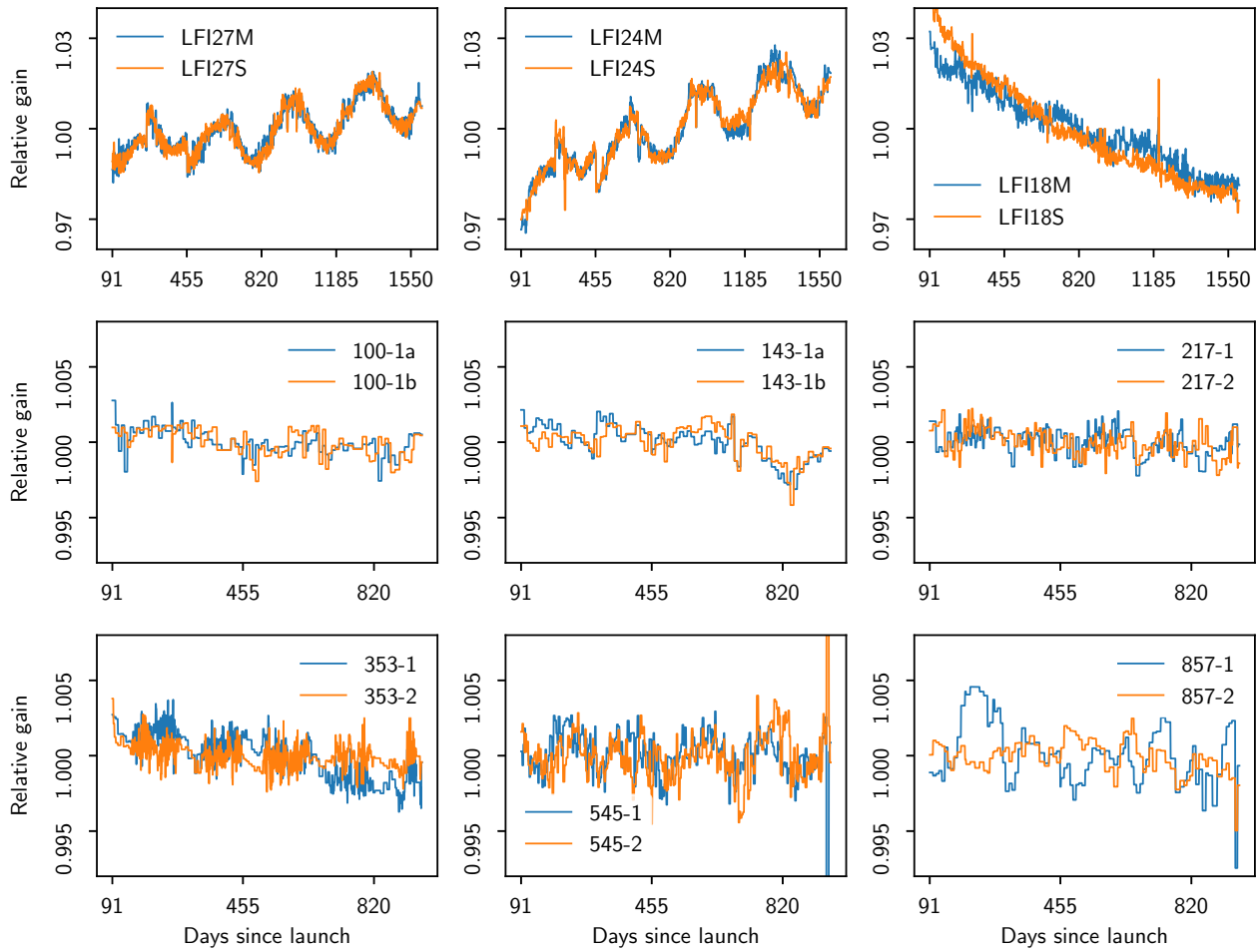


Fig. 9. Sampling of measured gain fluctuations at all *Planck* frequencies. The *top row* shows 30, 44, and 70 GHz, *from left to right*, for representative detectors, with "M" and "S" referring to "main" and "side" (see Planck Collaboration III 2020). The *second and third rows* show HFI frequencies, indicated in the detector names. The vertical lines indicate complete observing years. Seasonal effects due to Solar distance are apparent in the 30 and 44 GHz gains, as well as discontinuities from switching to the redundant 20 K sorption-cooler system on day 455 and changing the transponder setting on day 270.

in the cleaned TOD. That means adding the orbital dipole, far sidelobes, bandpass mismatch, HFI transfer-function residuals, and zodiacal light. It is true that a gain template sampled from the same sky map we are trying to estimate will have noise that is correlated with the detector TOD, but the effect is negligible because of the overwhelming number of samples accumulated into any sky pixel.

The gain template is split into disjoint chunks of time to trace changes in the gain. Each gain step is long enough to amount to roughly equal S/N, except for the few steps that are broken by known discontinuities or run into a preset maximum step length of 300 pointing periods (about 10 days). We have optimized the S/N threshold for each frequency to allow adequate tracking of gain fluctuations without excessively contributing to the map noise through noise-driven gain errors. We show the measured gain fluctuations for a subset of *Planck* detectors in Fig. 9. Unlike the LFI-DPC gain solution, the `NPIPE` gains are not filtered at all. Details of the filtering required to render LFI-DPC ring gains usable can be found in Planck Collaboration V (2016).

2.4.2. Residual HFI ADC nonlinearity

In previous *Planck* releases, residual HFI ADCNL was approximately addressed by correcting for apparent gain fluctuations (Planck Collaboration VIII 2016). Observed changes in the effective detector gain may be combined with the approximate level of the input signal to create an effective model of the residual ADC nonlinearity. This approach was shown to be effective in reducing large-scale polarization systematics in Planck Collaboration III (2020). We refer to this model of the ADCNL causing multiplicative, gain-like fluctuations in the TOD as the "linear gain model" of ADCNL.

Despite the success of the linear gain model in suppressing the nonlinearity errors, ADCNL was still the dominant source of large-scale polarization uncertainty in Planck Collaboration III (2020), prompting the use of inter-frequency cross-spectral methods for estimating the re-ionization optical depth from the maps (Planck Collaboration Int. XLVI 2016).

`NPIPE` makes two important updates to the handling of residual ADCNL. Firstly, we do not assume the apparent gain fluctuations to be exclusively sourced by the ADCNL. In this way, our gain steps are free to address both apparent and real gain fluctuations. Secondly, we add another calibration template that has all values above the median signal nulled. The function of this second "distortion" template is to enable us to adapt to signal distortions that arise from a different effective gain being imposed at opposite ends of the signal range. This is a natural extension of the linear gain model when the signal covers a wide dynamic range and the extremes of the signal probe different registers of

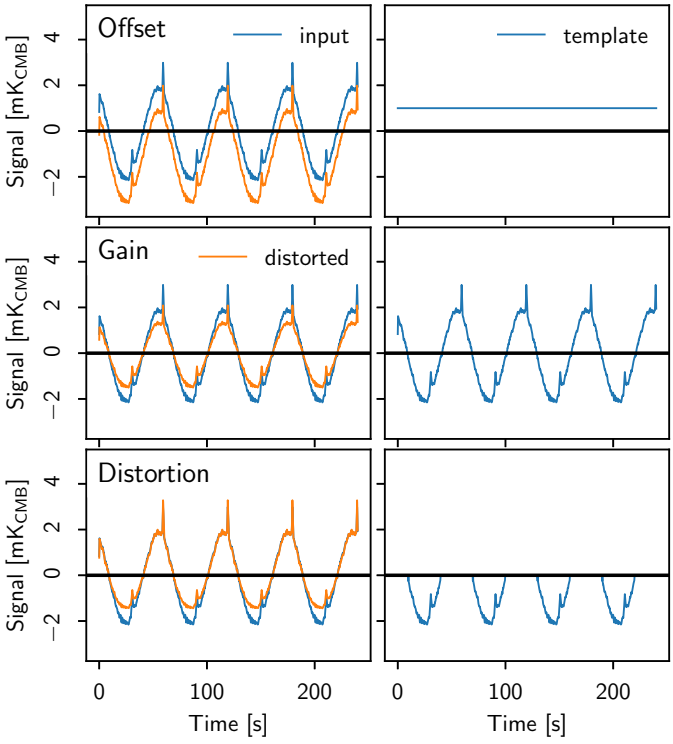


Fig. 10. Three degrees of signal distortion due to HFI ADC nonlinearity (*left panels*) and the NPIPE templates used to fit for them (*right panels*). The blue curves give an interval of data from a specific bolometer. The main variations seen come from the detector scanning across the dipole and the Galactic plane, including the ADCNL effects. The orange curves show the steps taken to fit the distortions. *Top*: (harmless) signal offset is captured by destriping baselines (i.e., moving the orange curve upwards). *Middle*: linear gain fluctuations are addressed by calibration. *Bottom*: linear gain varies as a function of signal level and requires a separate gain template for the opposite ends of the signal range. This correction is unique to NPIPE.

Table 2. HFI transfer-function bins.

Band	First harmonic [a]	$f_{\rm min}$ [Hz]
1	1	0.025
2	2	0.042
3	4	0.075
4	6	0.108
5	9	0.158
6	13	0.225
7	18	0.308
8	25	0.425
9	34	0.575
10	46	0.775
11	61	1.025
12	81	1.358
13	106	1.775
14	139	2.325
15	181	3.025
16	235	3.925

Notes. [a]The satellite spin rate, 16.7 mHz.

the ADC. We demonstrate the first three leading orders of signal error due to ADCNL in Fig. 10. Any fluctuation in the TOD that can be fitted with the gain and distortion templates, could also be fitted with two distortion templates having the opposite signal halves nulled. Arranging the templates into gain- and distortion-only parts simplifies the interpretation. The template corrections are applied to the TOD by multiplying the median-subtracted signal and half-signal (samples that have values below the median) with the appropriate factors. We show examples of the fitted distortion template amplitudes in Fig. 11.

It is worth emphasizing that the NPIPE model for the ADCNL is not a precise model that derives the correction from known inputs. Such a model is, in fact, impossible to construct, due to the down-sampling that occurred onboard the spacecraft. Instead, we simply fit piece-wise stationary time-domain templates to the mismatch between detectors, in an effort to reduce temperature-to-polarization leakage that would otherwise dominate the large-scale polarization. The shape of the templates is well motivated by our understanding of the ADCNL. The language of linear gain drifts is useful in describing the signal effects, but there is no way of disentangling actual changes in bolometer gain (expected to be of the order of 10^{-5}) from apparent drifts sourced by ADCNL.

2.4.3. Orbital dipole

The motion of *Planck* around the Sun causes a seasonal Doppler modulation of the CMB monopole. The spacecraft position is known to an accuracy better than 1 km and the velocity better than 1 cm s^{-1} (Godard et al. 2009). The CMB monopole is 2.72548 K ± 0.57 mK (Fixsen 2009). Because both the spacecraft motion and the CMB monopole are extremely well known, this ∼300 μK$_{\rm CMB}$ dipole[12] signature is our best absolute calibrator. We fit for an orbital dipole template in each detector TOD and find a per-frequency calibration coefficient that brings the average orbital dipole template to the expected level. We do not require that all detectors observe an identical dipole, since far sidelobe effects alter the perceived amplitude at the individual horn level. We do, however, subtract the same orbital dipole from both detectors in a horn. Far sidelobe effects are approximately corrected by convolving the orbital dipole template up to the second-order quadrupole (sometimes called the "kinematic quadrupole") with a GRASP[13] estimate of the far sidelobes.

The orbital dipole fits provide for absolute calibration for all frequency channels except for 857 GHz, where NPIPE uses the same planetary calibration as in PR3.

2.4.4. Far-sidelobe pickup

We calculate timeline estimates of the far-sidelobe (FSL) pickup by convolving PR2 full-frequency maps (with a Solar dipole added) with a model of the beam. We considered updating the convolved timelines from PR3 or NPIPE temperature maps, but the differences would have been negligible.

To perform this convolution, we used available GRASP estimates (which are monochromatic) of the 4π detector beams. Full spectral treatment of the frequency-dependent FSL would have amounted to a percent-level correction to a 10^{-3} level systematic. The estimates are only available for the 30–353 GHz channels, so we used 353 GHz beams as proxies for 545 and 857 GHz channels. Except at the submillimetre frequencies, the FSL template is too faint and degenerate to include in the fitting. We simply subtract it and correct the amplitude of the subtracted

[12] The range of the orbital dipole signature depends seasonally on the phase of the 7.5° spin axis precession and the position of each detector on the focal plane.

[13] The GRASP software was developed by TICRA (Copenhagen, DK) for analysing general reflector antennas (https://www.ticra.com/).

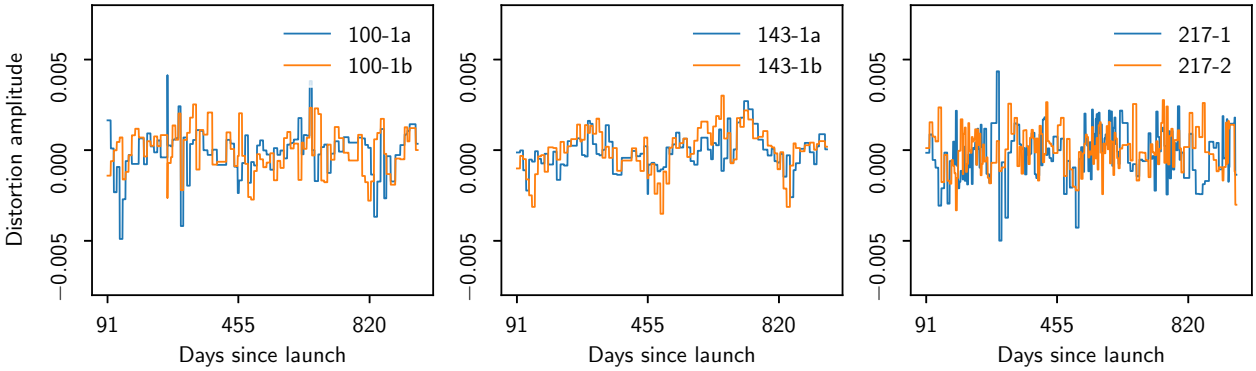


Fig. 11. Sampling of measured signal distortions (cf. Sect. 2.4.2) at 100, 143, and 217 GHz. These are the only frequencies where the distortion template is fitted. The distortion templates are fit at the same time steps as the gain templates in Fig. 9.

Table 3. Zodiacal template emissivities.

Frequency [GHz]	λ [μm]	Cloud	Band 1	Band 2	Band 3	Blob	Ring
857	350	0.285	0.53	0.274	0.833	0.317	0.782
545	550	0.182	0.973	0.32	1.37	0.214	0.438
353	849	0.105	0.954	0.207	1.00	0.0883	0.449
217	1382	0.0391	0.823	0.0999	0.685	0.0211	0.332
143	2096	0.0106	0.777	0.0693	0.494	0.0207	0.243
100	2998	0.0124	0.68	0.0779	0.523	−0.0327	−0.0282

Notes. The statistical uncertainties for these amplitudes are found to be negligible. See Planck Collaboration XIV (2014) for an explanation of the components. Note that `NPIPE` only fits and corrects for the time-varying part of the Zodiacal emission model.

template as our estimate of the gain fluctuations improves. For 545 and 857 GHz, the FSL template is successfully fitted with the individual SWBs, yielding roughly consistent best-fit amplitudes between the detectors.

2.4.5. HFI transfer-function residuals

The time response of the bolometers to sky signals is complicated. HFI uses planet observations and stacked cosmic-ray glitches to infer the shape of the bolometer transfer function and deconvolve it. These calibrators are excellent for probing the intermediate- and high-frequency components of the transfer function, but leave an uncertainty about the relative values of the low- and high-frequency components. The most striking effect is a relative change in phase between the Solar dipole and signals on much smaller angular scales. We construct templates that correspond to the real and imaginary parts of a binned residual transfer function, fit the templates, construct the transfer function from the fit amplitudes, and deconvolve it from the signal. The real part of the transfer function corresponds to a small, scale-dependent change in calibration, and can be fit reliably only at 353 GHz and above. The imaginary part introduces a phase shift between different scales, and is fit at all HFI frequencies.

We quantize the transfer-function bins to contain one or more harmonics of the 16.7 mHz spin frequency, and exponentially increase the bin width according to

$$w_b = \mathrm{floor}(\mathrm{e}^{b/4})/60\,\mathrm{s}, \tag{13}$$

where b is the band index, w_b is the width of the band in Hz, the first band begins at 0.5/60 s, and the "floor" function returns the largest integer value that is smaller than its argument. The last band is extended all the way to the Nyquist frequency. To avoid a degeneracy with overall calibration, we omit the first band of the real transfer-function templates. We find that we can reliably fit up to 16 bands at 353 GHz and above.

In the HFI 2018 processing, the empirical transfer function was fitted in only four bins. A pattern of "zebra stripes" was identified in the 353 GHz odd–even survey difference maps and shown to be caused by the inability of the coarse-grained transfer-function model to track the actual residual. The odd–even survey differences for both methods are described further in Sect. 6.3.1.

2.4.6. Zodiacal emission

Interplanetary dust in the Solar System forms a cloud through which we observe the CMB. Thermal emission from zodiacal dust is visible at high *Planck* frequencies (Planck Collaboration XIV 2014), and has characteristic seasonal variations as the satellite moves with respect to the cloud. The seasonal variations are of concern for our template fitting, since they might bias the gains and noise offsets. We use the six geometric components described by Kelsall et al. (1998) to construct six independent templates that capture the seasonal variations. To preserve the mean zodiacal emission in the signal, our templates are the differences of expected emission between the current spacecraft position and another position 180° along Earth's orbit. The opposite position is found by inverting the heliocentric position vector. Furthermore, to boost the S/N we use the same template amplitude for all detectors in a frequency channel.

Our recovered zodiacal template amplitudes are roughly in agreement with the 2018 results (Planck Collaboration III 2020), but offer no new information, owing to the removal of the common mode and the emphasis on cleaning the data of seasonal effects. We defer the actual zodiacal emission removal to the component-separation process, where multi-frequency

information offers better means of cleaning the frequencies where the emission is faint. The recovered template emissivities are listed in Table 3.

One may assess the success of the zodiacal emission removal by examining the difference between maps made from odd and even surveys. Comparison of NPIPE and PR3 odd–even survey differences (see Sect. 6.3.1) suggests that NPIPE processing achieves at least the same level of zodiacal residuals as the 2018 processing; however, it is hard to quantify the performance because of other features in the same survey-difference maps.

2.4.7. Bandpass mismatch

Planck detectors within a frequency band have differing bandpasses, causing them to see foreground components at different intensities. In theory, one can use a frequency-dependent model of the foreground emission, coupled with a set of measured detector bandpasses, to predict the amount of bandpass mismatch and derive correction maps. In practice, uncertainties in both the measured bandpasses and the sky model require fitting the corrections directly against the data. For all continuum-emission foregrounds (synchrotron, free-free, anomalous microwave emission, and dust), we use a high-resolution Commander sky model to predict the frequency derivative of the foreground emission, pixel-by-pixel, at the centre of the frequency band. The fitting amplitude directly gives us an estimate of the relative difference in centre frequency between the detectors. We also consider narrow-line emission from CO regions, which are fitted with a separate template. The bandpass templates are HEALPix maps that we bilinearly interpolate to the detector sample positions.

For added flexibility, we choose to include some of the foreground components as separate templates. This compromises any direct physical interpretation of the amplitude of the bandpass-mismatch template (i.e., difference in centre frequency), but leads to better agreement between full-frequency and single-detector maps. For the 30–70 GHz channels, these extra templates are the frequency derivatives of the Commander anomalous microwave emission (AME) and dust frequency components, chosen for their similarity to maps of the full-frequency versus single-detector mismatch. For 70 GHz we also fit for an HCN ($J = 1 \rightarrow 0$, 88.63 GHz) emission-line template. For 100–857 GHz, we use the frequency derivative of the free-free component instead of AME or dust.

2.4.8. Initial calibration

In order to speed up convergence, we perform a linear regression of the raw TOD against the 20% of the sky nearest to the dipole extrema. The target is the 2015 *Planck* dipole, which we use as a crude initial calibration. The fit provides an initial, fixed, calibration coefficient for each detector; these are overridden by the orbital dipole fit and relative gain corrections during the first reprocessing iteration.

2.4.9. Submillimetre processing

Planck 545 and 857 GHz frequency channels have unique processing requirements. Foregrounds dominate across the sky, wide bandpasses cover a number of potential emission lines, and the detectors, although nominally polarization-independent, are in fact weakly sensitive to polarization, with polarization efficiencies measured to be around 6%. In contrast, the HFI PSBs have values around 90%. For 545 GHz, we marginalize over the NPIPE 30, 217, and 353 GHz polarization maps, smoothed to 60′, to project out detector mismatch due to polarization. For 857 GHz, we found that this approach does not improve detector agreement, and chose to overlook the small polarization sensitivity in reprocessing.

2.4.10. Outlier detection

NPIPE reprocessing searches for outlier pointing periods in the data at two stages. First, a submatrix of F is fitted against the TOD, one detector and one pointing period at a time. Pointing periods showing anomalous fitting coefficients are flagged. Second, the destriper estimates the initial residual for every detector and every pointing period by binning, resampling, and subtracting the input TOD. Again, anomalous periods are flagged. These steps result in fewer than ten pointing periods permanently flagged for each detector. The total numbers of additional discarded pointing periods (across all detectors) at each frequency is 23, 25, 65, 9, 16, 18, 14, 0, and 11 for 30 GHz through 857 GHz, respectively.

2.4.11. Horn symmetrization

Each pair of polarization-orthogonal detectors that shares a horn forms a single polarization-sensitive unit, with nearly identical optical properties. We enforce this by:

- applying a joint set of quality flags to both detectors and dropping the same outlier data for each;
- using a common noise weight for both detectors;
- using the same pointing;
- using the same FSL template;
- using the same zodiacal-light templates and template amplitudes.

The purpose of the symmetrization is to limit the temperature-to-polarization ("T-to-P") leakage. Since the detectors in a horn share the same optical feed, the difference in samples collected simultaneously from the two is free from most of the optical mismatch that would otherwise cause leakage. In fact, our symmetrization scheme has the same effect as differencing the TOD from the two detectors, then using the resulting difference to solve for sky polarization.

One potential source of T-to-P leakage is the sub-pixel structure that includes sharp foreground features as well as the dipole gradient. Samples drawn from various corners of the pixel may indicate different signal levels, and if those observations are associated with different observing angles, the difference can be interpreted as polarization. To the extent that the polarized pairs of detectors are optically aligned, this source of leakage is blocked by our symmetrization procedures.

We have verified that the symmetrization reduces T-to-P leakage in the Galactic plane and in the vicinity of compact sources, where beam effects are pronounced. While the effect is small, it was nevertheless considered strong enough to merit the increase in noise from symmetrizing the flags and using slightly less optimal noise weights for the detectors.

2.4.12. Last iteration and processing masks

Each iteration of the reprocessing uses a processing mask to avoid issues arising from sub-pixel structure, variable unresolved ("point") sources, and small incompatibilities in our bandpass-mismatch templates. The same mask is not necessarily optimal

for all of the templates. In particular, the gains require only the cleanest and most reliable parts of the sky to be considered, while the bandpass corrections can greatly benefit from including as much as possible of the intense Galactic emission. To best serve both applications, we mask more of the sky while we iterate over the calibration, then freeze the gain solution and run the last iteration with more of the Galaxy exposed.

Point sources, especially variable radio sources, and other steep signal gradients on the sky may cause an error in the fitted template amplitudes. We use processing masks to avoid regions of the sky where the signal model is inaccurate.

NPIPE reprocessing uses apodized processing masks. This means that it is possible to down-weight regions of the sky without discarding them completely when solving for the template amplitudes. This approach is particularly useful for measuring the gain fluctuations both at the dipole extrema and over intense Galactic emission. The use of apodized processing masks is a new development in TOD processing, but has been used extensively in map domain analysis.

We build our apodized masks from an earlier iteration of NPIPE temperature maps. First, we subtract PR2 dipole and CMB map from each frequency map, and then find the scaling coefficients that suppress the foreground amplitude in each pixel below a fixed threshold. These thresholds are not rigorously optimized, but rather are chosen to visually suppress regions of the sky where single-detector maps show notable disagreement with the full-frequency map. Further adjustment is performed when the gain solutions show sharp and recurring seasonal trends. Finally, point sources are added to the masks from the second *Planck* compact source catalogue (Planck Collaboration XXVI 2016).

We present our processing masks in Fig. 12. The final short-baseline destriping uses binary masks, but we are able to use a large sky fraction due to the time-domain bandpass-mismatch corrections.

2.4.13. Polarization estimates

The full polarized solution of the template amplitudes is degenerate at CMB frequencies and subject to large noise uncertainty. Without external constraints, the template fitting converges to a self-consistent but sub-optimal solution, with large-scale polarization power leaking in from the dipole (see Sect. 6.3.1). These issues can be remedied by performing the template fits on a temperature-only sky. Polarization cannot be overlooked without biasing the solution, but it is possible to marginalize over the sky polarization by adding additional templates that model the foreground polarization to the template matrix. To this end, the 217 GHz reprocessing fits for 30 and 353 GHz polarization templates smoothed to 1°, while 44–143 GHz reprocessing fits for 30, 217, and 353 GHz. For each template, we only allow for a single template amplitude across all detectors. This improves the S/N at the cost of ignoring the polarized bandpass mismatch. We do not explicitly fit for bandpass mismatch in polarization, but the sky model does predict a percent level correction based on the central frequencies fitted with temperature which is included when applying the bandpass correction.

To construct the time-domain templates, we smooth each foreground frequency map with a 1°-FWHM Gaussian beam and sample according to the detector polarization parameters and orientation.

2.4.14. Degeneracies

Including sampled sky maps as templates in Eq. (9) creates a degeneracy. For instance, the same bandpass-mismatch template is sampled for every detector, and we can shift power between the template amplitudes in $\boldsymbol{a}$ and the sky map, $\boldsymbol{m}$. Such degeneracies prevents direct solution of Eq. (10); instead, we perform a conjugate gradient (CG) iteration for $\boldsymbol{a}$ that minimizes the residual in

$$\left(\mathsf{F}^{\mathsf{T}}\mathsf{N}^{-1}\mathsf{Z}\mathsf{F}\right)\boldsymbol{a} = \mathsf{F}^{\mathsf{T}}\mathsf{N}^{-1}\mathsf{Z}\,\boldsymbol{d}. \tag{14}$$

Once the CG iteration converges, we adjust the template amplitudes, $\boldsymbol{a}$, as follows:

– the bandpass-mismatch correction for each frequency must average to zero;
– relative gain corrections for each frequency must average to zero; and
– the average of all noise offsets must equal zero.

These adjusted template amplitudes are then used to correct the TOD. An alternative (but equivalent) approach would have been to modify the destriping equation itself to include these priors.

The matter of degenerate modes is a common topic in destriping. They arise due to the projection matrix, Z, which removes the sky-synchronous part of the signal. Even in the case of traditional destriping that only fits for noise offsets, an overall offset of the solved map is set by the initial guess rather than being solved from the data. Beyond that, any sky-synchronous mode in the templates is invisible to the destriper.

The unadjusted average of all of the gain fluctuations tends to be large (of order unity rather than zero) and matched by an opposite average value in the orbital dipole template. This simply reflects a degeneracy between the two templates, because the gain template includes the orbital dipole. To deal with this, we subtract the frequency average of the gain fluctuations and adjust the orbital dipole amplitudes accordingly to keep the level of total orbital dipole unchanged.

There is a degeneracy between the templates used to model foreground polarization and the templates used to capture bandpass mismatch. This degeneracy requires no active mitigation due to the way bandpass-mismatch correction is only applied in the final template fitting step that does not involve polarization templates. For more discussion, see Appendix H.

2.4.15. Polarization parameters

During the NPIPE development a number of diagnostics were produced to assess the performance of the template corrections. One very useful diagnostic is a set of polarization-corrected, single-detector maps reflecting the final template amplitudes. Any mismatch between the full-frequency temperature map and these single detector maps can be considered a residual and an indication of a deficiency in the data model or the fitting procedure. Once other sources of residuals have been addressed, the remainder, especially for some of the 353 GHz bolometers, exhibits a residual signal that closely resembles maps of polarized sky emission. We suggest that these residuals arise from small errors in the polarization efficiencies or angles that we adopted from the DPCs.

We derive two time-domain polarization templates by scanning the full-frequency polarization map, one with the usual Q and U weights, and another using the polarization angle derivative of the Q and U weights. We then bin these two

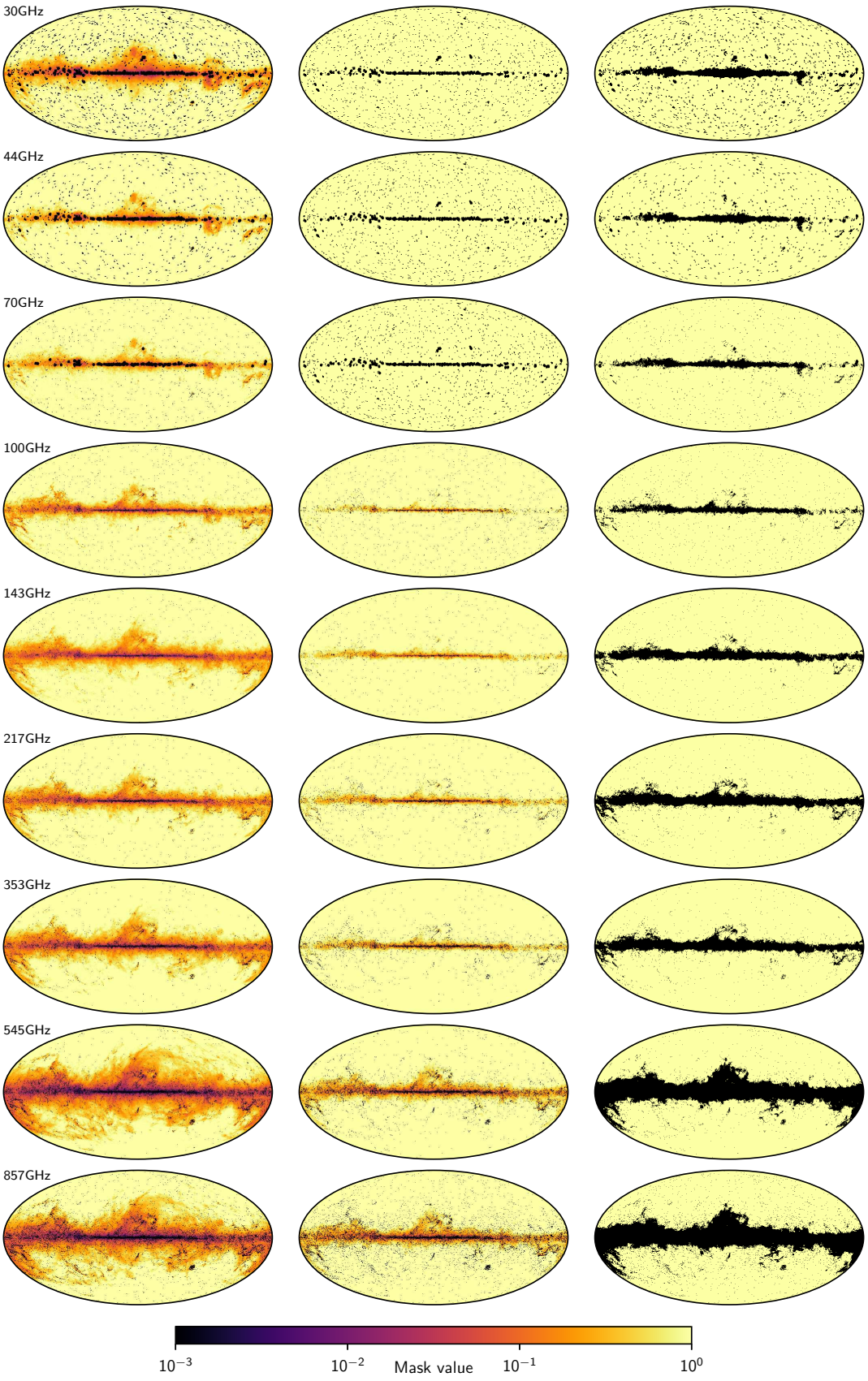


Fig. 12. Sharp and apodized processing masks used in NPIPE processing. The three columns are for the calibration (*left*), bandpass correction (*middle*), and destriping (*right*) masks. The destriping masks are explicitly binary because Madam (which we use for this step) does not support floating-point masks.

Table 4. NPIPE polarization efficiencies and angles.

Bolometer	Pol. efficiency [%]	Δ [%]	Pol. angle [deg]	Δ [deg]
100-1a	94.35 ± 1.03	−0.36	23.32 ± 0.28	+0.22
100-1b	96.98 ± 0.87	+2.67	111.93 ± 0.24	+0.13
100-2a	96.03 ± 0.45	−0.15	44.26 ± 0.18	−0.84
100-2b	89.74 ± 0.45	−0.50	134.57 ± 0.21	+0.27
100-3a	89.56 ± 0.71	−0.53	179.33 ± 0.14	−0.87
100-3b	93.23 ± 1.06	−0.22	90.29 ± 0.20	+0.19
100-4a	93.88 ± 0.63	−1.83	156.96 ± 0.16	+0.46
100-4b	93.02 ± 0.69	+0.75	67.95 ± 0.17	−0.06
143-1a	85.55 ± 0.20	+2.32	45.74 ± 0.09	+0.24
143-1b	86.43 ± 0.34	+1.84	137.66 ± 0.12	−0.04
143-2a	85.35 ± 0.28	−2.09	45.19 ± 0.11	−0.21
143-2b	88.97 ± 0.20	−0.21	135.49 ± 0.09	+0.29
143-3a	86.03 ± 0.45	+2.13	0.08 ± 0.09	+0.38
143-3b	88.69 ± 0.35	−1.25	92.74 ± 0.09	−0.36
143-4a	91.75 ± 0.40	−1.36	0.61 ± 0.08	−0.29
143-4b	93.72 ± 0.39	+0.88	89.61 ± 0.09	+0.31
143-5	7.52 ± 0.30	+1.25	68.95 ± 1.03	+3.25
143-6	3.87 ± 0.38	−0.30	69.71 ± 2.06	−0.89
143-7	3.06 ± 0.32	+1.74	95.24 ± 2.96	−7.56
217-1	2.34 ± 0.20	−1.50	97.11 ± 2.18	−1.29
217-2	3.33 ± 0.20	+1.34	95.87 ± 1.57	+13.37
217-3	1.23 ± 0.15	−2.72	163.02 ± 2.27	−7.88
217-4	3.78 ± 0.14	−0.61	121.20 ± 1.56	+1.20
217-5a	96.14 ± 0.12	+1.13	47.08 ± 0.04	+0.08
217-5b	97.35 ± 0.15	+2.15	135.88 ± 0.08	−0.22
217-6a	90.12 ± 0.09	−4.82	46.01 ± 0.06	−0.09
217-6b	93.91 ± 0.09	−1.48	135.97 ± 0.04	+0.07
217-7a	94.09 ± 0.15	+0.04	179.43 ± 0.02	−0.17
217-7b	95.15 ± 0.14	+1.48	90.53 ± 0.03	+0.03
217-8a	96.76 ± 0.20	+2.55	1.18 ± 0.04	+0.88
217-8b	95.80 ± 0.20	+1.68	89.90 ± 0.05	−0.80
353-1	0.23 ± 0.13	−2.97	101.74 ± 18.81	−1.36
353-2	7.49 ± 0.05	+2.78	120.68 ± 0.63	+6.08
353-3a	89.93 ± 0.09	+0.95	45.70 ± 0.06	+0.20
353-3b	90.90 ± 0.11	−1.17	133.64 ± 0.07	−0.26
353-4a	87.35 ± 0.13	+0.14	45.67 ± 0.05	−0.03
353-4b	89.85 ± 0.08	−1.74	135.52 ± 0.06	−0.08
353-5a	85.44 ± 0.11	+0.75	177.87 ± 0.02	+0.17
353-5b	90.95 ± 0.11	+3.35	89.43 ± 0.04	−0.17
353-6a	84.99 ± 0.58	−2.55	0.22 ± 0.12	+0.52
353-6b	92.56 ± 0.66	+3.79	88.86 ± 0.09	−0.64
353-7	5.31 ± 0.12	−2.68	126.81 ± 0.70	+5.31
353-8	9.88 ± 0.08	+2.06	139.30 ± 0.35	+6.30

Notes. We report corrections with respect to the ground-measured values used in PR3. In Rosset et al. (2010) the statistical uncertainties for polarization efficiency are estimated to be 0.1–0.6 % for PSBs and SWBs operating between 100 and 353 GHz. Polarization angle systematic uncertainties for the PSBs were estimated by Rosset et al. (2010) to be 0.°9 (not including a 0.°3 common error); uncertainties for the SWBs were estimated to be up to 5.°5. For NPIPE, statistical uncertainties in the PSB polarization angles are small, while those for the SWBs are up to 2.°3. The NPIPE uncertainties include statistical and systematic errors estimated using 60 simulations.

templates using the same intensity weights as with the polarization-corrected, single-detector TOD. Fitting the resulting two polarization templates against the single-detector residual map allows us to measure a correction to the polarization efficiency and polarization angle that we incorporate into a refined version of the instrument model. The template fits are limited by other residuals in the single-detector maps, so their uncertainties are systematic, not statistical, and are difficult to assess without simulations.

We have been able to reduce the single-detector/full-frequency residuals only from 100 to 353 GHz (as shown in Table 4). The LFI residual maps do not show a residual consistent with an error in the polarization parameters, either due to lack of such error or due to higher noise content.

One might argue that the binned full-mission maps are not appropriate for fitting the polarization parameters because they do not capture the fact that the detectors are rotated with respect to the pixels between different surveys. We tried fitting for the polarized templates during destriping (similar to other templates), but found that the change in position angle over most of the sky is insufficient and the template fits are unreliable. This is why the polarization parameter corrections are based on map domain fits.

3. Destriping

If the relevant systematics are sufficiently suppressed, low-frequency noise fluctuations in each detector begin to dominate large-scale uncertainties in the final maps. In an effort to extract maximal information from the data, we have adopted an aggressive approach to destriping, fitting each detector with 167 ms baseline offsets that correspond to 1° steps on the sky. These short baselines are used to capture as much of the low-frequency instrumental noise as is statistically possible, but they also adapt to other large-scale residuals such as ADC nonlinearities and gain fluctuations. The baseline-offset solution is regularized by a prior on the baseline distribution (Keihänen et al. 2005, 2010), which is derived from the detector-noise power spectral densities (PSDs; see Sect. 5.1).

We tested the efficacy of the short baselines by destriping and projecting dark-bolometer data from the two dark bolometers as if they were bolometers 857-1 and 857-3. Figure 13 shows the power spectra of maps destriped with various baseline lengths. Using the dark bolometer data instead of synthetic TOD has the advantage that they include realistic glitch and 4 K line residuals.

In choosing the destriping resolution, we have balanced S/N, which improves with lower destriping resolution, with signal error, which is best when there is no detectable sub-pixel structure. Our destriping resolution is $N_{\rm side} = 512$ for all LFI frequencies, $N_{\rm side} = 1024$ for 100–353 GHz, and $N_{\rm side} = 2048$ for 545 and 857 GHz. These resolutions allow for modest sub-pixel signal gradients outside the destriping mask, but do not lead to temperature-to-polarization leakage due to the horn symmetrization (Sect. 2.4.11).

Solving for short baselines also serves a secondary purpose in allowing us to apply the Madam approach for estimating the residual noise in the pixel-to-pixel covariance in a low-resolution version of the NPIPE data set. In Madam, the noise is broken down into baseline offsets and white noise. When the baselines are too long to capture the $1/f$ noise fluctuations, Madam can only accommodate the extra power in the white noise component of the signal model, leading to mismatch between the noise matrix and the actual noise in the maps.

4. NPIPE data set

We now describe the contents of the NPIPE data release, highlighting differences between NPIPE and other *Planck* public releases that may be relevant to the user. For information regarding the dissemination of the maps and the related software, see Appendix A.

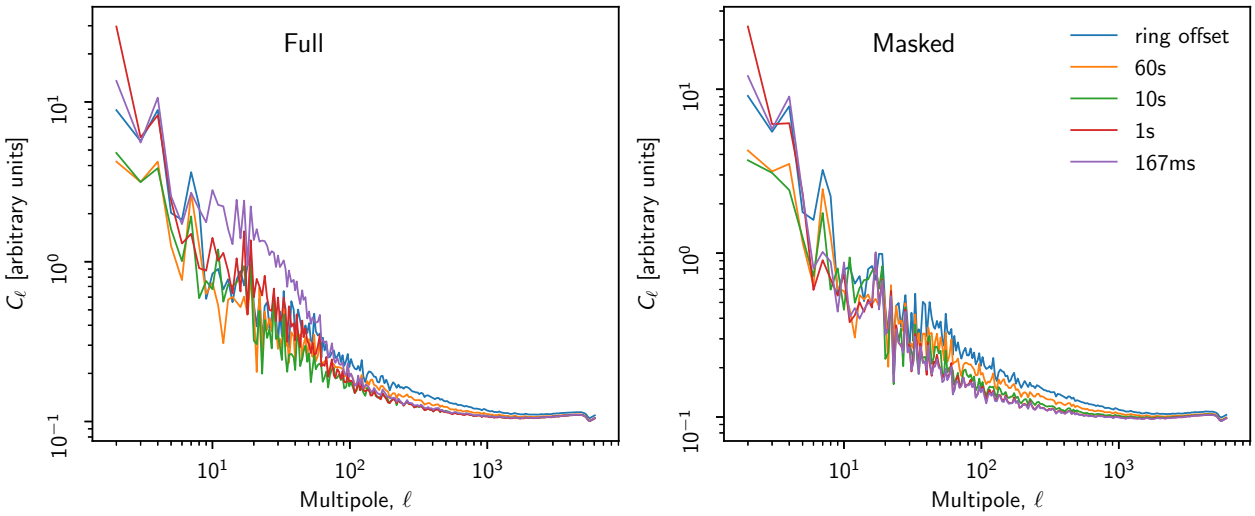


Fig. 13. Power spectra of destriped dark-bolometer maps using various baseline lengths. The "full" case on the *left panel* shows the pseudo-C_ℓ power spectrum over the entire map, including the Galactic plane that was masked out while solving for the baseline offsets. *Right panel*: power spectrum taken over the sky pixels that were considered when solving for the baselines. Shorter baselines clearly suppress noise above $\ell = 20$. The 167 ms case shows an elevated residual at $\ell < 100$ inside the destriping mask, likely due to the filter's inability to constrain the solution. The residual at $\ell < 20$ is dominated by the irreducible uncertainty set by the *Planck* scan strategy, and cannot be improved without external information.

4.1. Maps

All NPIPE maps are calibrated to thermodynamic temperature units in kelvins (K_{CMB}). The 857 GHz calibration is the same as in PR3, but is converted into K_{CMB} temperature units using the measured bandpasses. Note that NPIPE maps include the Solar dipole; it may be removed using the parameters supplied in Sect. 8.

The conversion factor from thermodynamic temperature to flux density units is $58.04\,\mathrm{MJy\,sr^{-1}}/\,K_{CMB}$ for the 545 GHz maps and $2.27\,\mathrm{MJy\,sr^{-1}}/\,K_{CMB}$ for the 857 GHz maps (see Planck Collaboration IX 2014).

4.1.1. Frequency maps

We present temperature maps at all nine *Planck* frequencies in Fig. 14, and with the Solar dipole removed in Fig. 15. The zero levels of the maps are adjusted for plotting by evaluatings the monopole outside a Galactic mask. The mapmaking procedure leaves the true monopole undetermined. Stokes Q and U polarization maps and the polarization amplitude at the seven polarized frequencies are shown in Fig. 16. Polarization maps smoothed to 3° are shown in Fig. 17.

4.1.2. Half-ring maps

The repetitive *Planck* scanning strategy (each scanning circle is repeated 30–75 times) allows us to build subset maps that have effectively identical systematics but independent instrument noise. We split each pointing period into two half-rings, and assign the halves to separate subsets, which we call "HR1" and "HR2". The half-ring difference (i.e., HR1−HR2) is a useful measure of the instrumental noise, but should not be assumed to be unbiased for large-scale cross-spectral analysis. The half-ring maps share all the calibration, bandpass mismatch, and other template residuals.

The noise in the HFI half-ring maps contains a small amount of correlated error between the two ring halves (cf. Sect. 2.3.4 and Fig. 18). This error comes from a noisy signal estimate used in the removal of the glitches.

4.1.3. Single-detector maps

NPIPE builds unpolarized single-detector maps from the reprocessed TOD by subtracting an estimate of the polarization response from the full-frequency map. The single-detector maps are destriped with Madam independently, using the same destriping parameters as for the full-frequency data. The maps are provided with and without bandpass-mismatch corrections. Mismatch-corrected maps provide a useful diagnostic for reprocessing performance, while the uncorrected maps are used as inputs to temperature component separation. This provides component-separation algorithms with samples of the sky for each detector bandpass, which is necessary, for example, to measure the CO line emission. Sampling and subtracting the polarization signal from the full-frequency map injects a small amount of correlated noise into the single-detector maps. This noise is attenuated by the fact that we downgrade the frequency map to the same resolution as our bandpass-mismatch templates, i.e., $N_{side} = 512$ for 30–100 GHz and $N_{side} = 1024$ for 143–353 GHz. Most of this correlated noise is cancelled in the single-horn maps (Sect. 4.1.4), so applications that are particularly sensitive to small-scale noise correlations are best served using the horn maps instead.

4.1.4. Single-horn maps

In addition to polarization-corrected single-detector maps, we also provide unpolarized single-horn maps that naturally cancel most if not all of the polarization. These maps are linear combinations of the polarization-orthogonal detector data in the polarized *Planck* horns, with the weights chosen to maximally cancel polarization in the averaged map. The process also eliminates noise present in the resampled polarization estimate.

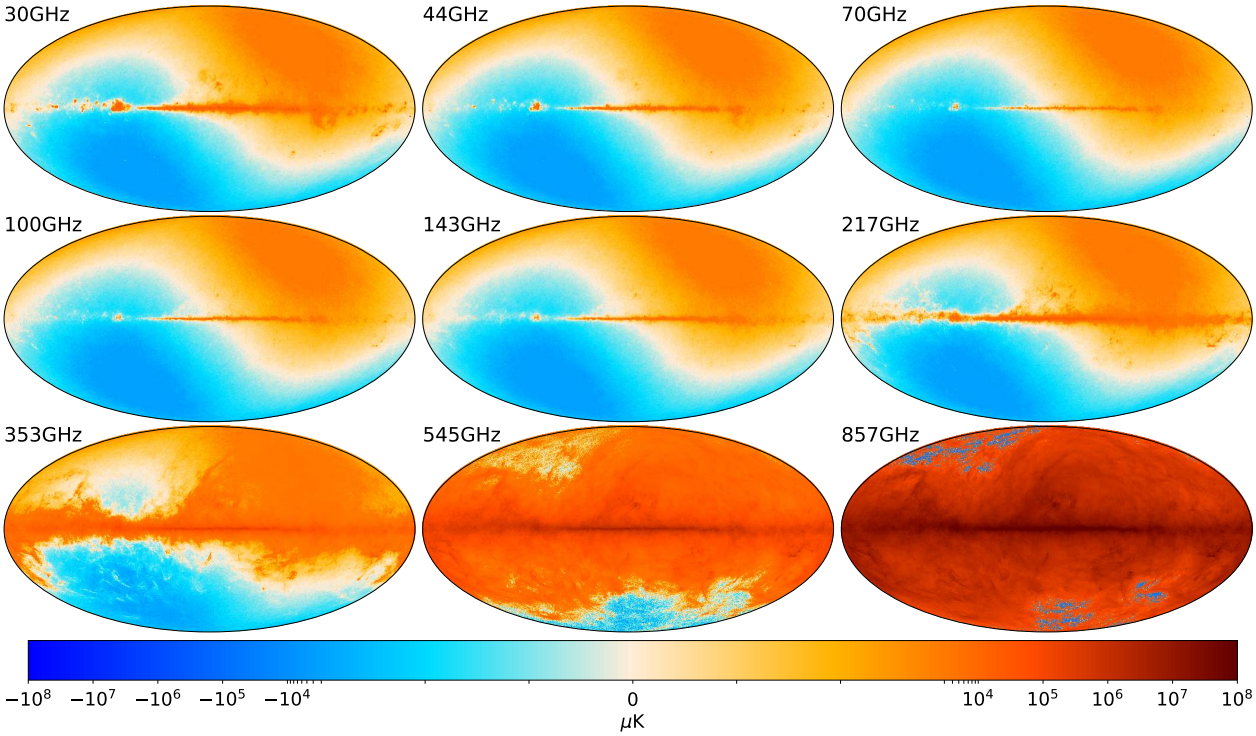


Fig. 14. NPIPE temperature maps, including the Solar dipole. The scaling is linear between −3 and 3 mK.

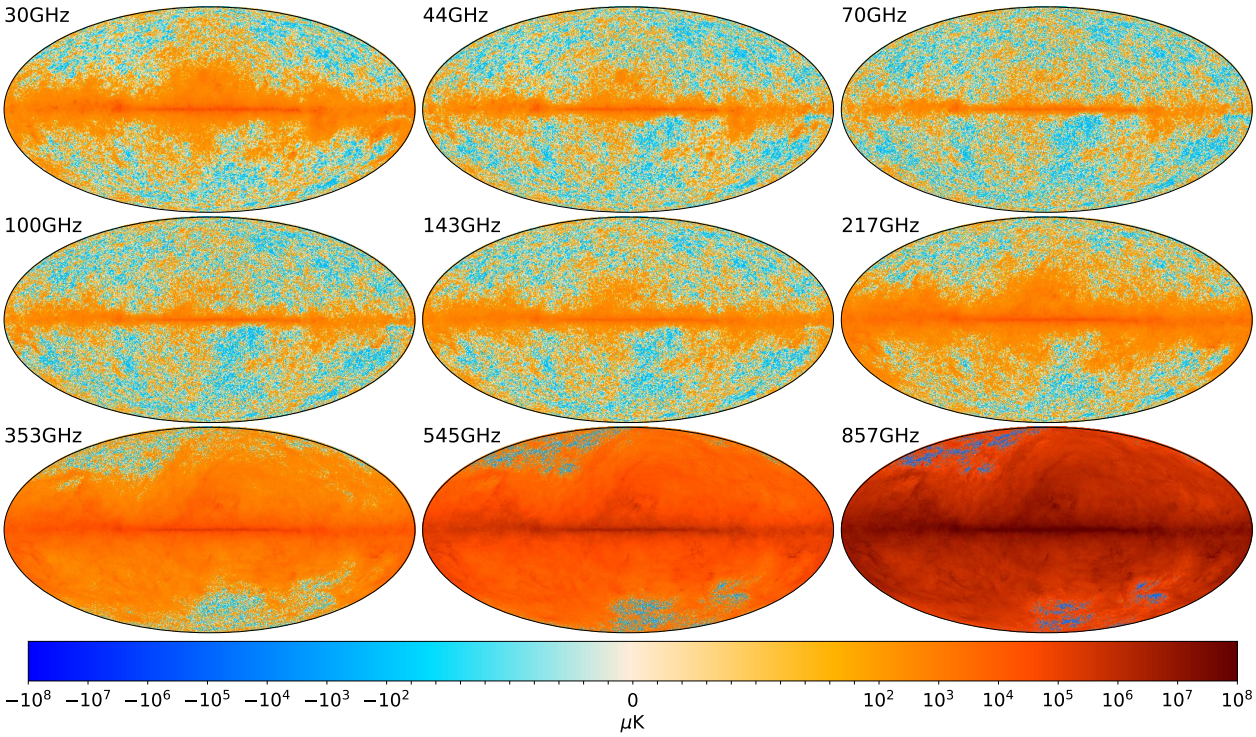


Fig. 15. NPIPE temperature maps, with the *Planck* 2015 dipole removed. The scaling is linear between −100 and 100 μK.

4.1.5. A/B maps

NPIPE provides only one data split where the systematics between the splits are expected to be uncorrelated (unlike the case for the half-ring splits discussed in Sect. 4.1.2). We have split the horns in the focal plane into two independent sets, A and B, and performed reprocessing independently on each set.

The *Planck* scanning strategy requires at least two polarized horns to solve for a full-sky polarized map. The detector-set splitting was not possible at either 30 or 44 GHz because of the lack of redundant polarized horns. Instead, for these two frequencies, the split is done time-wise: set A comprises operational years 1 and 3; while set B has years 2, 4, and the month of integration time from the final, fifth observing year.

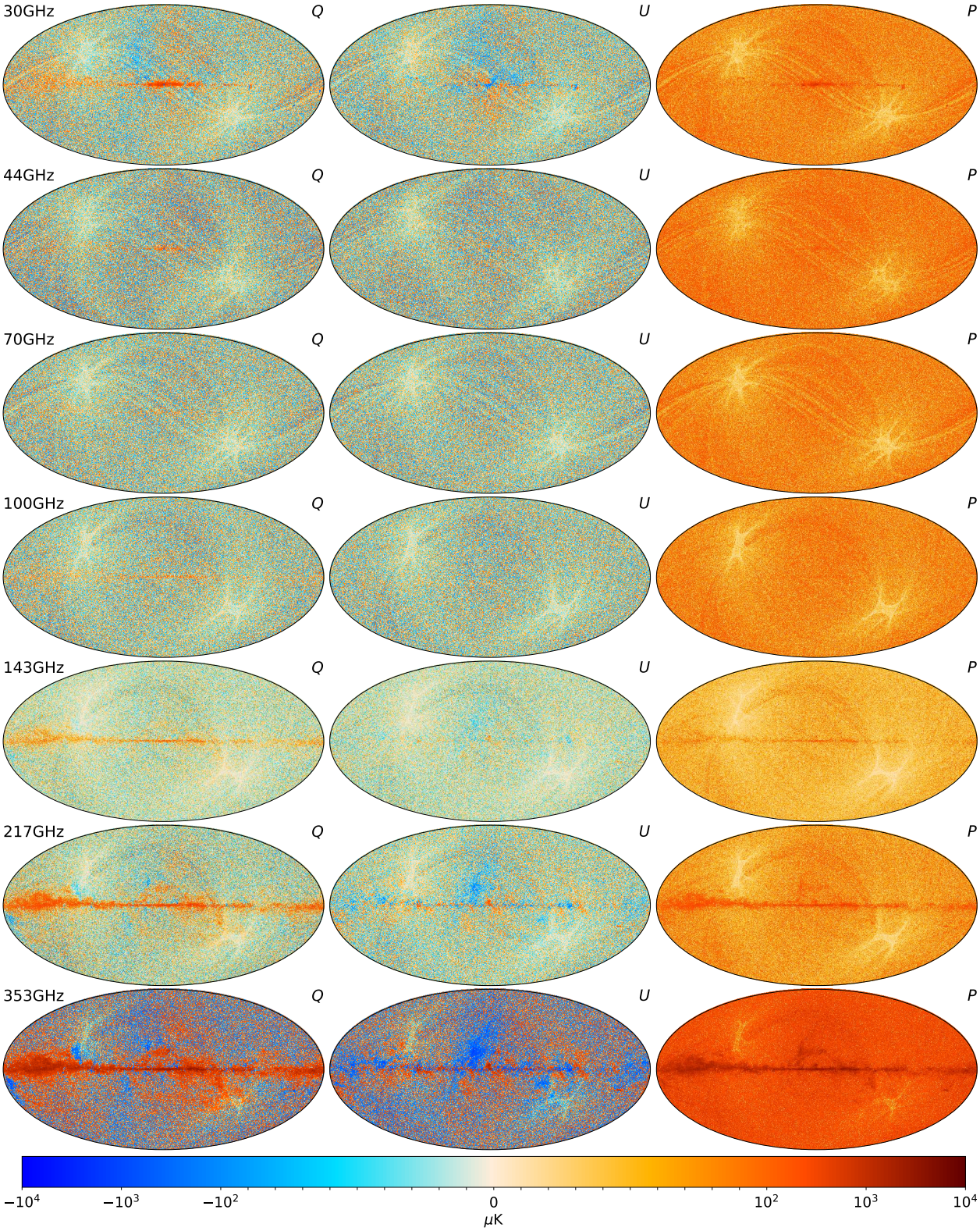


Fig. 16. NPIPE polarization maps. The scaling is linear between −100 and 100 μK. Here $P = (Q^2 + U^2)^{1/2}$.

(A half-mission split is not feasible because the second half mission does not have full sky coverage.) The time-wise split is not expected to have independent systematics, since the two subsets will share initial beam and bandpass mismatch; however, the instrument noise and gain fluctuations will be uncorrelated between the two subsets, making the time-wise split at these frequencies a reasonable compromise to provide two consistent and maximally disjoint sets of *Planck* data. Full descriptions of the A/B splits are provided in Table 5.

The NPIPE approach of processing maximally-independent subsets of data is orthogonal to the approach adopted by HFI in PR3. There the systematic templates were fitted on the entire

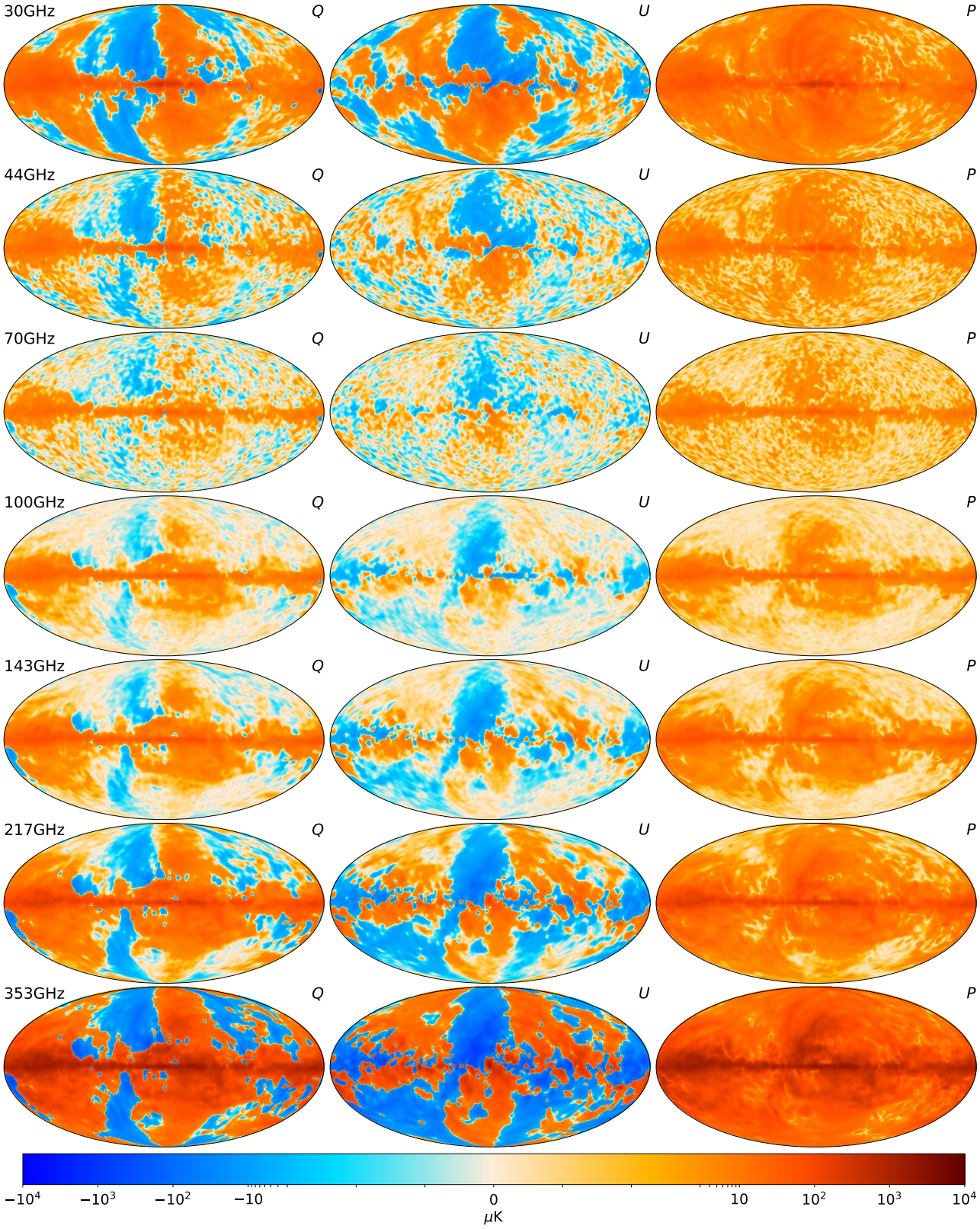


Fig. 17. NPIPE polarization maps smoothed with a Gaussian beam with FWHM 3°. The scaling is linear between −3 and 3 μK.

frequency data set, and the cleaned data were split only for the purpose of projecting the TOD into subset maps. Such processing introduces correlations between systematic residuals in the subset maps, making it necessary to estimate and correct for noise bias even in cross-spectra between the subsets. This is part of the reason Planck Collaboration III (2020) emphasizes the importance of using inter-frequency cross-spectra in extracting scientific information from PR3 maps.

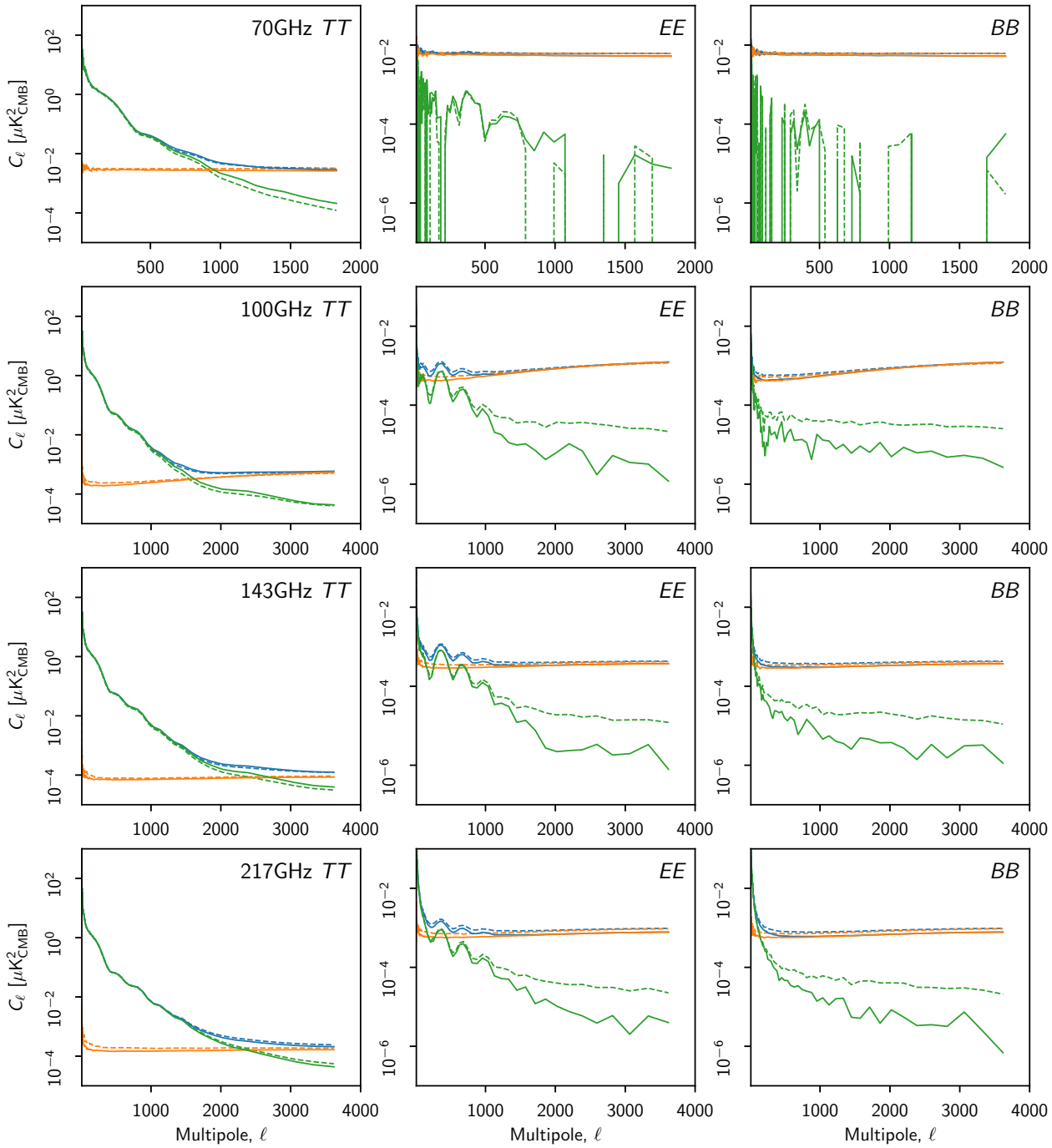


Fig. 18. Half-ring sum (blue), difference (orange), and cross (green) power spectra from NPIPE (solid lines) and PR3 (dashed lines). Improvements in gap-filling and transfer-function deconvolution reduce the small-scale half-ring correlations in NPIPE data from PR3 levels. The power spectra are binned into 100 logarithmically-spaced bins, and corrected for sky fraction. The 70 GHz results are shown as a reference case without half-ring correlations.

4.2. Low-resolution data set

We provide low-resolution versions of the NPIPE maps for maximum-likelihood pixel-domain analysis of the large scales. The maps are downgraded to HEALPix resolution $N_{\rm side} = 16$ and convolved with a cosine apodizing kernel:

$$b_\ell = \begin{cases} 1, & \ell \le \ell_1, \\ \frac{1}{2}\left(1 + \cos\left[\pi \frac{(\ell-\ell_1)}{(\ell_2-\ell_1)}\right]\right), & \ell_1 < \ell \le \ell_2, \\ 0, & \ell > \ell_2, \end{cases} \quad (15)$$

with the choice of $\ell_1 = 1$, $\ell_2 = 3N_{\rm side}$. The lower threshold has been decreased from the usual $\ell_1 = N_{\rm side}$ to reduce ringing in the 857 GHz maps. In addition to the smoothing kernel, the maps include the standard $N_{\rm side} = 16$ pixel window function. Our downgrading tool noise-weights each high resolution pixel with the estimated noise level and parallel-transports the polarization parameters between pixel centres. Initial high resolution pixel window function was not deconvolved as it is negligible at the angular scales that are represented in the low resolution maps.

The low-resolution maps are accompanied by pixel-pixel noise-covariance matrices that reflect the baseline uncertainties and the high-frequency instrumental noise approximated as uniforms white noise. Since the high-frequency noise in NPIPE is

Table 5. Details of the A/B subset split.

Frequency [GHz]	Set A	Set B
30	Years 1 and 3	Years 2, 4, and start of 5
44	Years 1 and 3	Years 2, 4, and start of 5
70	Horns 18, 20, and 23	Horns 19, 21, and 22
100	Horns 1 and 4	Horns 2 and 3
143	Horns 1, 3, 5, and 7	Horns 2, 4, and 6
217	Horns 1, 3, 5, and 7	Horns 2, 4, 6, and 8
353	Horns 1, 3, 5, and 7	Horns 2, 4, 6, and 8
545	Horn 1	Horns 2 and 4
857	Horns 1 and 3	Horns 2 and 4

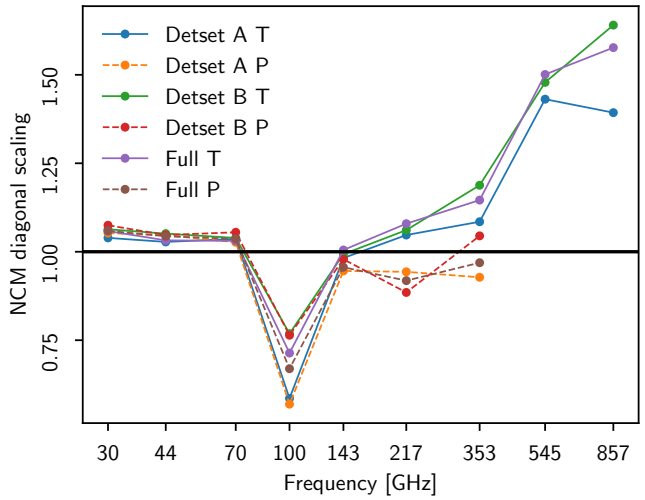


Fig. 19. Scaling factors applied to the Madam noise matrices to match the variance in half-ring-difference pixels. We apply different scalings to the temperature and polarization parts of the noise matrices, and scale the detector-set matrices and the full-frequency matrices independently. The 100 GHz channel stands out because of the prominent bump at the high-frequency end of the noise PSD (cf. Fig. 27).

not uniform for any of the *Planck* detectors, the approximation is inexact and, as a result, requires an extra scaling step performed on the diagonal of the matrix to match the half-ring difference maps. The off-diagonal elements correspond to low frequency noise and need not be scaled. We show the scaling factors in Fig. 19. For the CMB channels, these factors are close to unity except at 100 GHz.

4.3. Large-scale polarization transfer function

NPIPE calibration at the CMB frequencies (44–217 GHz) acts like a constrained filter. The polarized sky model that we derive from the extreme frequencies contains, for all practical purposes, only the foregrounds. This leaves NPIPE actively trying to suppress the CMB polarization with the available templates. While obviously undesirable, the effect of this filtering is kept small by the number and structure of the templates. It is also straightforward to measure the amount of suppression using simulations and incorporate the effect into the transfer functions.

Allowing for the non-trivial transfer function in NPIPE calibration is a compromise between measuring very noisy but unbiased large-scale polarization from all large-scale modes, and filtering out the modes that are most compromised by the *Planck* calibration uncertainties left in the data by the *Planck* scan strategy (see Figs. 40 and 41). Based on similarities between LFI 2018 processing and NPIPE it is possible that the LFI 2018 results were also subject to filtering. Unfortunately the accompanying simulations were not sufficient to determine this as only one CMB realization was processed with the calibration pipeline. Algebraically there is no reason to anticipate the HFI 2018 results to contain a similar effect but, again, the 2018 simulations cannot be used to demonstrate this.

The filtering approach is typically adopted by ground and balloon-borne experiments to suppress atmospheric noise in their data. The only requirement for the filtering to be workable is that the resulting ℓ-dependent transfer function be measurable. We show in Fig. 20 that it is well characterized in the case of NPIPE analysis of *Planck* data. Tabulated values of the transfer function can be found in Appendix G.

4.3.1. CMB transfer function

We measure the transfer function from our simulations by comparing the input CMB map to the output foreground-cleaned map. We smooth and downgrade both maps and compare the input (CMB × CMB) cross-spectrum to the CMB × cleaned map. If the effect is multiplicative, as is evident from Fig. 22, we expect to find

$$C_\ell^{\mathrm{CMB\times output}} = k_\ell\, C_\ell^{\mathrm{CMB\times CMB}}, \tag{16}$$

where k_ℓ is a potentially scale-, frequency-, and mask-dependent suppression factor. The left panel of Fig. 20 shows the k_ℓ for each CMB frequency. The right panel of Fig. 20 compares the 100 GHz transfer functions evaluated over sky fractions ranging from 0.1 to 0.9 and shows to what degree the shape of the transfer function depends on the mask. These transfer functions, like the beam window functions, must be squared to obtain the full impact on a power spectrum. We find that the EE quadrupole and octupole are significantly suppressed by our calibration, with about 64% of the quadrupole and 73% of the octupole missing in the NPIPE power spectra. Multipoles 4 and 5 are missing about 30% of the power, and $\ell = 6$ is missing about 20%. The simulations do not include enough CMB B modes to measure the B-mode transfer function and changing the signal content would compromise our interpretation due to the nonlinear calibration process.

The suppression of EE power is limited to frequencies that are calibrated with the polarization prior. There is no suppression at 30 GHz or 353 GHz (Fig. 21). Measuring the transfer function using the CMB polarization at these frequencies is much harder because of the foregrounds and noise. This is evidenced by the large upwards fluctuations at 353 GHz, particularly at $\ell = 4$ and 9. These fluctuations are associated with excess power at the level of the CMB E-mode fluctuations, rather than suppression of power. Their importance diminishes when these frequency maps are scaled to the CMB frequencies to estimate the foregrounds.

Figure 22 shows individual input-output EE multipole pairs for 100 GHz. The data points show a great deal of scatter owing to the limited S/N. Nevertheless, fitting for a linear model between the input and output values robustly identifies the degree to which the input EE power is suppressed by the NPIPE calibration. It is also apparent that a simple multiplicative correction will remedy the bias for all realizations.

Frequency and detector-set cross-spectra may be impacted by a different transfer function. The measurement also potentially differs, since we must compare the input spectrum to the foreground-cleaned cross-spectrum:

$$C_\ell^{\mathrm{output1\times output2}} = k_\ell^2\, C_\ell^{\mathrm{CMB\times CMB}}. \tag{17}$$

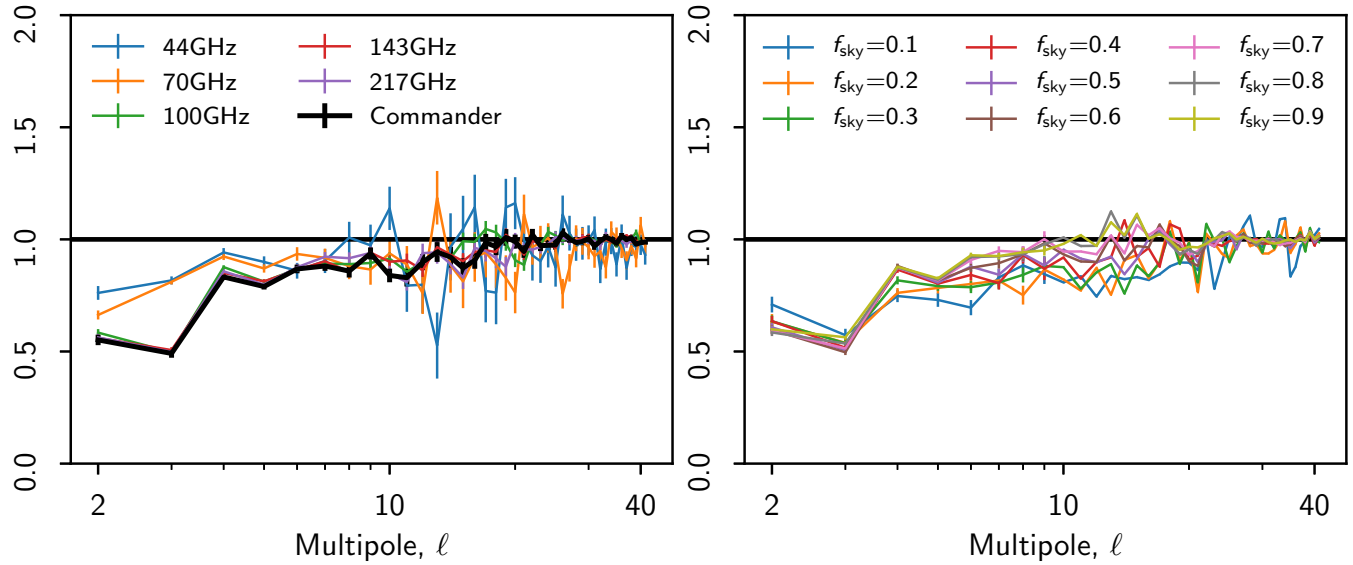


Fig. 20. NPIPE *E*-mode transfer functions measured by comparing simulated CMB input and foreground-cleaned output maps. *Left*: CMB frequencies and component-separated Commander maps (Sect. 7) over 60% of the sky. The apparent mismatch between the LFI and HFI transfer functions results from the quantity and structure of the template corrections; templates that are specific to HFI, especially the ADC distortion, provide more degrees of freedom to suppress the CMB power. The 44 GHz transfer function is closer to unity because the 30 GHz template shields about 22% of the CMB polarization. The error bars reflect the statistical uncertainty of the measured transfer function, not the total Monte Carlo scatter. Tabulated values of the transfer functions are listed in Table G.1. *Right*: *E*-mode transfer function for 100 GHz over multiple sky fractions. The error bars at $\ell \geq 10$ were suppressed to show more structure.

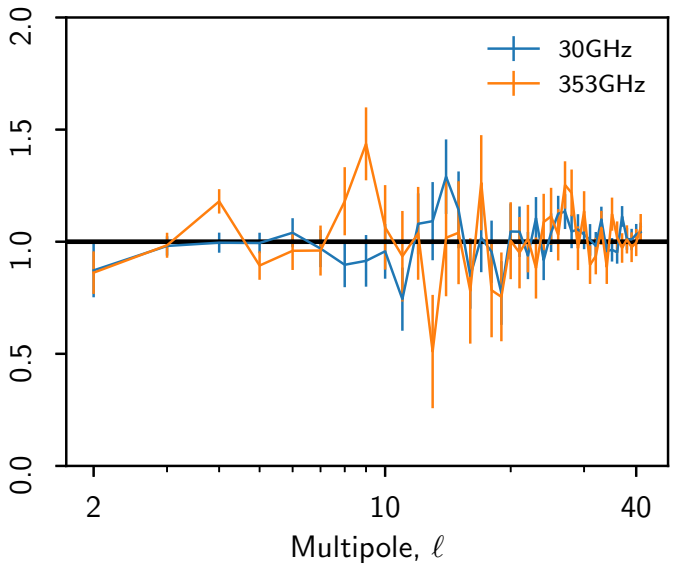


Fig. 21. NPIPE *E*-mode transfer functions, measured by comparing simulated CMB input and foreground-cleaned output maps over 60% of the sky. The 30 and 353 GHz frequencies are not expected to have a measurable transfer function because they are not calibrated with a polarization prior. These two transfer functions are merely of diagnostic value (to demonstrate the absence of signal suppression) and are not applied in any analysis. The error bars reflect the statistical uncertainty of the measured transfer function, not the total Monte Carlo scatter. Tabulated values of the transfer functions are listed in Table G.1.

We find that the use of the simulated cross-spectra to be problematic for two reasons: they are much noisier than the CMB×output spectra; and they may contain persistent bandpass-mismatch residuals. For these reasons, we instead derive the cross-spectral transfer functions as the geometric mean of the individual transfer functions. The geometric mean was found to be in agreement with the Eq. (17) and offered superior noise performance. We show the measured *E*-mode transfer functions for detector-set and frequency cross-spectra in Fig. 23.

We also considered measuring the full, anisotropic transfer function, a complex multiplier applied to every $a_{\ell m}$ mode of the input CMB sky. The results are shown in Fig. 24. We cannot reliably expand the foreground-cleaned CMB map from the simulations without masking out the strongest Galactic residuals. This is problematic for a study of anisotropy, because the mask already includes a preferred orientation. Instead, we write a least-squares minimization problem in terms of a transfer function acting on the input CMB expansion:

$$\chi^2 = \boldsymbol{r}^{\mathsf{T}}\boldsymbol{r} \quad \text{with} \quad \boldsymbol{r} = \boldsymbol{m} - \sum_{X\ell m} f^X_{\ell m} a^X_{\ell m} Y^X_{\ell m}, \quad X \in [T, E, B] \tag{18}$$

where the residual, $\boldsymbol{r}$, is estimated by subtracting a CMB expansion ($a^X_{\ell m} Y^X_{\ell m}$) convolved with a transfer function ($f^X_{\ell m}$). In this shorthand notation each basis function, $Y^X_{\ell m}$, is a full *IQU* map with either the *I* or the *QU* part identically zero. The sum in Eq. (18) is conveniently carried out using the HEALPix alm2map facility. This formulation allows us to use full-sky input expansions and still evaluate the residual over a masked sky.

The anisotropic transfer functions in Fig. 24 show a statistically significant anisotropy, especially at $\ell = 3$. This suggests that fully isotropic methods for power spectrum estimation and transfer function correction are not optimal in analysing the large-scale polarization in the NPIPE maps. Certain modes could be down-weighted to minimize uncertainty. Nevertheless, specializing the methods to utilize this information is complicated and not attempted in this paper.

To gain insight into the magnitude of the effect, we take a simulated CMB sky with a realistic amount of large-scale polarization power, and apply the 143 GHz anisotropic transfer

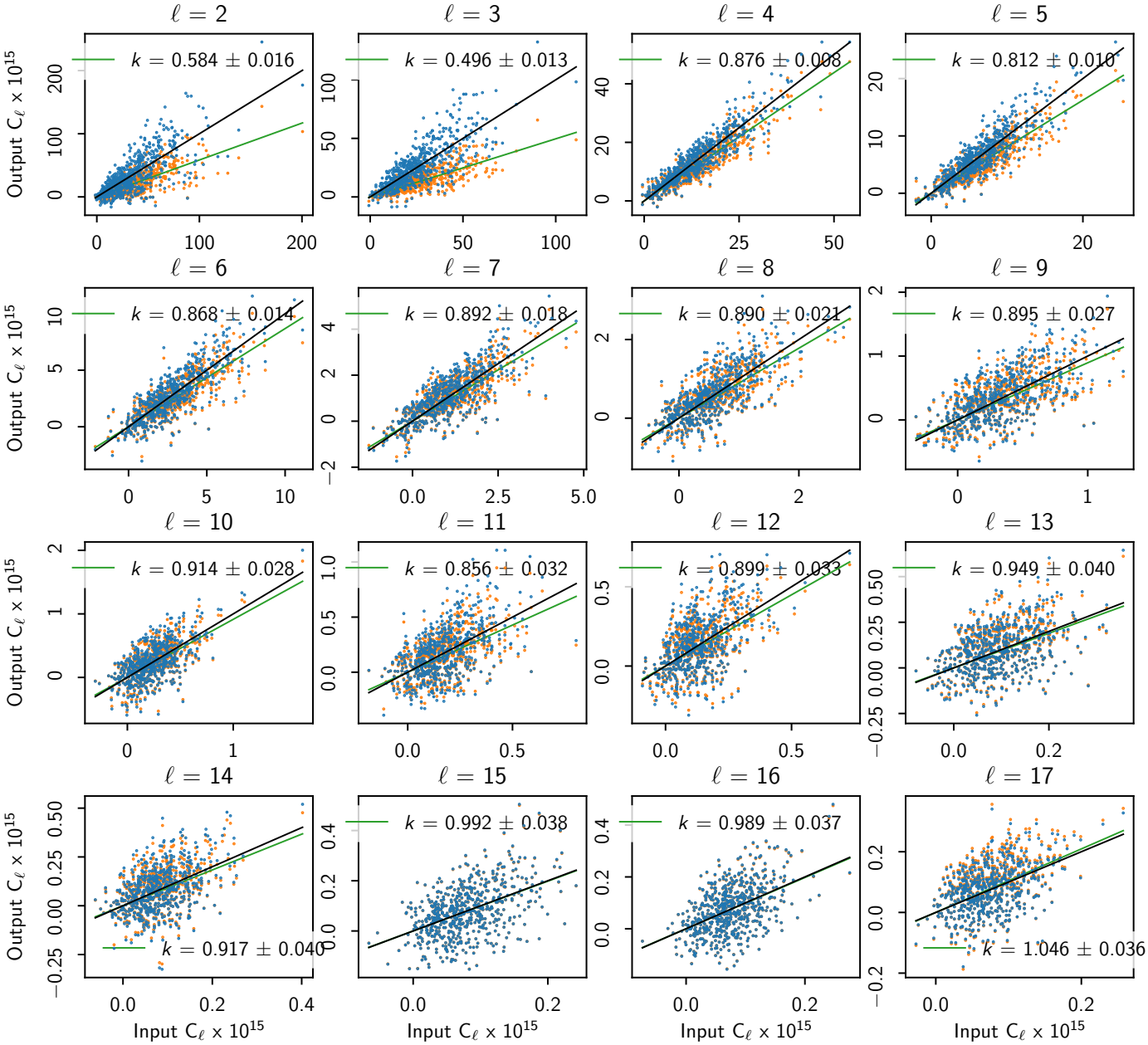


Fig. 22. Measuring the NPIPE large-scale polarization transfer function by comparing the simulated input and output CMB power multipole by multipole at 100 GHz and with 60% sky fraction. The output power spectrum is measured by removing the foreground from the simulated frequency map and then measuring the cross-spectrum with the input CMB map. The orange data points show input versus output before the transfer function correction. Their slope is shown in green. The blue data points show the corrected input versus output. Their (unity) slope is shown in black.

function from Fig. 24 to it. The input, output, and difference maps are shown in Fig. 25.

We provide the measured E-mode transfer functions for all frequency auto-spectra, frequency cross-spectra, and detector-set cross-spectra. The format of the files is exactly the same as the QuickPol beam-window-function files used in Planck Collaboration V (2020). In these files, the transfer function is unity everywhere except for the $\ell < 42$ E modes. These transfer functions were measured with the 60% sky mask. It is impossible to anticipate all sky masks and map combinations that a user may require, so we also provide software that enables measurement of the transfer function for arbitrary sky masks. To ensure fidelity of the science results, any statistic that employs large-scale ($\ell < 10$) CMB polarization should be corrected for the transfer function or otherwise calibrated with the simulations. This is the baseline for all NPIPE products.

4.3.2. Non-CMB transfer function

There is no algebraic reason to expect that the suppression of large-scale CMB polarization would extend to Galactic or extragalactic foregrounds. Indeed, measurement of these signals is driven by the 30 and 353 GHz channels, neither of which is subject to the polarization prior in the calibration. Furthermore, the presence of polarized foregrounds is accounted for in the calibration process by marginalizing over the polarization templates.

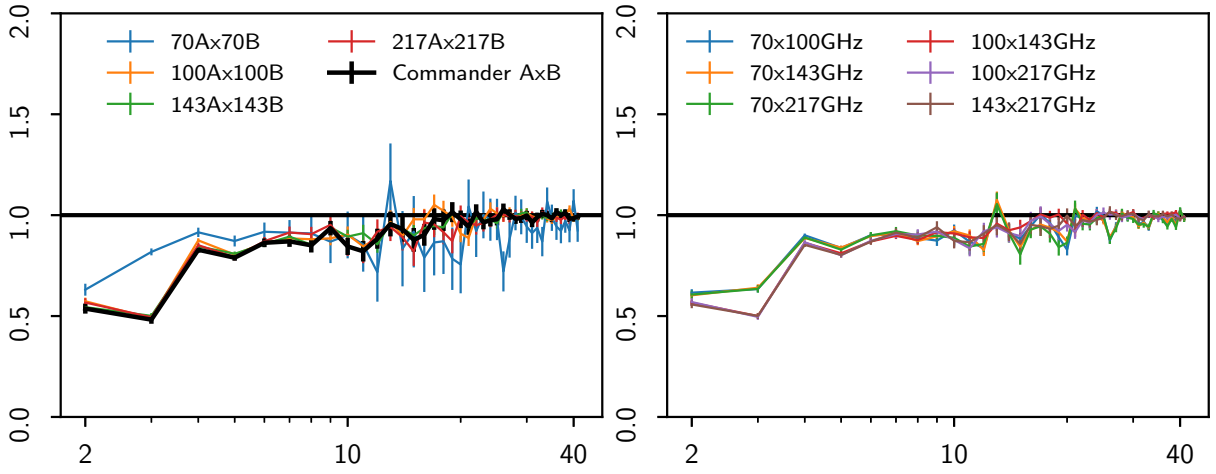


Fig. 23. NPIPE *E*-mode cross-spectrum transfer functions measured by comparing simulated CMB input and foreground-cleaned output maps over 60% of the sky. *Left*: detector-set cross-spectra. *Right*: frequency cross-spectra.

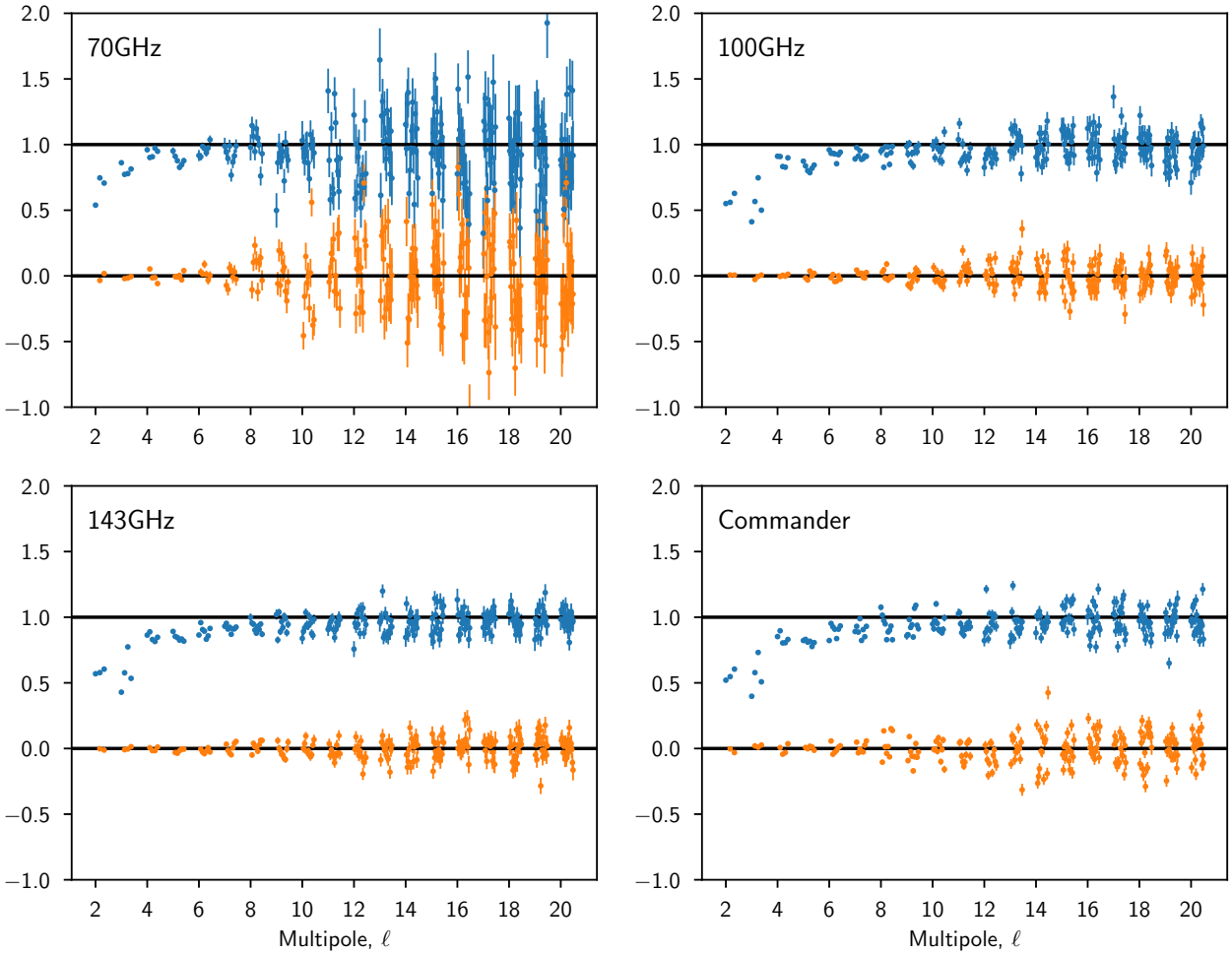


Fig. 24. Full anisotropic transfer functions for 70, 100, and 143 GHz and Commander. These demonstrate that the signal suppression is not completely uniform for the quadrupole and the octupole. The real part of the transfer function is shown in blue. The imaginary part (orange) of the transfer function is consistent with zero. These transfer functions correspond to minimizing Eq. (18) over 80% of the sky. The different m modes are slightly staggered here, with $m = 0...\ell$ plotted left to right, to aid visualization.

Foreground experts may object to the use of foreground templates to model the polarized foregrounds in the calibration process. Such templates are inherently incapable of supporting spatial variation of the spectral index of the foregrounds. When these templates make up the polarization prior, the calibration process does have the potential to suppress true spatial variation of the spectral index. We are confident that the limitations of our polarization prior have not compromised the foreground polarization for three reasons.

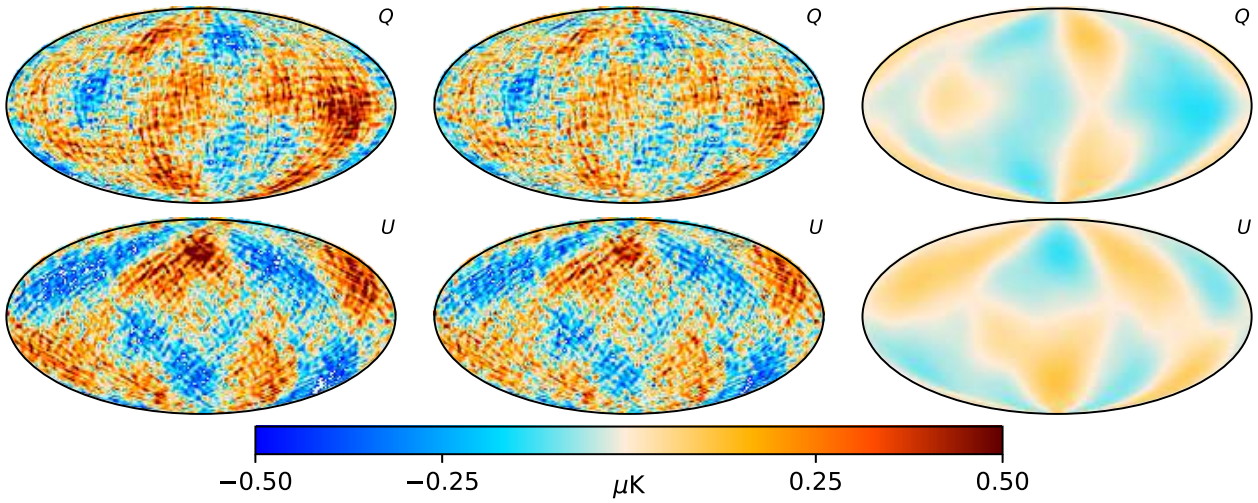


Fig. 25. Impact of the anisotropic NPIPE transfer function (Fig. 24) on a simulated CMB sky smoothed to 3°. The columns (*left* to *right*) are the input CMB, the transfer-function-convolved CMB, and the difference.

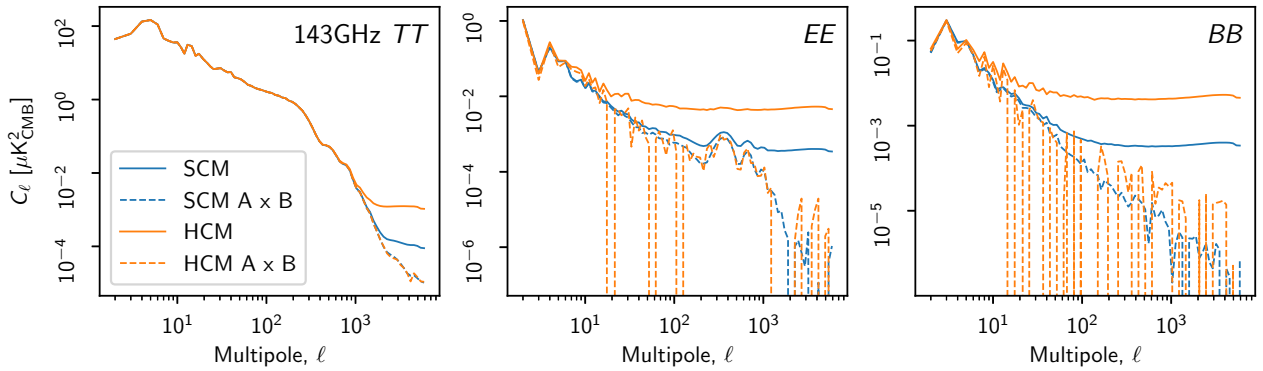


Fig. 26. Comparing the stable science scan (SCM) and the repointing manoeuvre (HCM) data. We find excellent consistency between the two disjoint data sets. The solid lines are auto spectra computed over 50% of the sky. The dashed lines are corresponding A/B cross-spectra. The spectra are binned into 100 logarithmically-spaced bins. The dashed orange spectrum in the *EE* panel demonstrates that even the 143 GHz repointing-manoeuvre data alone (which were excluded from all previous releases) are sensitive enough to probe at least the first three acoustic *EE* spectrum peaks. The agreement in small-scale *TT* cross-spectra indicates that the HCM pointing reconstruction is accurate enough not to widen the effective beam. The agreement in large-scale *EE* and *BB* power suggests that potential thermal effects from the thruster burns do not leak into large-scale polarization. The small scale noise in the HCM data set is higher than in the SCM data set from having only 8% of the integration time.

1. The calibration is performed using a Galactic mask, excluding the more intense Galactic plane and the source of more probable spatial variation of the spectral index.
2. Spatial variation in the dust spectral index has limited support by fitting both 217 and 353 GHz templates.
3. Our simulations that are based on the spatially-varying Commander sky model do not indicate any suppression of the spatially varying spectral indices (see Fig. 74).

4.4. Data taken during repointing manoeuvres

It is not immediately obvious that including data taken during the 4 min repointing manoeuvres is going to improve the resulting frequency maps and hence it is important to check for consistency. The three thruster burns of a given manoeuvre lead to a faint but measurable impact on the HFI thermal baseline, and the attitude reconstruction during the manoeuvre is admittedly less robust due to the dynamic nature of the data. We explored the consistency between "stable science mode" (labelled "SCM" in the *Planck* attitude history files) and the repointing mode ("HCM") by building 143 GHz full-frequency and detector-set maps exclusively from either SCM or HCM data. We show the power spectra of such maps in Fig. 26. These power spectra demonstrate excellent consistency between the two data subsets. Accordingly, data taken during pointing manoeuvres are used in all the NPIPE data products. *Only* SCM data were used in previous *Planck* products and results. This additional integration time reduces the small-scale noise uncertainty in NPIPE by approximately 9%.

5. Simulations

NPIPE is released with 600 high fidelity Monte Carlo simulations that include CMB, foreground, noise and systematics. These simulations allow testing and debiasing analysis tools with life-like frequency and detector-set maps with known inputs. For further details of the release, see Appendix A.

Calibrating against the same sky signal we are trying to measure is an inherently nonlinear problem and calibration and other template uncertainties in the NPIPE maps are not negligible. For these reasons we found it necessary to simulate the entire reprocessing part (Sect. 2.4) of the NPIPE processing. Each Monte Carlo iteration begins with a simulated CMB sky, represented as an expansion in spherical harmonics. We convolve

Table 6. Cosmological parameters of the FFP10 simulations.

Parameter	Value
$h^{(\star)}$	0.6701904
$\omega_{\rm b}^{(\dagger)}$	0.02216571
$\omega_{\rm c}$	0.1202944
τ	0.06018107
$A_{\rm s}$	2.119631×10^{-9}
$n_{\rm s}$	0.9636852
r	0.01

Notes. $^{(\star)}h$ is the dimensionless Hubble parameter: $H_0 = h \times 100\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$. $^{(\dagger)}\omega \equiv \Omega h^2$.

the CMB with high-order expansions of the final *Planck* scanning beams ($\ell_{\rm max} = 2048$ for LFI and 4096 for HFI). The convolution is carried out at the sample level (Prezeau & Reinecke 2010), respecting the actual beam orientation as a function of time. Foregrounds are simulated by evaluating the Commander sky model at the target frequencies. Static zodiacal emission is included by adding the same nuisance templates that Commander marginalized over. When components of the model are measured to much higher resolution than the target frequency (e.g., dust at 30 GHz), we smooth the foreground component with an azimuthally-symmetric beam expansion (Hivon et al. 2017) specifically calculated for the NPIPE data set. Bandpass mismatch is included by adding appropriately-weighted maps of foreground frequency derivatives and CO, customizing the foreground for each detector. The foreground map is then sampled into TOD, using the appropriate detector pointing weights, and co-added with the CMB TOD. Finally, we add the expected dipole and far-sidelobe signals. Noise is simulated from the measured noise PSDs as in Planck Collaboration XII (2016), including a correlated noise component in each polarized horn.

The simulated CMB sky includes lensing and frequency-dependent frame-boosting effects, as described in Planck Collaboration XII (2016). The CMB realizations are the FFP10 simulations used in PR3, and the cosmological parameters are listed in Table 6.

The gain fluctuations of several percent seen in LFI detectors are applied to the simulated TOD once all the components are added. Rather than trying to develop a statistical model that captures all the relevant features of the measured fluctuations, we apply the same smoothed version of the measured gain fluctuation to all Monte Carlo realizations.

HFI ADC nonlinearity is included into the simulations using the same tools as for LFI gain fluctuations, i.e., we apply a time-dependent gain fluctuation derived by smoothing the measured apparent gain fluctuations in the flight data. Similarly, we apply the measured transfer-function residuals to the simulated bolometer TOD.

5.1. Instrument noise

The $1/f$ instrument noise is simulated using the inverse fast Fourier transform (FFT) technique. A vector of complex Gaussian random numbers is generated, multiplied by a measured power spectral density (PSD), and then Fourier transformed into the time domain. We treat each pointing period as independent, enforcing no continuity across the pointing period boundaries.

Our noise estimation uses the same technique as described in Planck Collaboration XII (2016), i.e., we subtract an estimate of the signal by resampling the full-frequency map into the time domain, evaluate the correlation function over unflagged and unmasked samples, and Fourier transform to obtain the PSD. The low-frequency part of the PSD is measured from a downsampled version of the timestream to lessen the cost of estimating the covariance function. The traditional method of estimating the PSD directly from the TOD by FFT scales as $O(N \log N)$, but is forced to use all of the timestream samples, even the ones that contain a constrained realization of the data (e.g., to fill gaps) or are subject to steep gradients in the sky signal. The covariance function estimation used in NPIPEscales as $O(N^2)$, but allows for omission of the flagged samples and use of only the part of the sky where the signal estimate is robust.

The noise estimates are derived from preprocessed TOD to best match the noise seen during reprocessing. We apply the measured LFI gains during the estimation to remove the several-percent gain fluctuations from the noise estimates.

We have further developed the noise estimation code to use the covariance function method for measuring the correlated noise between detectors. Of particular interest is the correlated noise within each polarized horn (between the two polarization-orthogonal receivers). The correlated noise affects the sum of the two receivers, and thus the estimation of sky temperature; however, it cancels in the difference of the two receivers, leaving the polarization estimate unaffected. The noise estimates in Fig. 27 demonstrate that all polarized horns show significant internal noise correlations at low frequency, but only the HFI horns have significant noise correlations at high frequency. The HFI noise correlations are thought to result from coincident cosmic-ray glitches. The correlated HFI modes also exhibit several 4 K line residuals, which is understandable, given that the changes in the measured line-template amplitudes are highly correlated.

We simulate the correlated modes using the same inverse FFT technique as used for the uncorrelated modes, and then co-add them to the uncorrelated timestreams. Only the intra-horn component is simulated. We ignore the smaller and less relevant correlation between horns.

5.2. What is not included in the simulations

NPIPE simulations make an effort to capture all of the relevant systematics and their coupling in the final maps; however, the prohibitive cost of preprocessing precludes running the preprocessing module on simulated timelines. This leads to some notable omissions:

– ADC nonlinearity residuals beyond the linear gain fluctuation model (the higher-order distortion fluctuations are shown to be small in Fig. 11);

– 1 Hz and 4 K line residuals, except for what is captured by noise estimation;

– cosmic-ray glitch residuals, except for what is captured by noise estimation and modelled as Gaussian noise;

– measuring and updating the detector polarization parameters; and

– errors in the bandpass-mismatch templates, i.e., the sky model and the CO maps that are solved with Commander and fed back in future iterations of NPIPE.

Despite these omissions, we take the general agreement between flight data and simulated residuals in Figs. 28–31 as a validation of our approach.

HFI ADC nonlinearities are simulated in NPIPE using the linear gain model. This notably excludes the higher-order distortion modes discussed in Sect. 2.3.3. These modes are very faint and the fits include a fair amount of statistical uncertainty.

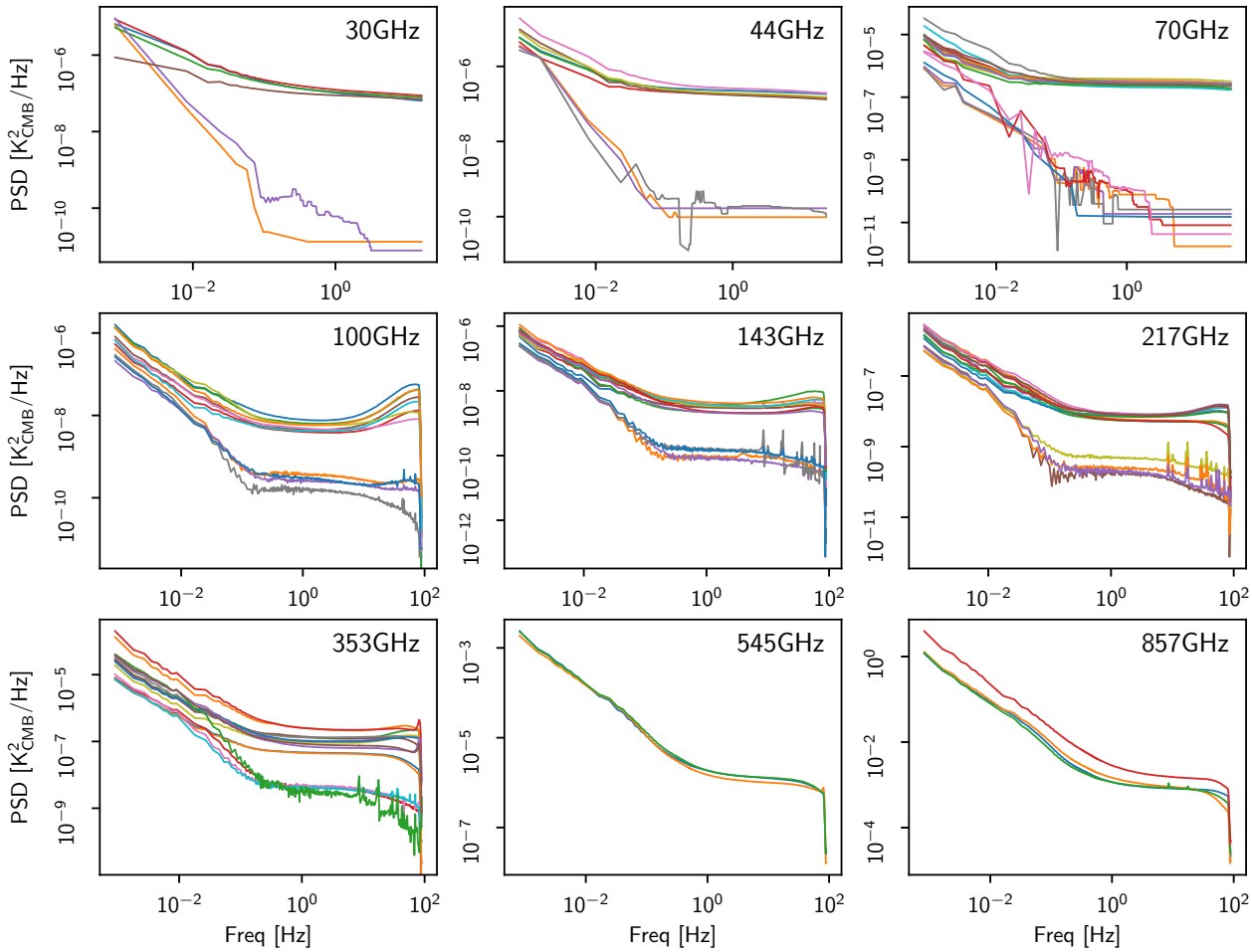


Fig. 27. Averaged noise PSDs for each detector (upper curves) and correlated-noise modes for each polarized horn (lower curves). The total noise power is the sum of the correlated and uncorrelated modes. These noise PSDs are measured from the data by subtracting a signal estimate and then evaluating the sample-sample covariance function. The HFI noise is suppressed near the Nyquist frequency (≈90 Hz) by the bolometric transfer function filtering. The PSDs are used for simulating the $1/f$ noise fluctuations, and as inputs to the `Madam` noise filter for destriping.

Without these higher modes of distortion, the simulated maps are affected only by the statistical uncertainty associated with the distortion correction. The excellent agreement between the large-scale polarization residuals in flight data and simulations suggests that the statistical uncertainties are sufficient to describe the large-scale polarization uncertainties in the `NPIPE` maps.

The line and glitch residuals would manifest themselves as small-scale excess power in the `NPIPE` maps. We only model these residuals to the extent that they can be measured in the detector noise PSD. The binned PSD cannot fully capture the Poissonian nature of the glitch residuals, nor the finely localized features associated with the lines. However, since the small-scale power of our simulated maps agrees very well with the real maps, we deduce that the mismatch is negligible.

Updating the instrument model using the single-detector residual maps and polarization estimates has not been formally included in the `NPIPE` processing plan, but has been done incrementally as post-processing over multiple revisions of the `NPIPE` maps. Simulating this process was deemed unfeasible given the time constraints. `NPIPE` simulations are therefore based on the final instrument model, including the adjustments to polarization efficiency and angle determined from flight data processing and shown in Table 4.

`NPIPE` HFI simulations differ from the FFP10 simulations used in Planck Collaboration III (2020) in some important ways. FFP10 implemented a light-weight version of the low-level data processing (without glitch or 4 K line removal) allowing for application of the ADCNL directly in the fast, modulated samples. It also allowed simulating the information loss from compression and decompression. A noteworthy complication of this approach was that the instrumental noise also needed to be simulated in the raw data, necessitating an extra noise-alignment step later in the pipeline. Analysis of the FFP10 results showed that the ADCNL was by far the most important systematic included in the simulation. `NPIPE` uses a more approximate approach, simulating the data directly in their preprocessed form based on noise estimates from the same stage data and adding ADCNL consistent with the linear gain model. Despite their differences, most effects do get included in both pipelines but using very different ways to express them. Our simulations are the first *Planck* simulations to apply the asymmetric scanning beam to each CMB realization separately in the time domain. In FFP10 only one beam-convolved CMB realization was passed through the full processing pipeline.

5.3. `NPIPE` simulation results

We compare the simulated CMB power to the simulated residuals in Figs. 32 and 33. They demonstrate that we are able to suppress large-scale noise and systematic uncertainties at or below

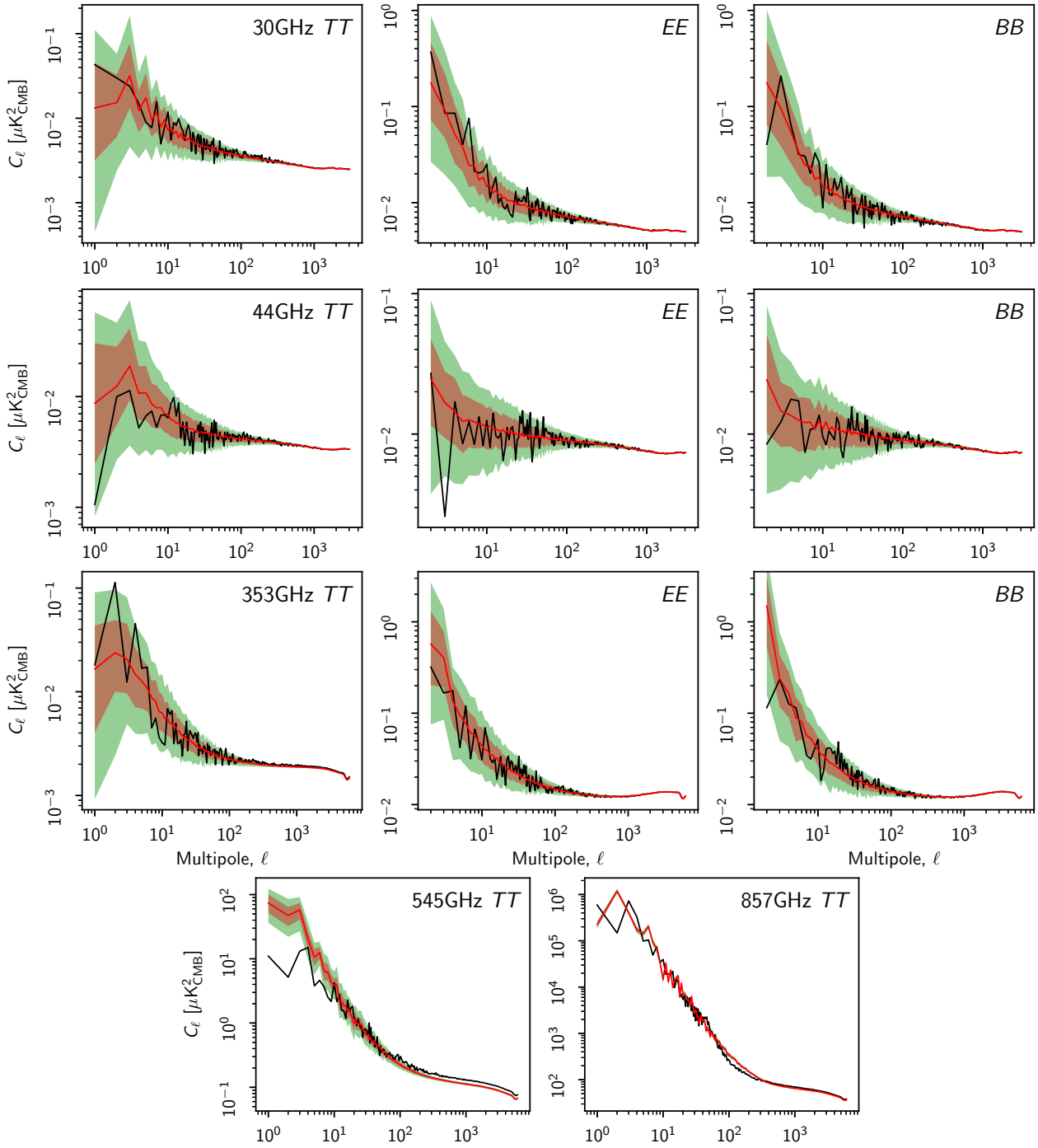


Fig. 28. Simulated A/B difference versus flight data A/B difference at 30, 44, 353, 545, and 857 GHz. The flight data are shown in black, and the median simulated power in each bin is plotted in red. The coloured bands represent the asymmetric 68% and 95% confidence regions. The power spectra are binned into 300 logarithmically spaced bins. These spectra are shown again in Fig. 34, but divided by the median simulated spectrum.

the faint large-scale *EE* power at the *Planck* CMB frequencies between 70 and 217 GHz.

5.3.1. Null maps and null map power spectra

We can demonstrate that the simulations succeed in capturing the relevant residuals by comparing real and simulated detector-set-difference power spectra in Figs. 28, 29, 34, and 35 (for *TE*, *TB*, and *EB* spectra, which are expected to be null, see Figs. 30 and 31). The detector-set differences exhibit excellent agreement between the flight data and the simulated detector-set differences in polarization. The temperature difference is not as well matched at the largest scales, owing to three factors. First, at 143 and 217 GHz the solved gain fluctuations seem to converge to a slightly different large-scale temperature sky for the two detector sets. The resulting roughly $1\,\mu K_{\rm CMB}$ mismatch is not fully captured in the simulations, suggesting that it is driven by a systematic that is not simulated. This residual and the associated mismatch are roughly 100 times fainter than the simulated CMB anisotropy at the same angular scales.

Second, at 545 GHz the simulations suggest almost an order of magnitude more large-scale temperature mismatch than is

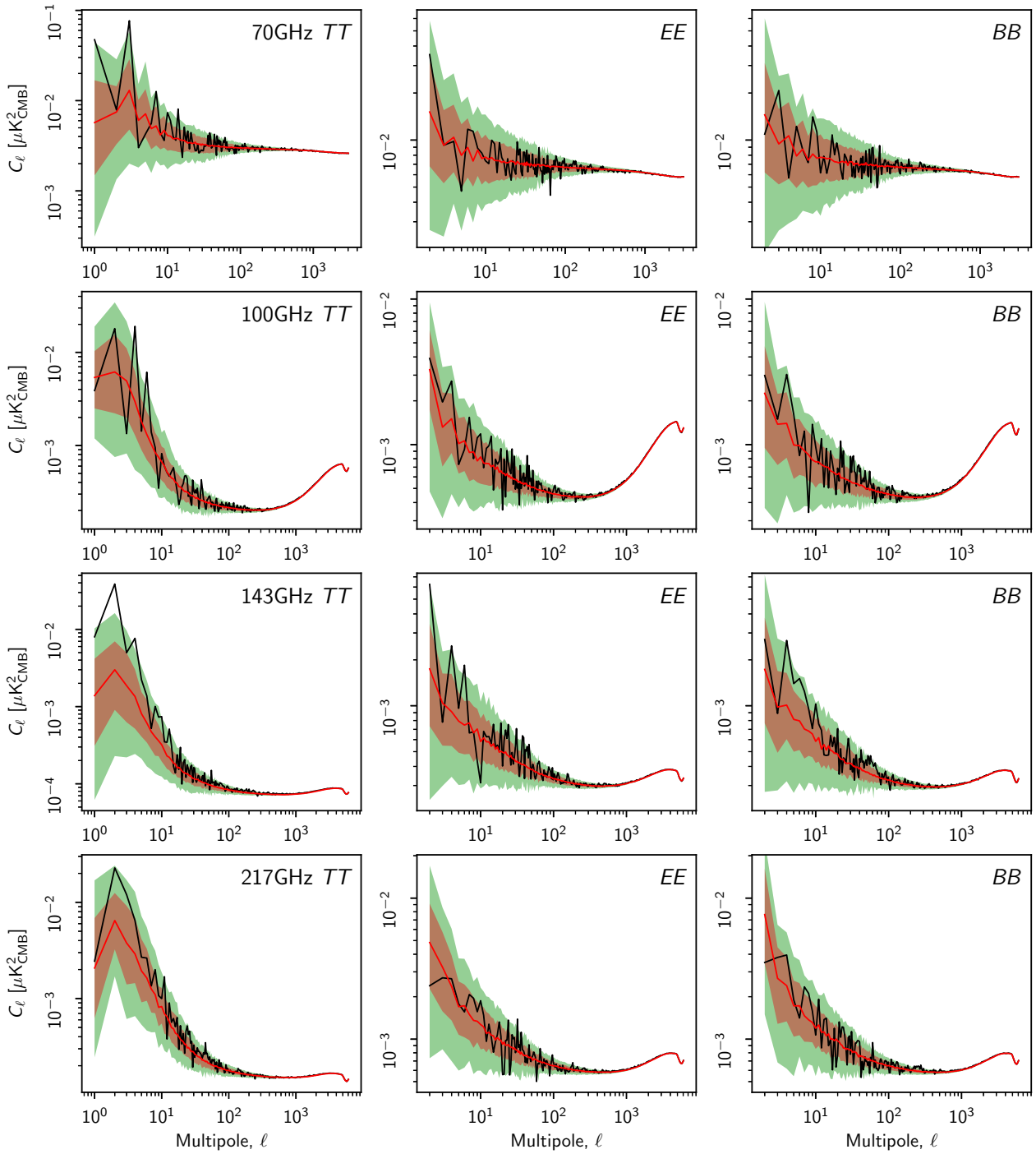


Fig. 29. Same as Fig. 28, at 70, 100, 143, and 217 GHz. These spectra are shown again in Fig. 35 but divided by the median simulated spectrum.

found in the flight data. This is explained by the crude bandpass-mismatch model (different delta-function bandpasses), coupled with the lack of redundancy at 545 GHz, where the "A" detector set comprises only a single bolometer. Even the amplified detector-set mismatch is 1000 times fainter than the dust in the 545 GHz band.

Third, at 857 GHz the dominant failure mode is the mismatch at intermediate scales between $\ell = 100$ and $\ell = 1000$. This mismatch is not seen in the half-ring differences, so it is not associated with the detector noise, but rather with systematic residuals related to the bandpass mismatch and bolometer transfer function. In the power-spectrum domain, this mismatch is 100–10 000 times fainter than the sky signal over the 50% of sky with the lowest dust emission. The 857 GHz detector sets only consist of two SWBs.

5.3.2. Noise alignment

The detector-set difference power spectra in Figs. 34 and 35 testify to a good overall agreement between the simulations and the flight data. Upon closer inspection it is possible to identify a percent level, scale-dependent deficit in simulated noise power at $\ell > 100$ affecting the polarized HFI frequencies (100–353 GHz). The deficit is exclusive to the HFI simulations, suggesting that its

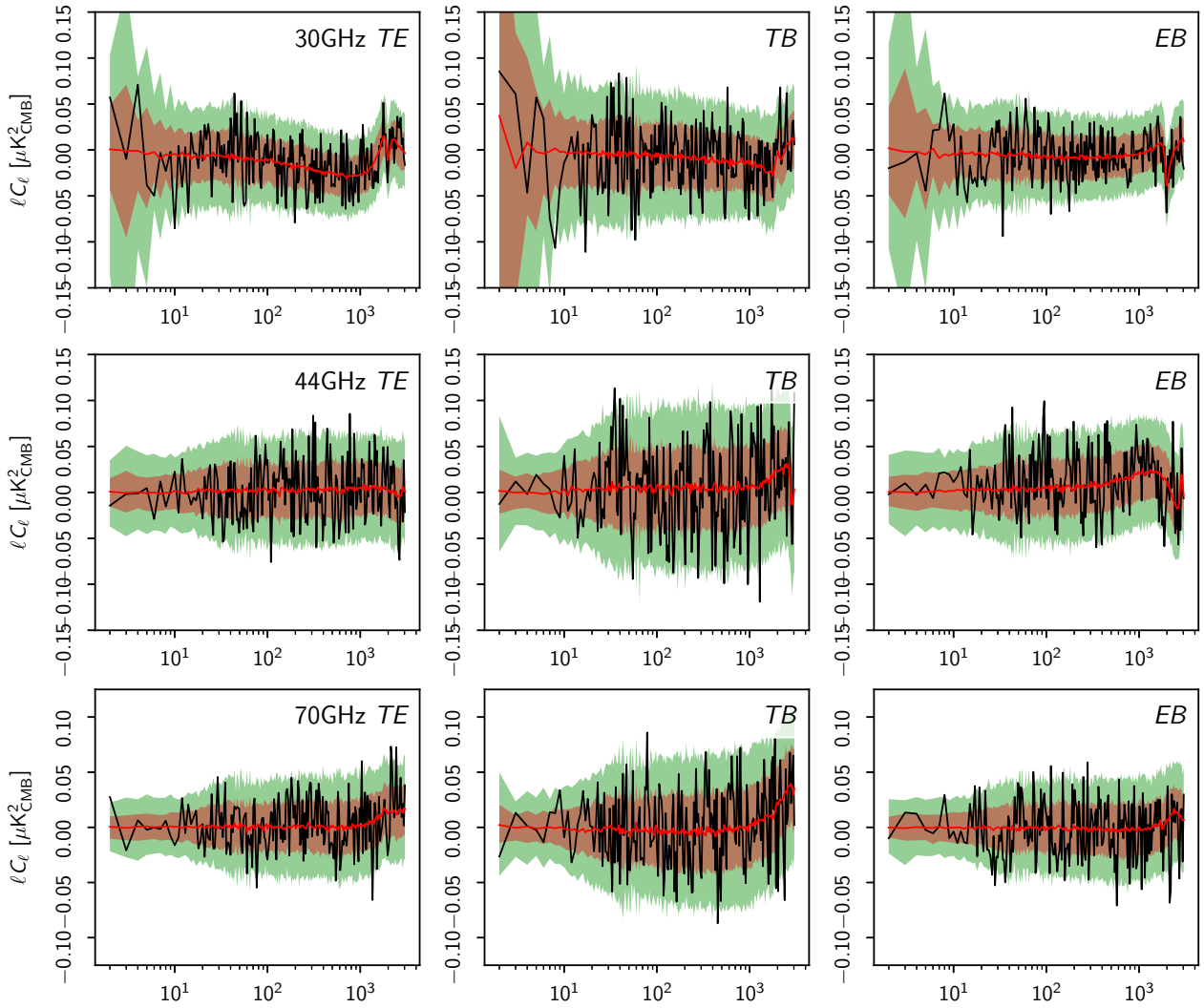


Fig. 30. Simulated A/B difference versus flight data A/B difference for TE, TB, and EB at 30, 44, and 70 GHz. Similar to Fig. 28, except for the ℓ-scaling.

origins lie in HFI specific phenomena, such as Poissonian glitch residuals that are not included in the simulations and could not be reproduced by a Gaussian noise simulation.

Overlooking the noise mismatch implies up to a percent level bias in statistical error measures that are derived from the simulations. For some analyses this degree of uncertainty is acceptable. For estimators that are sensitive to the mismatch in total power between the flight data and simulations, we offer a set of additional, 100–353 GHz, small-scale noise maps that can be added to the simulated NPIPE maps to align the total noise power. We stress that the additional noise maps only adjust the overall noise power, without any effort to match the spatial structure of these missing noise modes. We refer to this as "noise alignment". We demonstrate the effect of the noise-alignment maps in Fig. 36.

6. NPIPE and the earlier *Planck* releases

NPIPE differs from the earlier *Planck* releases in several ways. We compare NPIPE to both PR2 and PR3, allowing the reader to assess the magnitude of the changes against the 2015/2018 differences, which are themselves substantial. In the 2018 release, the LFI team made a critical change in calibration procedure, taking into account foreground polarization and thus removing a source of bias that generated spurious polarization. Also in 2018, HFI adopted a new mapmaking and TOD cleaning method called SRoll, which included time-domain corrections for bandpass mismatch, far sidelobes, and bolometer transfer-function residuals.

6.1. Calibration

NPIPE photometric calibration is measured by fitting the orbital dipole template (Sect. 2.4.3) to the data. The overall calibration is based solely on matching the far-sidelobe-corrected orbital dipole template against the data, rather than fitting a combination of the orbital dipole and some prior estimate of the total dipole. Our orbital dipole template includes the faint, frequency-dependent, dipole-induced quadrupole.

We follow the LFI approach, correcting the orbital dipole template for subtle suppression and phase shifts induced by the far sidelobes, and extend this treatment to the HFI frequencies. Details of the convolution are presented in Appendix C.

NPIPE processing effectively deconvolves the far sidelobes from the data, which are scaled to have unit response to the main beam rather than the 4π beam. This is reflected in the effective beam window functions, which have unit response to the

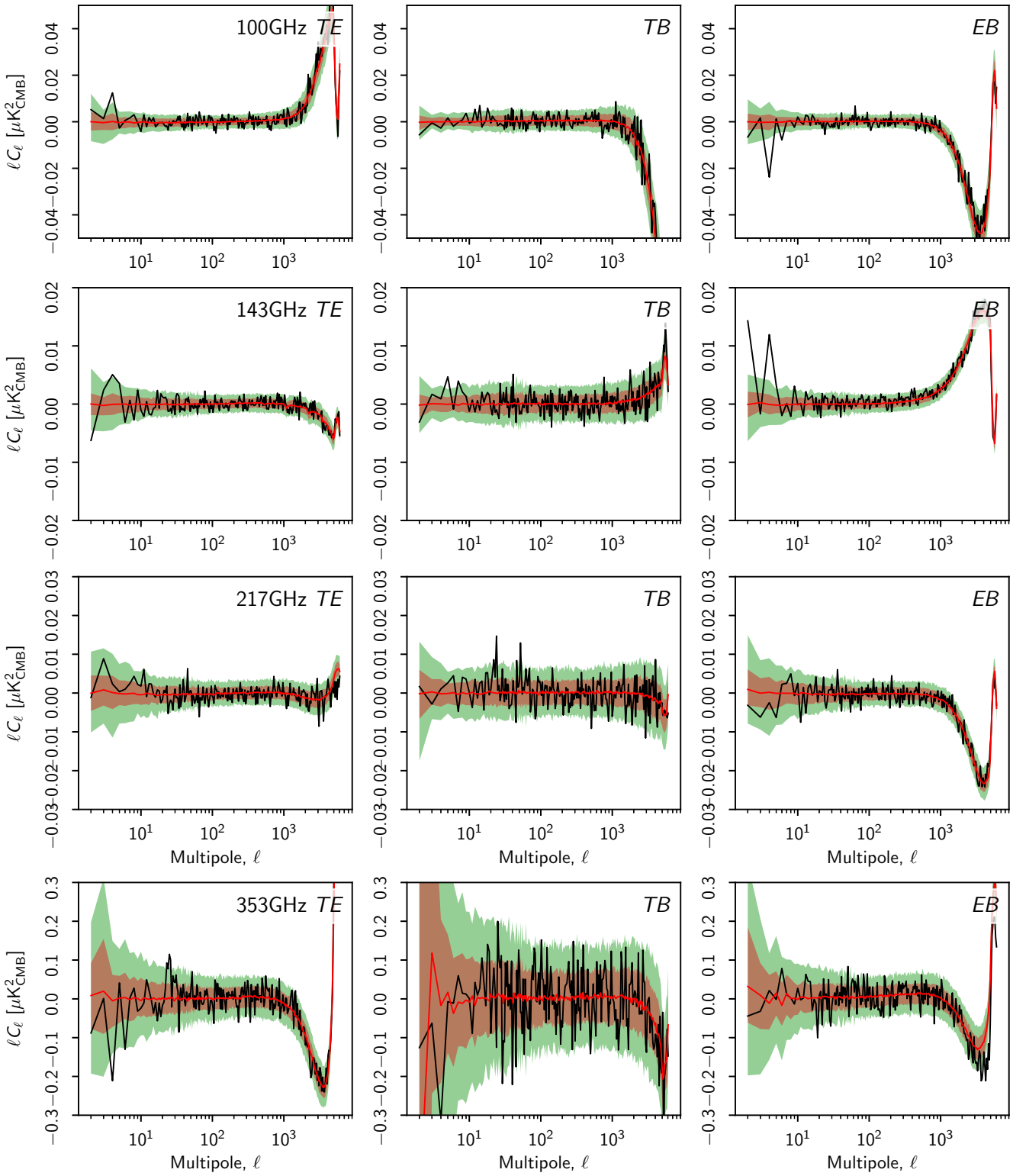


Fig. 31. Same as Fig. 30, for 100, 143, 217, and 353 GHz.

monopole and near unit response to the dipole. The uncorrected main-beam efficiencies are shown in Fig. 37.

We measure the calibration uncertainty and potential bias from template degeneracy, using the full-frequency simulated maps. To measure the calibration of each simulated map, we subtract the input CMB and foreground map, mask out 50% of the sky most affected by foreground residuals, and regress the input dipole template against the cleaned map. The resulting overall gain distributions are shown and compared to PR3 calibration in Table 7. The simulations indicate that we recover the input dipole at better than 1 μK (combined statistical and systematic uncertainty) in each of the CMB frequencies. The table also shows that there are small but statistically significant biases in the procedure, which cause the HFI CMB frequencies to be calibrated

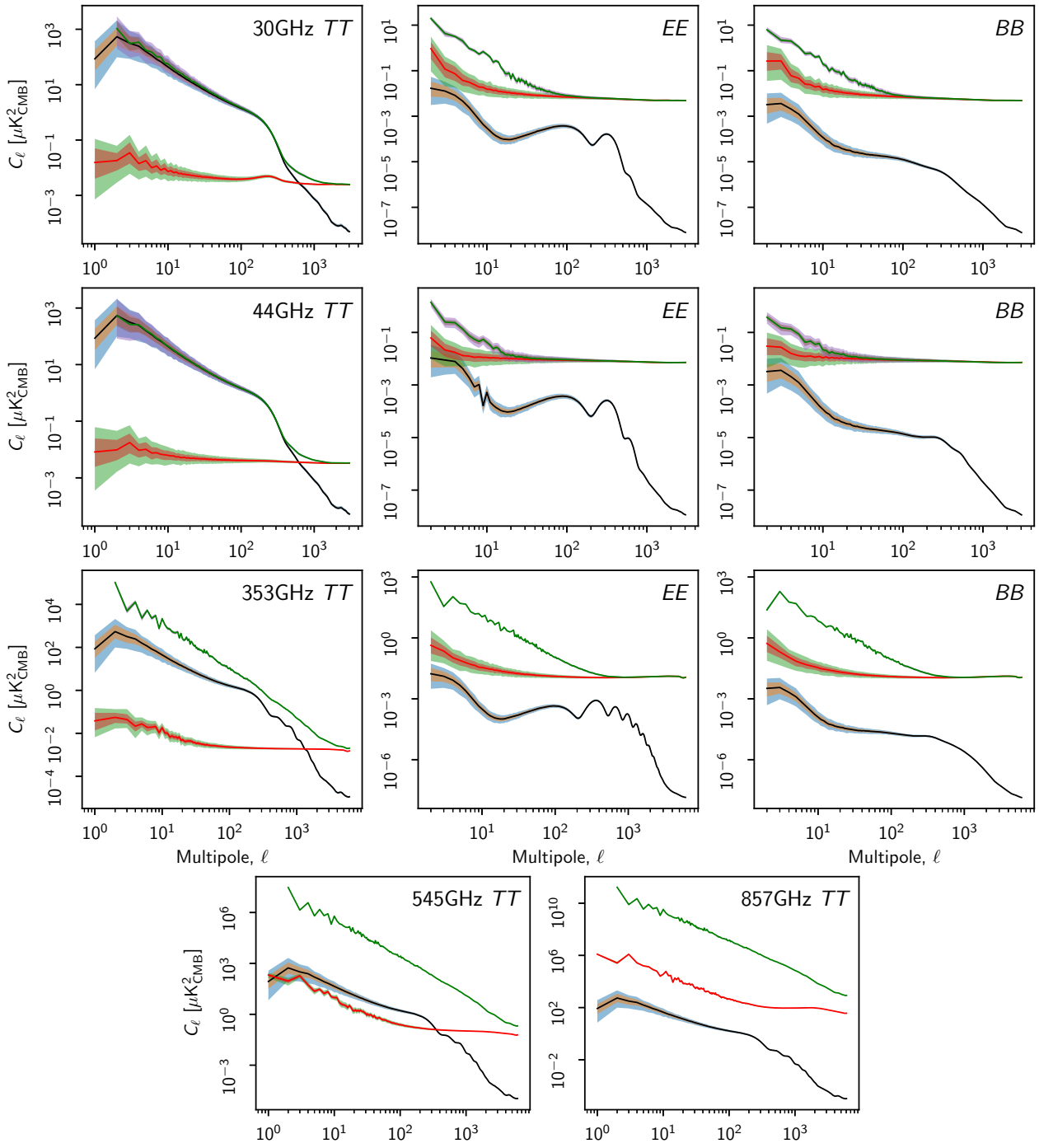


Fig. 32. Simulated CMB, noise, and systematics pseudo-spectra at 30, 44, 353, 545, and 857 GHz. The median CMB power in each bin is plotted in black, and the median noise and systematics power (difference between the input and output maps) in red. The coloured bands represent the asymmetric 68% and 95% confidence regions. The power spectra are binned into 300 logarithmically-spaced bins. The CMB is convolved with the beam and E-mode transfer function (Sect. 4.3). The large-scale CMB B-mode power in these plots follows from mode coupling and is not intrinsic to the simulations.

high by less than 0.03%, amplifying the Solar dipole amplitude by less than 1 μK.

6.2. Noise and systematics

We measure the level of residual noise and systematics in the *Planck* maps by splitting the available detectors into disjoint sets and performing the same processing on each detector set. The independent processing guarantees that the residuals between the detector sets are not correlated. Unlike making a split in the time domain (the so-called half-mission split), the detector-set difference does not cancel systematics that vary detector by detector, such as bandpass mismatch, far sidelobes, or the HFI transfer function residuals.

The definition of the detector-set null test has been somewhat ambiguous in past *Planck* publications. In an independent, local processing, the template corrections (calibration, transfer function residuals, etc.) and noise offsets are fitted to each detector set

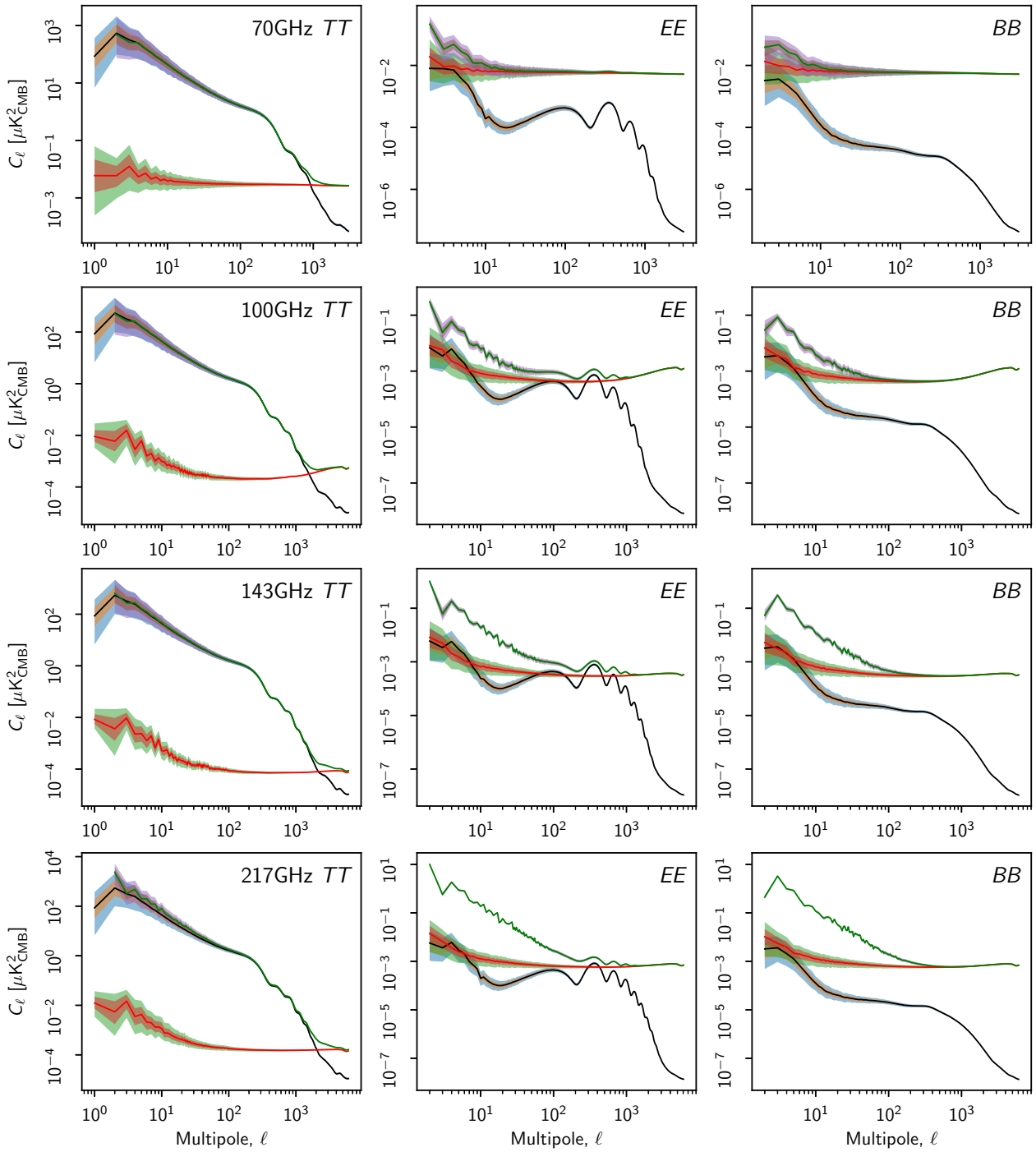


Fig. 33. Same as Fig. 32, for 70, 100, 143 and 217 GHz.

separately. Alternatively, it is possible to fit the templates globally to all of the data and project subsets of the cleaned data into detector-set and half mission maps. The global fit (sometimes referred to as "turbo destriping" because it only requires one run of the destriper) is suitable for measuring the achieved internal consistency in the cleaned TOD. However, residual systematics in turbo-destriped subsets are highly correlated, making it impossible to infer the total level of residuals from their differences. In this paper, all turbo-destriped subset maps are presented with an explicit mention of the global fit being used.

We compare the *EE* and *BB* total power in each polarized *Planck* frequency in Fig. 38. In Figs. 39 and 40 we study the detector-set null maps for the polarized HFI frequencies. These maps and spectra indicate that `NPIPE` products have less noise than PR3 maps. There are several factors that contribute. The `NPIPE` maps include data taken during repointing manoeuvres, which account for roughly 8% of the *Planck* mission. LFI noise levels are reduced further by the adaptive sky-load differencing (Sect. 2.3.6). HFI noise levels are improved by the extended flagging kernel during transfer-function deconvolution (Sect. 2.3.9). Large-scale HFI residuals, particularly in polarization, are significantly suppressed by our use of the polarized-foreground prior (Sect. 2.4.13) during calibration, as is demonstrated by the increase in `NPIPE` large-scale polarization uncertainty when

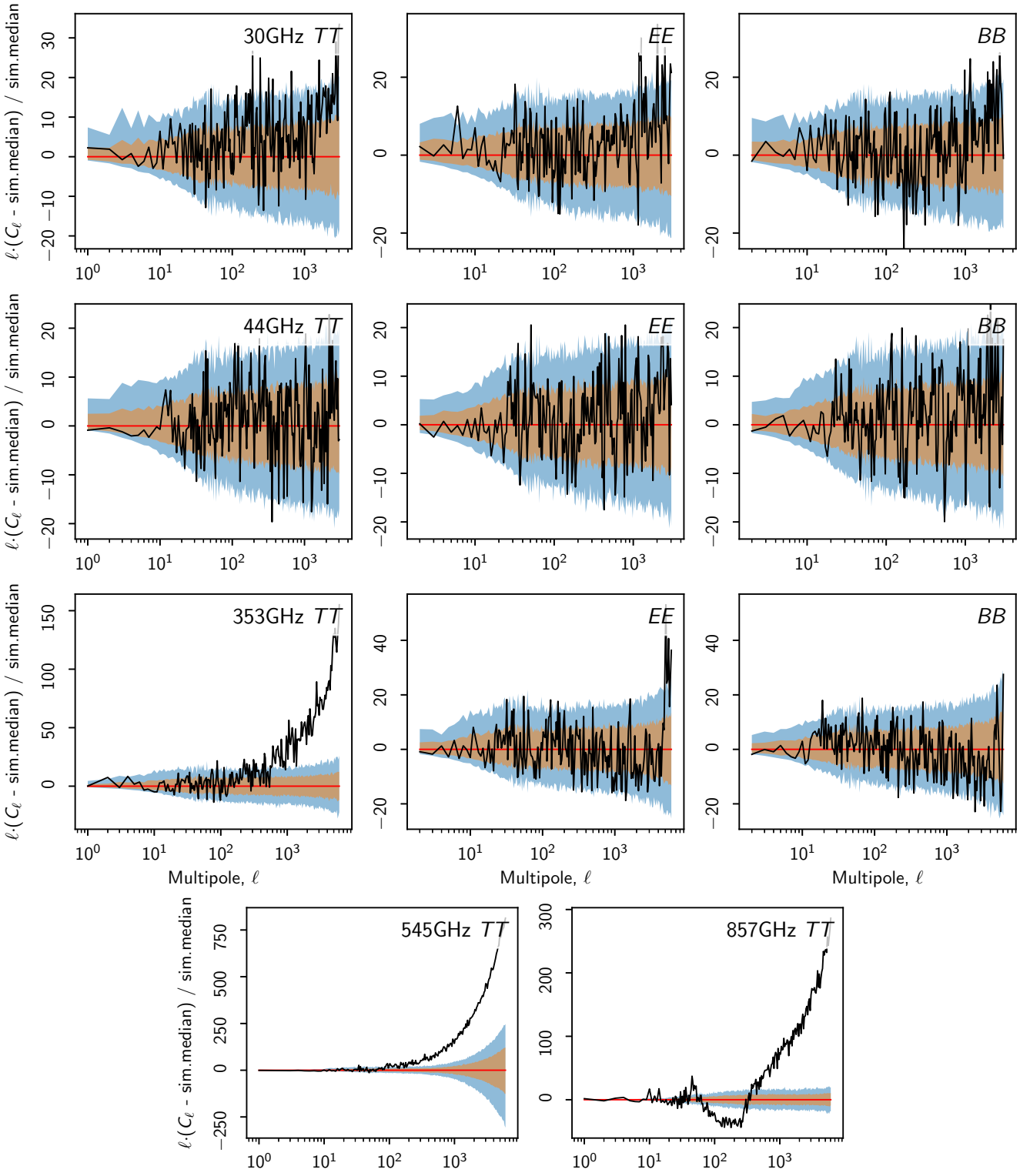


Fig. 34. Simulated A/B difference versus flight data at 30, 44, 353, 545, and 857 GHz. All spectra are differenced and divided by the median simulated spectrum. The flight data are shown in black, and the median simulated power in each bin is plotted in red. The coloured bands represent the asymmetric 68% and 95% confidence regions. The power spectra are binned into 300 logarithmically-spaced bins. The spectra shown here are the same as in Fig. 28. Notice the ℓ-scaling.

the prior is disabled (Fig. 41). These improvements hold even after deconvolution of the E-mode transfer function (Sect. 4.3). Finally, we fit the $1/f$ noise fluctuations with 167 ms baseline steps, compared to the 0.25 s (30 GHz), 1 s (44 and 70 GHz), and full-pointing-period (35–75 min at HFI frequencies) baseline steps used in PR3.

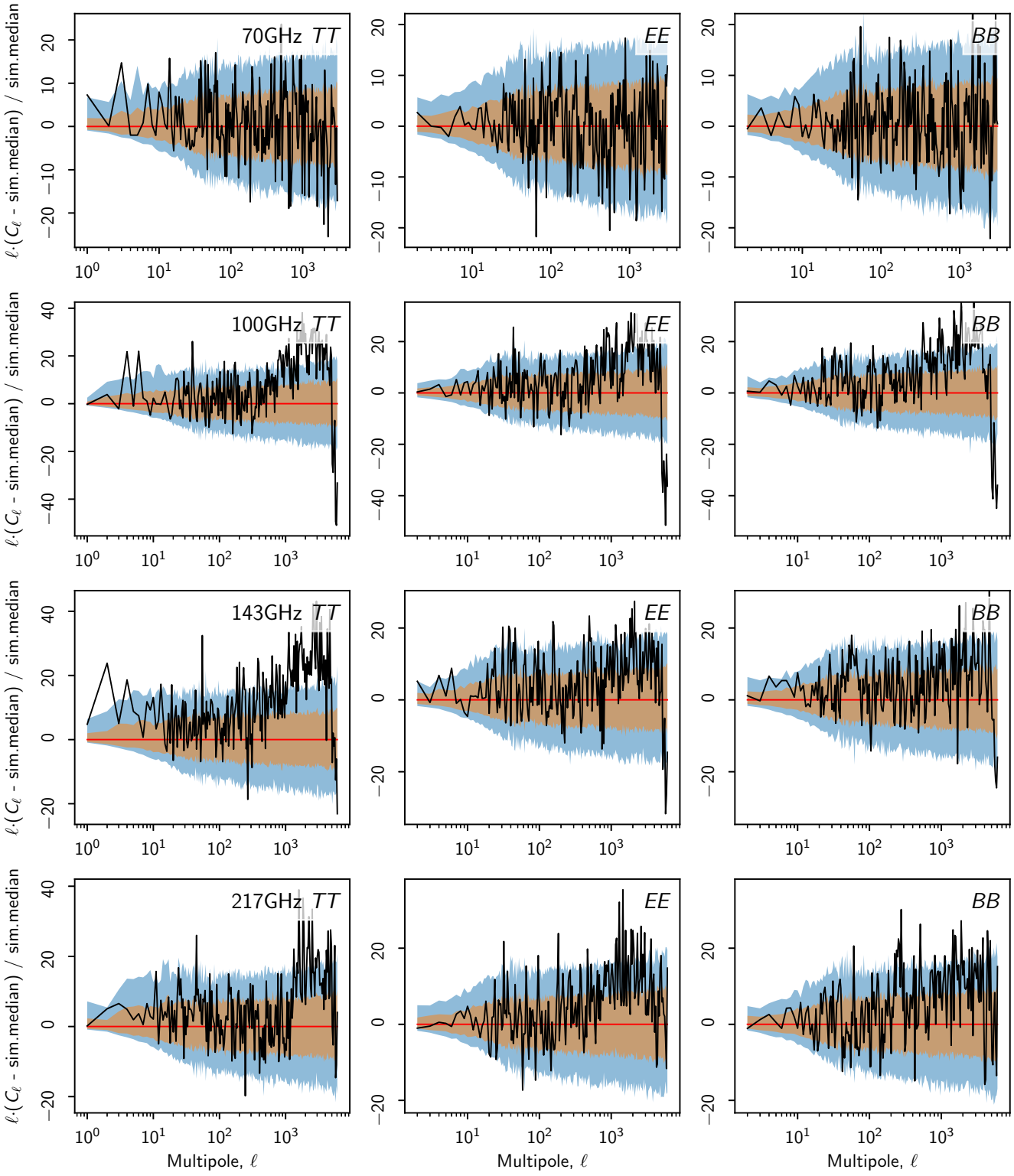


Fig. 35. Same as Fig. 34, but at 70, 100, 143, and 217 GHz. The spectra shown here are the same as in Fig. 29.

6.3. Signal

We present the temperature and polarization difference maps between NPIPE and PR3 in Fig. 42. To make the temperature map differences more informative, we have projected out the relativistic dipole, zodiacal emission, and an overall relative calibration mismatch. The 353 GHz temperature difference has a peculiar ringing pattern, understood to originate from the coarse-grained transfer function model used in the 2018 processing. For comparison, we show select NPIPE–2015 differences in Fig. 43.

While our work was being prepared for publication, a new processing of *Planck*-HFI polarization data, known as SRoll2, was published in Delouis et al. (2019). It features a new model of the ADCNL that does not rely on the linear gain fluctuation approximation. The new ADCNL model is able to capture the bulk of the residual ADCNL with considerable economy, thus

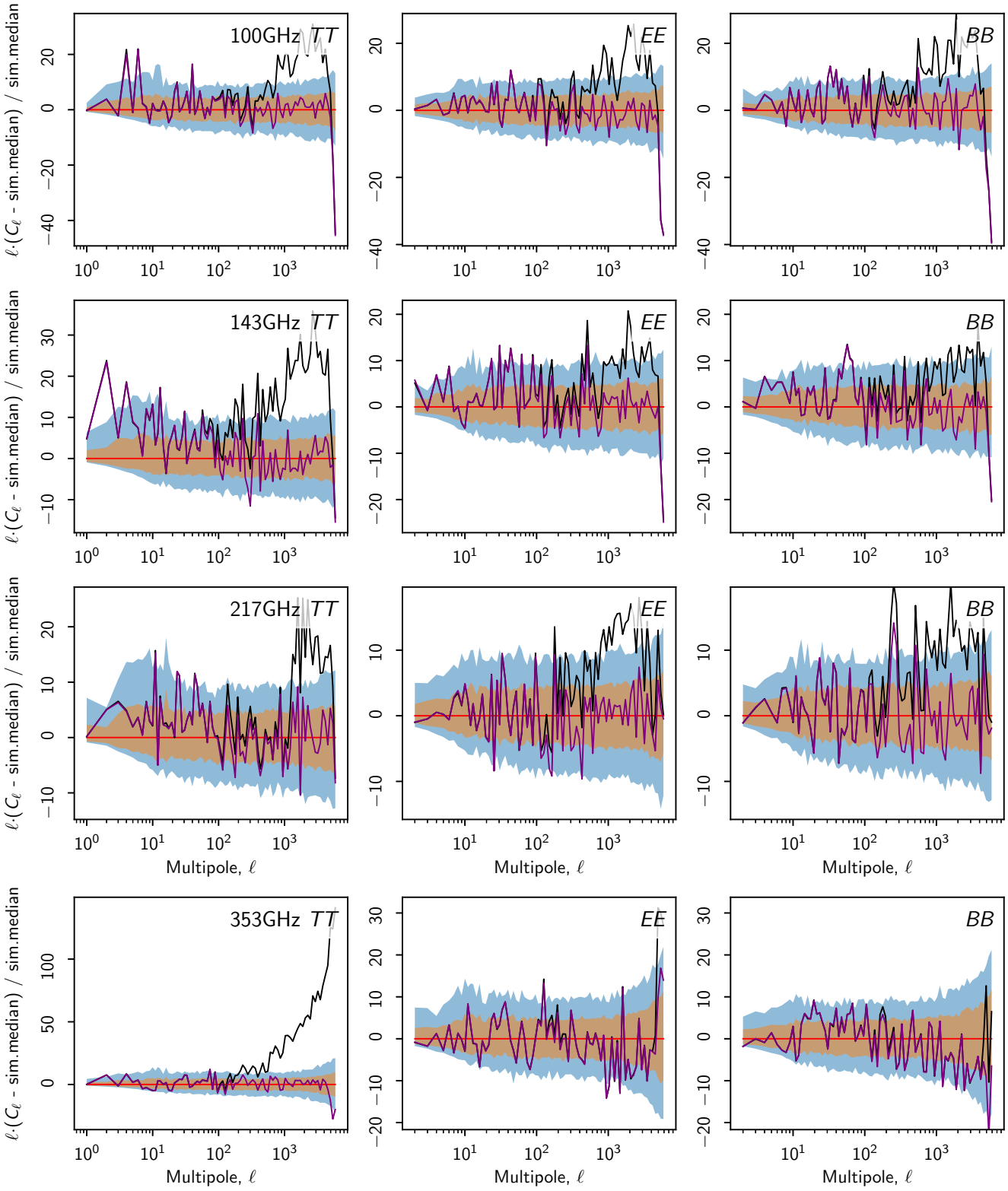


Fig. 36. Simulated A/B difference versus flight data at the HFI frequencies where we identify a percent-level, scale-dependent deficit in simulated noise power. The flight data power spectrum is differenced and divided by the median of the uncorrected simulations (black) and, alternatively, by the simulated maps that include the noise correction (purple). The coloured bands represent the asymmetric 68% and 95% confidence regions. The power spectra are binned into 100 logarithmically spaced bins. Apart from the wider binning, the black curves and the confidence limits are the same as those shown in Figs. 28 and 29.

limiting the degeneracy between the IQU map and the ADCNL correction. We show in Fig. 44 the NPIPE−SRoll2 difference maps, and observe that for polarization, the NPIPE and SRoll2 maps are considerably closer to each other than NPIPE and PR3, except at 143 GHz. But even the NPIPE−PR3 differences are very small. In temperature, maps from 100 to 217 GHz have

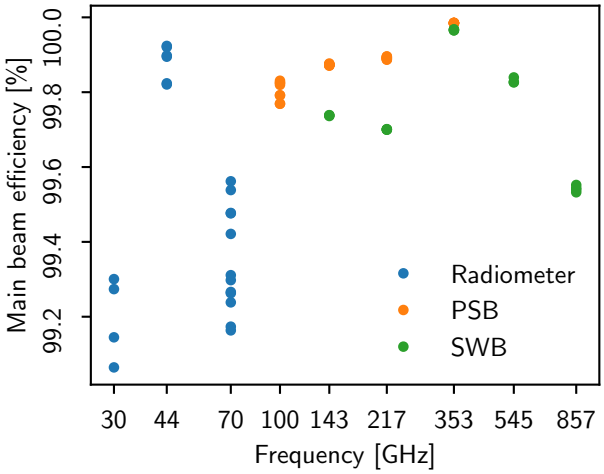


Fig. 37. Uncorrected main-beam efficiencies for all *Planck* detectors. The NPIPE calibration procedure corrects each frequency map to have full main-beam efficiency. The 545 and 857 GHz efficiencies are less certain due to uncertainties in main beam and sidelobe estimates, and are based on fitting the 353 GHz sidelobe templates to the data. The 100–353 GHz sidelobe estimates are based on first-order GRASP simulations, and overestimate the main-beam efficiency by a small but unknown amount.

Table 7. Relative calibration between NPIPE and PR3 maps and the NPIPE gain uncertainty estimated from recovery of the injected dipole in simulated full-frequency maps.

Frequency [GHz]	2018 gain [a] [%]	MC gain [b] [%]	MC bias [c] [%]
30	99.793	99.975 ± 0.045	−0.025 ± 0.005
44	99.763	99.990 ± 0.040	−0.010 ± 0.004
70	99.959	99.997 ± 0.030	−0.003 ± 0.003
100	100.061	100.016 ± 0.023	0.016 ± 0.002
143	100.082	100.016 ± 0.018	0.016 ± 0.002
217	100.005	100.024 ± 0.024	0.024 ± 0.002
353	100.463	100.052 ± 0.053	0.052 ± 0.005
545 [d]	115.123	99.586 ± 0.571	0.414 ± 0.057

Notes. [a]Relative calibration, NPIPE /2018, measured over 50% of the sky after adding the Solar dipole back into 2018 maps and smoothing the maps with a 1° Gaussian beam. [b]Distribution centre and width, measured relative to the input signal. [c]This is the average calibration error and the associated uncertainty. A 0.03% bias in calibration corresponds to about 1 μK_{CMB} in Solar dipole amplitude. [d]The NPIPE 545 GHz calibration is based on the orbital dipole, and does not suffer from the 10% modelling uncertainty affecting the planet-based calibration in the 2018 release. We also calibrate directly into K_{CMB}, without needing a unit conversion factor with additional uncertainty.

diverged in the Galactic plane, owing to the exclusion of all SWBs from SRoll2 processing. The 353 GHz zebra stripe pattern is gone for both NPIPE and SRoll2.

6.3.1. Internal consistency

Each *Planck* operational year divides into two sky surveys. The two surveys are complementary in the sense that most sky regions are scanned in approximately opposite directions between the two surveys. This makes the difference between odd and even surveys sensitive to a number of systematic effects, such as pointing error, far sidelobes, and transfer-function residuals. Processing the odd and even surveys into separate maps would be possible, but not representative of the full-map uncertainty, since the opposite scanning directions are essential in fitting the systematic templates. Instead, we can project the calibrated and template-corrected data into odd and even survey maps, and use their difference to gauge the level of internal consistency achieved. This test reveals if the considered templates have enough freedom to model the systematics that are responsible for the odd–even survey differences.

We compare LFI odd−even survey differences between PR2, PR3, and NPIPE in Fig. 45. This figure shows a clear improvement in internal consistency between the 2015 and 2018 maps, owing to the improvements in calibration. The NPIPE survey differences show further improvement, likely due to the full-frequency calibration scheme that corrects for integrated bandpass mismatches.

We compare the 100–217 GHz survey differences in Fig. 46. These figures contrast the apparent lack of large-scale structure in the NPIPE null map with the visible residuals in the earlier releases. The improvement is due to the polarization prior, short-baseline destriping, and additional bins in the transfer function correction. It is worth pointing out that the NPIPE 143 GHz map has a noticeable residual in the Galactic plane, possibly from the transfer-function correction that down-weights these pixels when fitting the templates. The level of the residual is negligible given the strength of the foreground signal in these pixels. The 217 GHz 2018 map outperforms the 2015 results due to improvements in FSL and zodiacal-light removal.

In Fig. 47 we extend the comparison from 353 to 857 GHz. The 353 GHz results are particularly notable because they show an apparent degradation in the null map between the 2015 and 2018 releases, visible in a pattern referred to as "zebra stripes" in Planck Collaboration III (2020). This is a result of insufficient granularity in the measured and applied transfer-function correction. The regression is corrected in NPIPE by increasing the number of spectral bins.

In Fig. 48 we consider polarization difference maps of the form $\boldsymbol{m}_{217}$ to 0.128 $\boldsymbol{m}_{353}$ as derived by SRoll, SRoll2, and NPIPE, where the scaling factor corresponds to a modified-blackbody spectral energy distribution (SED) ratio between 217 and 353 GHz evaluated for $\beta_d = 1.6$ and $T_d = 19.6$ K. These maps are thus designed to suppress thermal dust emission without having to resort to component-separation methods, and highlight potential residual systematic effects in the high-frequency channels. Focusing first on high Galactic latitudes, we first note that all three map solutions exhibit notable large-scale fluctuations. Some fluctuations are expected simply from true CMB fluctuations (which are only suppressed by 12.8% in these difference maps), and some fluctuations are expected from random statistical noise. However, some fluctuations are also likely due to residual instrumental systematics. For comparison, we note that a standard ΛCDM CMB sky with $\tau \approx 0.06$ exhibits peak-to-peak fluctuations of about 0.5 μK, with an overall standard deviation of 0.15 μK. At low Galactic latitudes, the residuals are dominated by possible spatial variations in the spectral index not captured by the constant scaling factor of 0.128 and temperature-to-polarization leakage. We note that NPIPE exhibits notably lower fluctuations at high latitudes than both SRoll and SRoll2, and Galactic-plane residuals appear more strongly correlated with the morphology of thermal dust.

In Fig. 49 we show corresponding angular power spectra, evaluated as cross-spectra between detector-split difference maps (for NPIPE) and half-mission split maps (for SRoll and SRoll2) outside the *Planck* 2018 polarization analysis mask (Planck Collaboration IV 2020). Consistent with the above

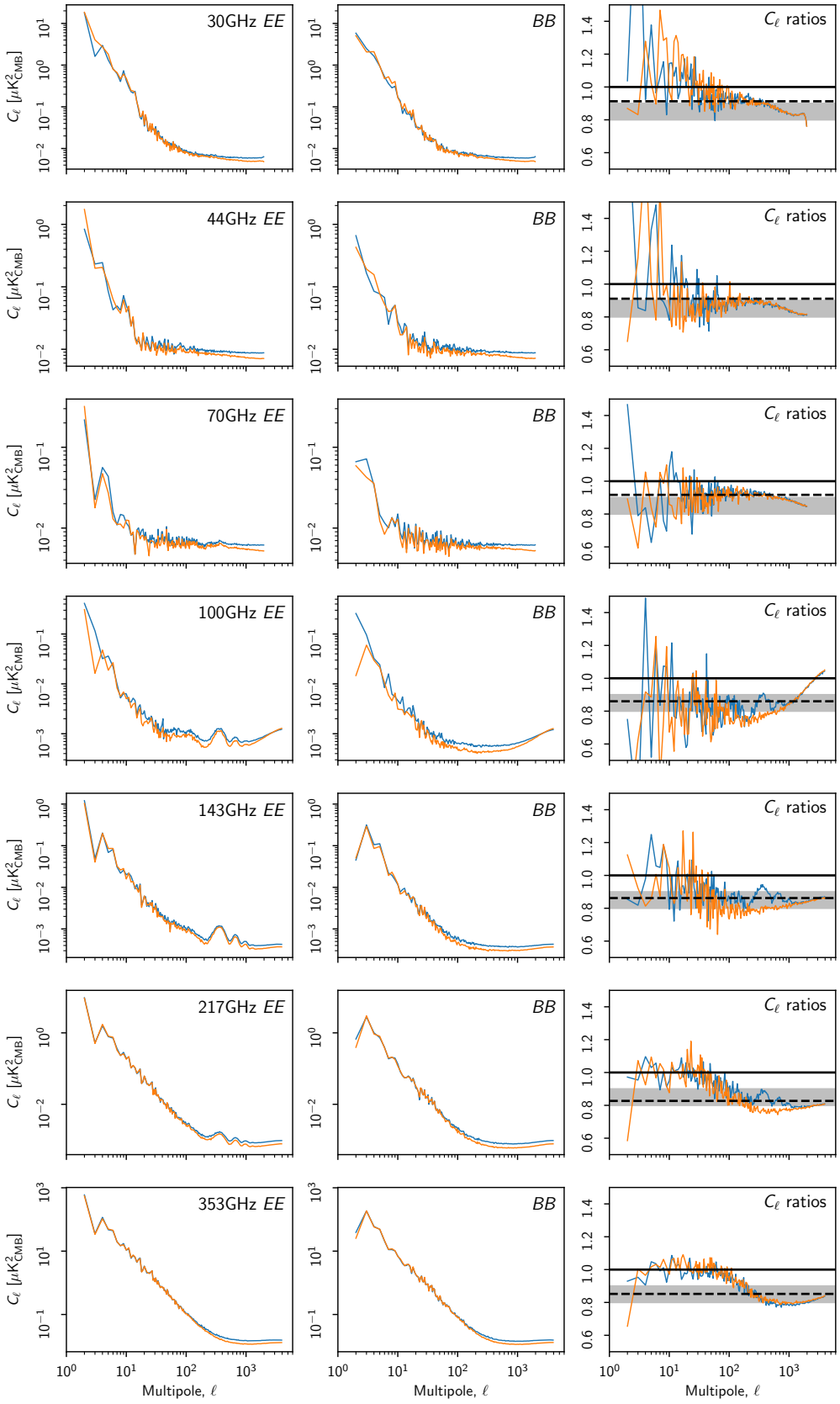


Fig. 38. *EE* and *BB* auto-spectra of the polarized frequency maps. The *first two columns* show PR3 (blue) and NPIPE (orange) auto-spectra, while the *third column* shows the ratios (NPIPE/2018) with *EE* in blue and *BB* in orange. For noise-dominated angular scales, NPIPE maps have 10–20% lower noise variance, indicated by the grey band in the ratio plot. We show a naive estimate of the ratios, based on the ratio of masked pixel hits in Table 1, as a dashed black line. These spectra are computed over 50.4% of the sky, corrected for the sky fraction and binned into 300 logarithmically-spaced bins.

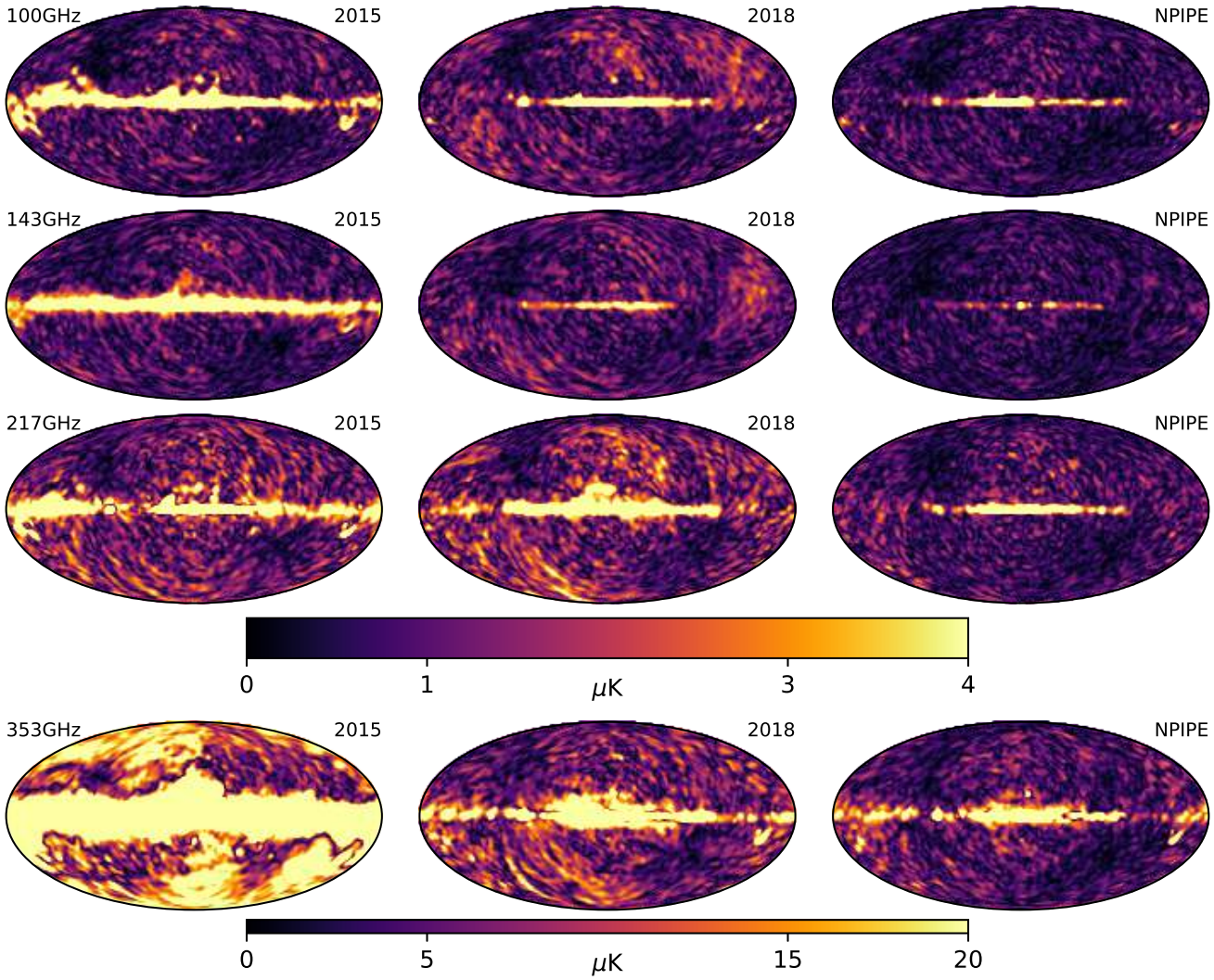


Fig. 39. Polarization amplitudes of the detector-set difference null maps. The angular power spectra of PR3 and NPIPE maps are shown in Fig. 40. Independent processing of the two detector-sets means that these maps reflect the level of total residuals in the frequency maps.

visual impression, we see that NPIPE exhibits significantly lower *EE* power at large angular scales, while at small scales all three maps appear broadly consistent.

The observed differences between *Planck* polarization maps illustrated in Figs. 48 and 49 may reflect elements specific to the NPIPE data processing. As discussed in Sect. 4.3.2, polarized Galactic emission is not affected by the same transfer function at low multipoles as the CMB, because polarization templates are part of the data model in Eq. (9). However, this approach introduces an interdependence between the NPIPE frequency maps, which may impact studies of Galactic polarization (see Appendix H). Notably, this is an issue to have in mind when using the NPIPE frequency maps to characterize the frequency correlation of dust polarization (Planck Collaboration XI 2019), an essential question in the search for primordial CMB *B*-modes.

The bandpass mismatch template used in NPIPE is based on the same sky model as the simulated sky signal, and (as discussed in Sect. 5.2) we do not simulate errors in the template itself. The net effect is that the uncertainty in the polarized emission by foreground Galactic dust is underestimated in the simulations, by the amount that errors in the sky model affect the bandpass mismatch correction. Since the magnitude of this uncertainty is unknown, it may be misleading to rely on simulations alone to assess uncertainties in polarized Galactic emission on large scales, or to investigate how Galactic polarization decorrelates with frequency. NPIPE is not unique in its approach of not fully sampling the space of template errors. Issues related to bandpass mismatch are a generic feature in *Planck* polarized mapmaking, and caution should be exercised when analysing other Planck releases as well.

6.3.2. External consistency

Planck is calibrated without reference to WMAP (Bennett et al. 2013) polarization, although in NPIPE, WMAP temperature data do contribute to the sky model used to derive bandpass-mismatch templates. It is informative to compare the two experiments for agreement in synchrotron polarization. The comparison is particularly interesting due to differences in the two experiments. WMAP's differencing-assembly design allows for gain and bandpass mismatch to be separated into special spurious maps, while *Planck* requires an estimate of the foreground intensity to correct for bandpass mismatch.

We measure the agreement by smoothing the *Planck* and WMAP *K*-band (23 GHz) maps with a 5° Gaussian beam and then regressing the WMAP *K*-band and *Planck* 353 GHz maps

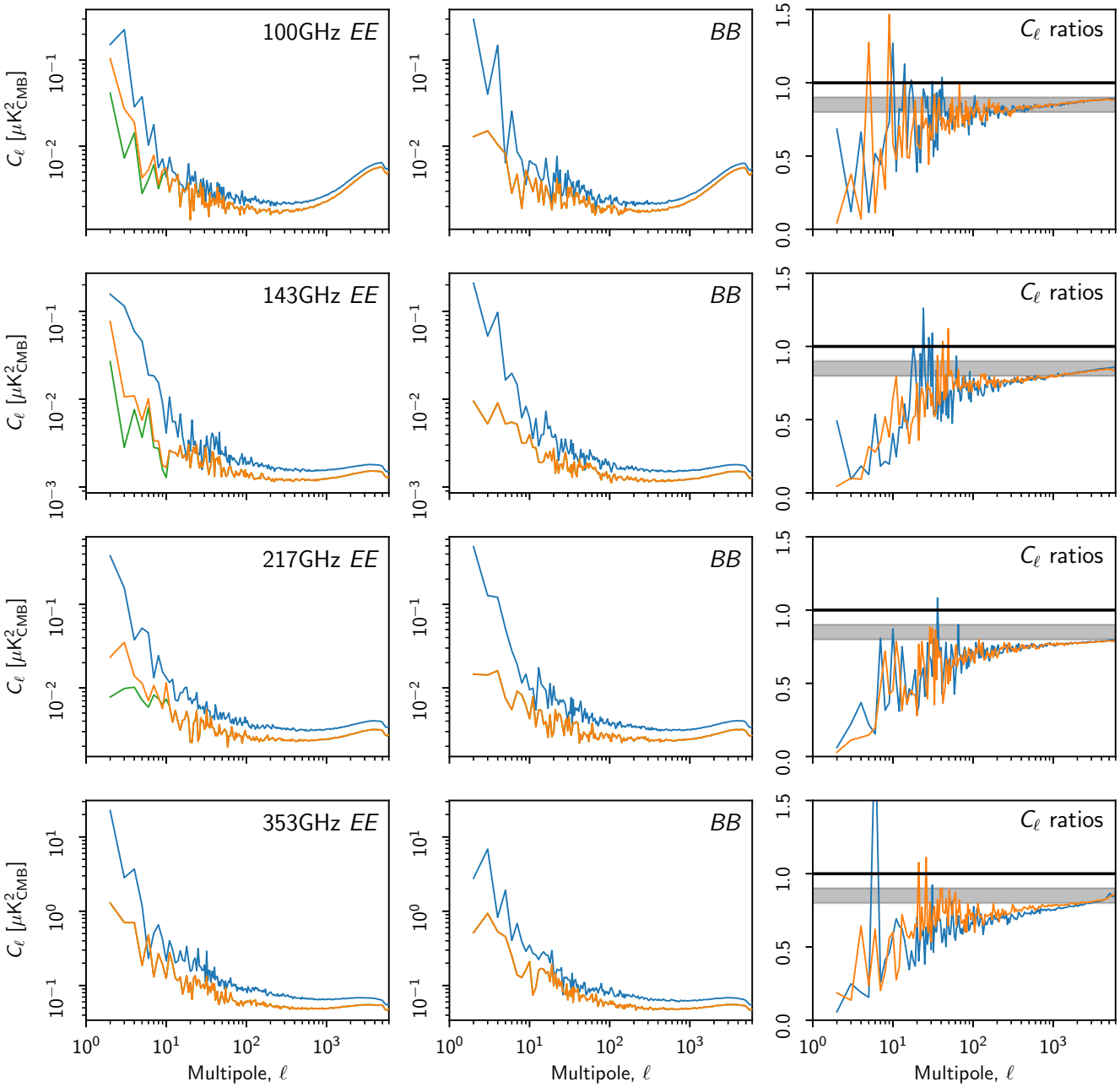


Fig. 40. *EE* and *BB* detector-set difference power spectra. The *first two columns* show PR3 (blue), raw NPIPE (green), and transfer-function-corrected NPIPE (orange) null-map power spectra. Note that PR3 detector sets are *not* the same as were differenced for Fig. 14 in Planck Collaboration III (2020), but rather ones that were destriped independently. The *third column* of panels shows the transfer-function-corrected NPIPE/2018 *EE* and *BB* ratios in blue and orange, respectively. NPIPE has notably less power at all angular scales. The grey band in the third column indicates a 10–20% improvement in power. These spectra are computed over 50.4% of the sky, corrected for the sky fraction and binned into 300 logarithmically-spaced bins. The polarization amplitudes of 2015, 2018, and NPIPE detector-set difference maps are shown in Fig. 39.

from the *Planck*-LFI maps. The residuals for 30 GHz are shown in Fig. 50, and for 44 GHz in Fig. 51. These figures show that improvements in the LFI calibration procedure between the 2015 and 2018 releases significantly improved the agreement between *Planck* and WMAP. They also show that the NPIPE large-scale polarization is more compatible with WMAP than is that of PR3. The fitting was carried out on the 30% of the sky that has the highest polarization amplitude in the smoothed *K*-band map, and we masked out a further 5% of the sky with the highest foreground intensity to reduce bandpass-mismatch-leakage effects, leaving 25.7% of the sky for fitting.

7. Component separation

We now derive new astrophysical component maps from the NPIPE data. For CMB extraction, we employ two different algorithms, Commander (Eriksen et al. 2004, 2008; Seljebotn et al. 2019) and SEVEM (Leach et al. 2008; Fernández-Cobos et al. 2012); for foreground estimation, we only use Commander. Both methods have been used extensively in previous *Planck* publications; for full details see Planck Collaboration IV (2020) and references therein. The main motivation for the present analysis is to characterize the internal consistency of the NPIPE

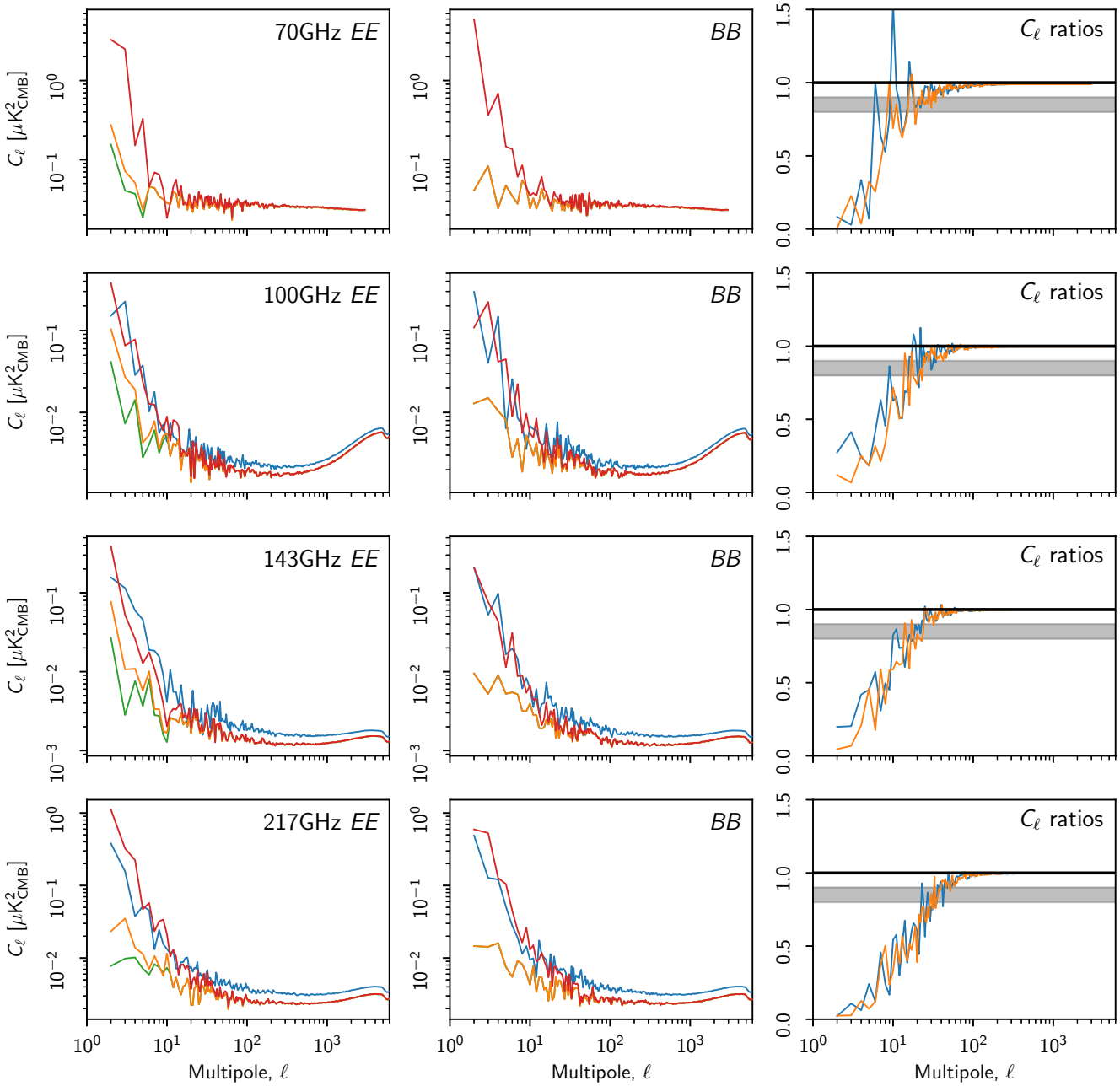


Fig. 41. Effect of the polarization prior on *EE* and *BB* detector-set difference power spectra. For 100–217 GHz, the green, orange, and blue lines on the *EE* and *BB* plots are the same as in Fig. 40 but here we add a red line, showing the power spectra for an alternative version of the NPIPE detector-set maps that are computed *without* the polarization prior. There is no blue line at 70 GHz because there is no comparable detector-set split in PR3, and 353 GHz is not shown because it is always calibrated without the polarization prior.

data themselves, rather than to derive an ultimate NPIPE sky model. In particular, external data sets are not considered in this paper. Instead, combined analyses are deferred to future studies.

7.1. Commander methodology

As described by Eriksen et al. (2008), Commander adopts a parametric approach to component separation. The first step in the process is to specify an explicit parametric model for the data. Since the current analysis considers only *Planck* data, we adopt a model similar to the one employed for the *Planck* 2018 analysis (Planck Collaboration IV 2020). In temperature, this includes CMB, a power-law low-frequency component, and a single modified-blackbody thermal-dust component. Since NPIPE produces single-detector temperature maps, it can also support individual CO components at all relevant frequencies (i.e., 100, 217, and 353 GHz). This is in contrast to PR3, which includes only integrated full-frequency maps. We therefore model CO emission using three independent components, with single-detector line ratios taking into account the relative bandpass differences between detectors.

One important variation with respect to previous *Planck* releases is the fact that the NPIPE maps retain the CMB Solar dipole. This allows for joint component separation and high-precision relative calibration, eliminates the need to model

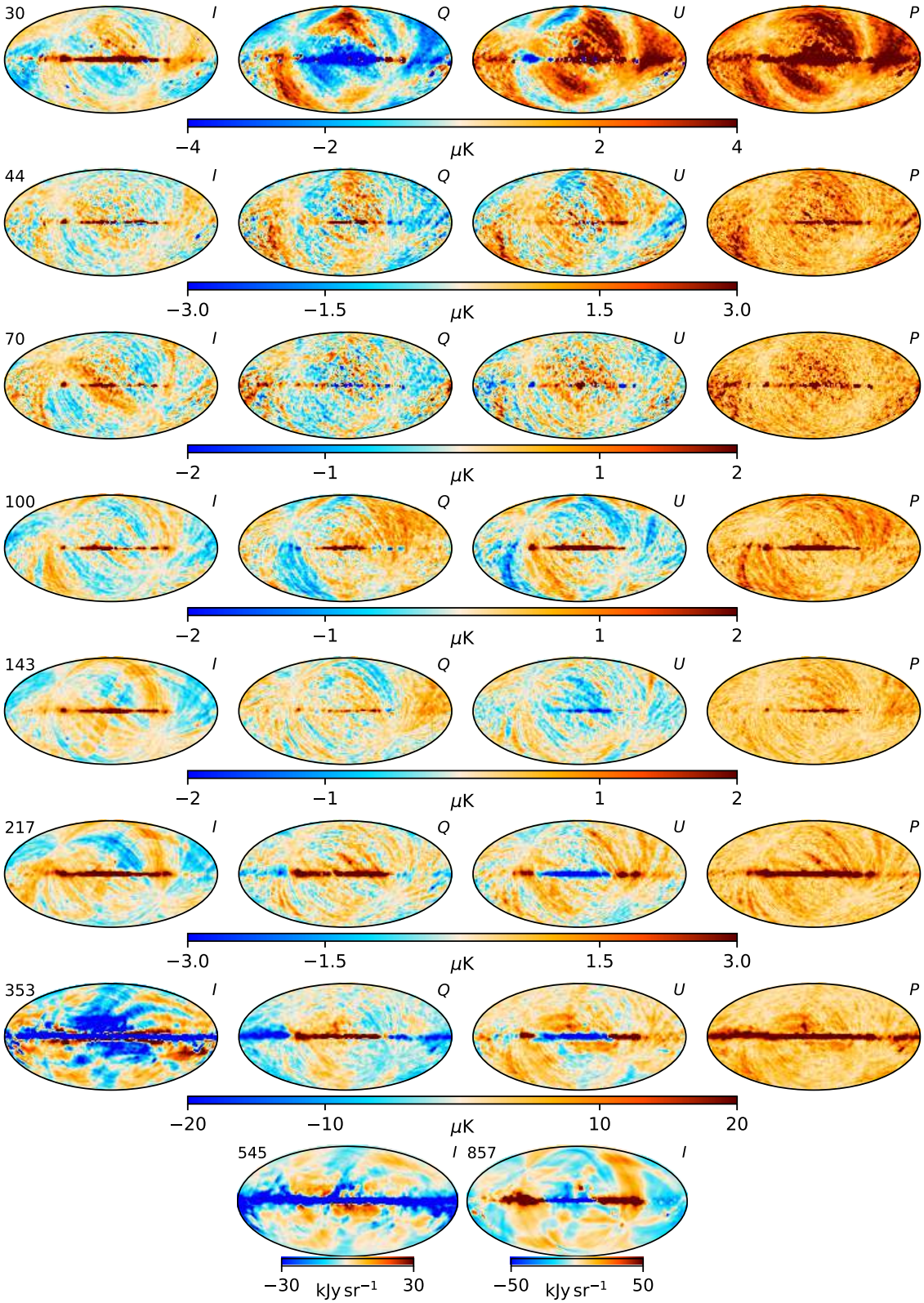


Fig. 42. NPIPE−2018 release difference maps in temperature and polarization. We have projected out the Solar dipole and zodiacal emission templates from the temperature differences, and performed a relative calibration using 50% of the sky to highlight differences beyond these trivial mismatch modes. All maps are smoothed with a 3° Gaussian beam to suppress small-scale noise.

additive dipole components in the frequency maps (Wehus et al. 2014; Planck Collaboration X 2016; Planck Collaboration IV 2020), and removes important large-scale degeneracies from the model. As a consequence, only three global parameters are fitted per sky map in the following analysis, namely a monopole/zero-level, an absolute calibration, and an overall bandpass correction (Planck Collaboration X 2016), compared to six parameters in previous analyses.

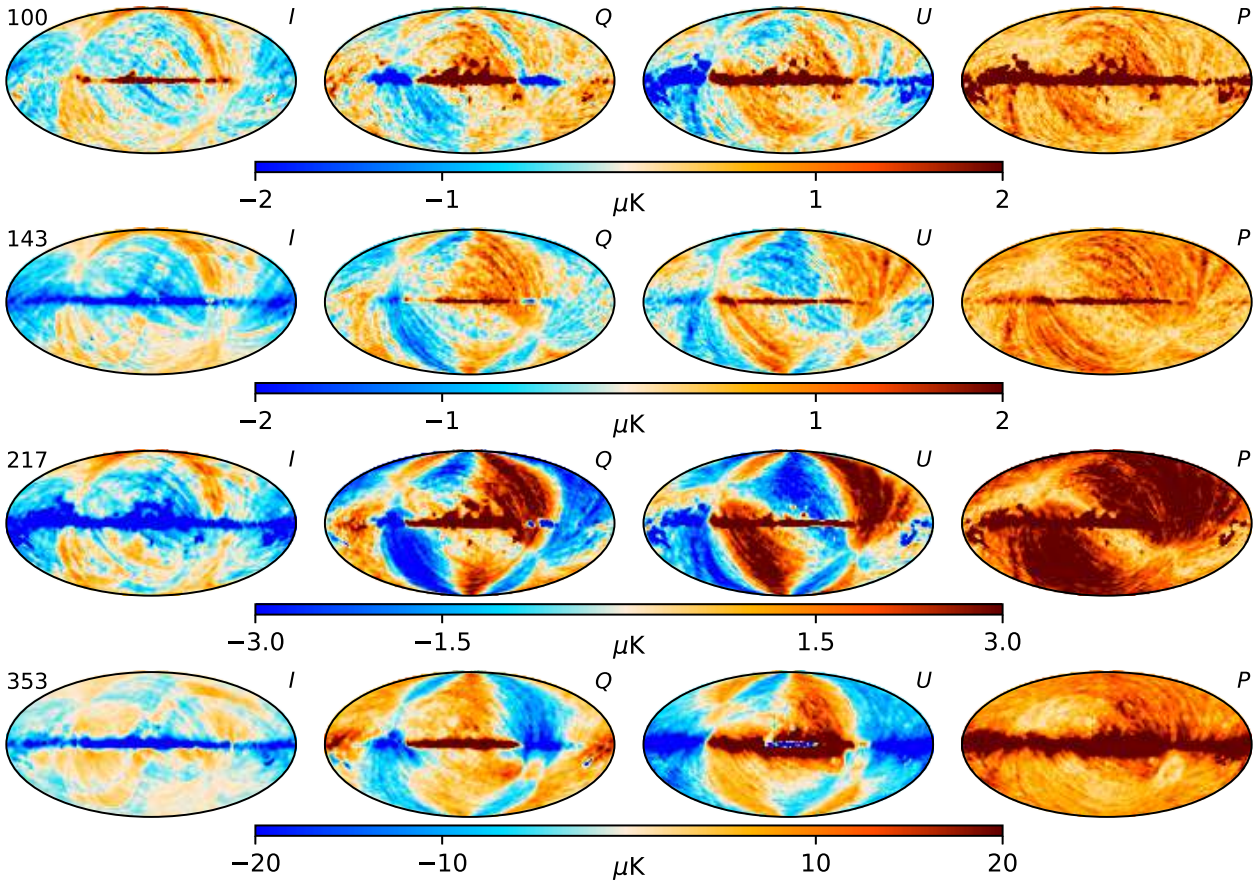


Fig. 43. NPIPE−2015 release difference maps in temperature and polarization to compare to Fig. 42. Note that the Solar dipole model for PR2 did not include the frequency-dependent part of the quadrupole term (Table C.1), so we also omit that correction from the dipole template here. The 353 GHz difference shows that the "zebra" pattern is exclusive to the 2018 temperature map. The polarization differences are indicative of an overall calibration mismatch between polarization-sensitive bolometers, causing substantial temperature-to-polarization leakage from the Solar dipole in the 2015 maps.

The temperature model employed in this paper may be written as a vector in pixel space:

$$\begin{aligned} \boldsymbol{s}_\nu = u_\nu\, g_\nu\, \mathsf{B}_\nu \Bigg[& \boldsymbol{a}_{\rm cmb}\gamma(\nu) \\ & + \boldsymbol{a}_{\rm lf}\left(\frac{\nu}{\nu_{\rm lf}}\right)^{\beta_{\rm lf}(p)} \\ & + \boldsymbol{a}_{\rm d}\left(\frac{\nu}{\nu_{\rm d}}\right)^{\beta_{\rm d}(p)+1}\left(\frac{e^{h\nu_{\rm d}/kT_{\rm d}(p)}-1}{e^{h\nu/kT_{\rm d}(p)}-1}\right) \\ & + \boldsymbol{a}_{\rm co}^{100}h_\nu^{100} + \boldsymbol{a}_{\rm co}^{217}h_\nu^{217} + \boldsymbol{a}_{\rm co}^{353}h_\nu^{353}\Bigg] \\ & + m_\nu, \end{aligned} \tag{19}$$

where ν denotes channel and p denotes pixel number. The terms correspond, from top to bottom, to the cosmological CMB signal, a combined power-law low-frequency component with reference frequency $\nu_{\rm lf}$, a single modified-blackbody thermal-dust component with a references frequency $\nu_{\rm d}$, three CO components ($J=1\rightarrow0$, $J=2\rightarrow1$, and $J=3\rightarrow2$), and, finally, a monopole "offset". The pre-factors before the astrophysical components are, from right to left: a per-channel beam-convolution operator B_ν, accounting for the azimuthally symmetric component of the true beam; an overall relative calibration factor g_ν per channel, with the 143 GHz calibration fixed to unity; and a unit conversion factor u_ν, scaling from brightness to thermodynamic temperature.

Each component is defined in terms of an amplitude map $\boldsymbol{a}_i$ at some reference frequency, multiplied by an SED that translates this amplitude map to any other frequency. All SEDs are defined in brightness temperature units. For the CMB component, $\gamma(\nu) = x^2e^x/(e^x-1)^2$ is therefore simply the conversion factor from thermodynamic to brightness temperature, where $x = h\nu/k_{\rm B}T_0$, h is Planck's constant, $k_{\rm B}$ is Boltzmann's constant, and T_0 is the mean CMB temperature. For the low-frequency component, the SED is given as a power law with a free spectral index, $\beta_{\rm lf}$. For thermal dust emission, the SED has two free parameters, a spectral index $\beta_{\rm d}$ and a temperature $T_{\rm d}$. Finally, the multiplicative CO line ratios h_j^i are set to zero for sky maps outside the bandpass range of a given CO component, and set to unity for one arbitrarily-chosen reference sky map inside the bandpass range of the same component (i.e., for the map from one detector at the relevant frequency); all other line ratios for the same component are fitted freely. For instance, h_j^{217} is set to unity for the 217-2 channel, fitted freely for all other 217 GHz channels, and set to zero for all non-217 GHz channels.

For notational simplicity, all unit-conversion and bandpass-integration effects are suppressed in the above model. However, all SEDs are integrated over the full bandpass $\tau(\nu)$ of each detector, using the formulae described in Planck Collaboration V (2014) and Planck Collaboration IX (2014), ensuring that the

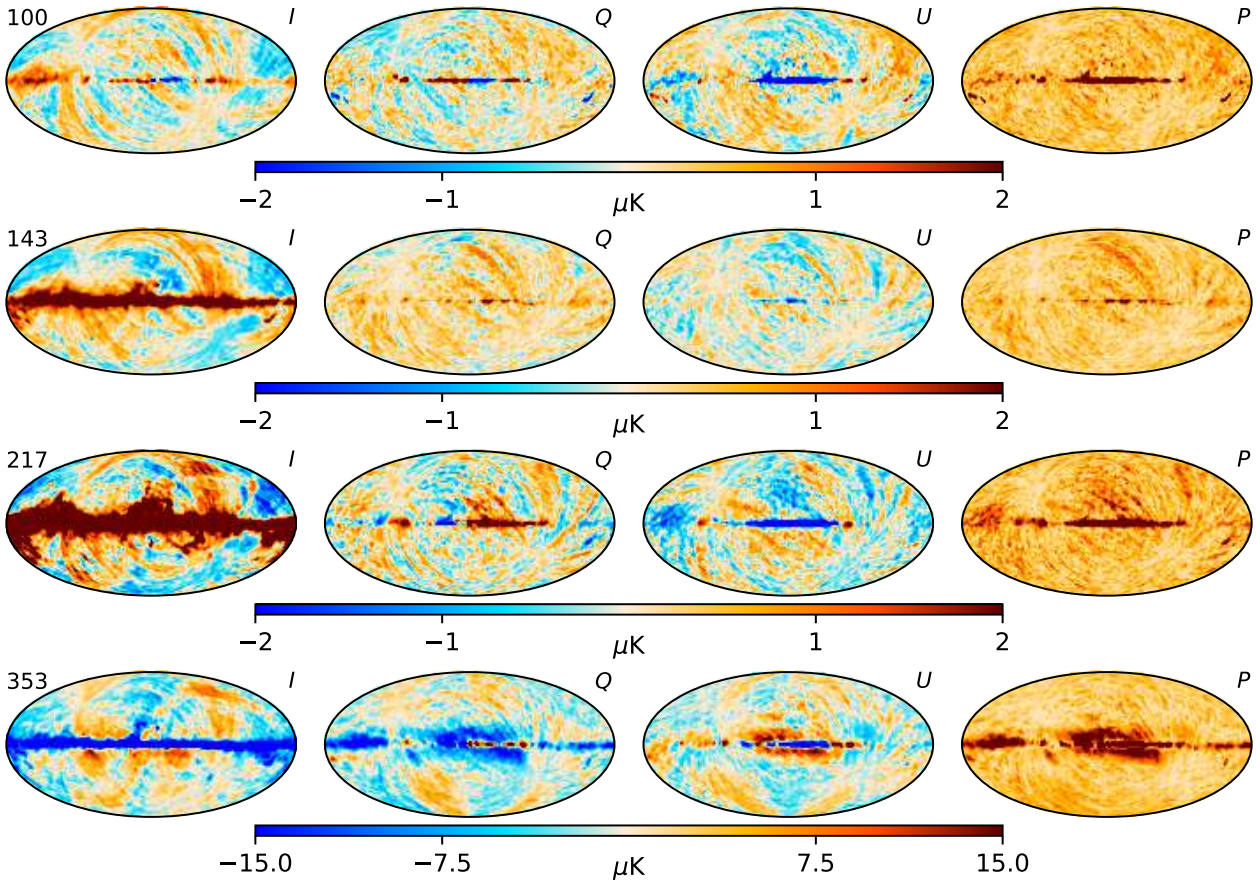


Fig. 44. NPIPE−SRoll2 difference maps in temperature and polarization to compare to Fig. 42.

amplitude maps correspond directly to the signal that would be observed at a sharp reference frequency, not as bandpass-integrated quantities. To account (at least partially) for the uncertainties in the bandpasses measured on the ground Planck Collaboration X (2016), we allow for an overall additive shift, Δ_ν, in each bandpass, such that $\tau(\nu) \rightarrow \tau(\nu + \Delta_\nu)$.

For polarization, we adopt a similar, but simpler, model:

$$\begin{aligned} \boldsymbol{s}_\nu = u_\nu\, g_\nu\, \mathsf{B}_\nu \Bigg[& \boldsymbol{a}_{\rm cmb}\gamma(\nu) \\ & + \boldsymbol{a}_{\rm s} \left(\frac{\nu}{\nu_{\rm s}}\right)^{\beta_{\rm s}(p)} \\ & + \boldsymbol{a}_{\rm d} \left(\frac{\nu}{\nu_{\rm d}}\right)^{\beta_{\rm d}(p)+1} \left(\frac{e^{h\nu_{\rm d}/kT_{\rm d}(p)} - 1}{e^{h\nu/kT_{\rm d}(p)} - 1}\right) \Bigg]. \end{aligned} \tag{20}$$

No CO components or monopoles are fitted for polarization, and the general low-frequency component has been renamed to just "synchrotron", since neither free-free nor spinning dust emission are expected to be significantly polarized (Planck Collaboration XXV 2016). Additionally, the thermal dust temperature $T_{\rm d}$ is fixed to the values derived from the temperature analysis, since *Planck* is not sensitive to polarization above 353 GHz. Note, however, that $T_{\rm d}$ is spatially varying even though it is fixed throughout the analysis. In contrast, the two spectral indices, $\beta_{\rm s}$ and $\beta_{\rm d}$, are fitted directly with polarization data.

All free parameters in these models are fitted jointly using the Commander software suite, following the same procedures as detailed in Planck Collaboration X (2016) and Planck Collaboration IV (2020). The starting point of these fits is (as usual) Bayes theorem,

$$P(\theta|\boldsymbol{d}) = \frac{P(\boldsymbol{d}|\theta)P(\theta)}{P(\boldsymbol{d})}, \tag{21}$$

where $P(\theta|\boldsymbol{d})$ is the posterior distribution, $P(\boldsymbol{d}|\theta) = \mathcal{L}(\theta)$ is the likelihood, $P(\theta)$ is a set of priors, and $P(\boldsymbol{d})$ is, for our purposes, an irrelevant normalization constant. In most of the following analysis, we report the maximum-posterior solution as our best-fit estimate, and we use corresponding simulations to quantity uncertainties.

We adopt the same set of priors as in Planck Collaboration IV (2020). Most importantly, for SED parameters such as β_i or $T_{\rm d}$, we adopt a product of informative Gaussian priors, with central values informed by the high signal-to-noise parts of the data set, and a non-informative Jeffreys prior to account for posterior volume effects due to nonlinear parameterizations. For spatial priors, we adopt angular-power-law-spectrum priors on all foreground amplitudes, but no spatial prior on the CMB amplitudes (for further details, see Appendix A in Planck Collaboration IV 2020). For global parameters, we impose no active priors on either calibration factors or offsets (g_ν and m_ν, respectively, in Eqs. (19)–(23)), except that we fix one overall calibration factor (at 143 GHz), and one offset per free diffuse component (30 GHz for the low-frequency-component; 100-1a for CO $J = 1 \rightarrow 0$; 143 GHz for CMB; 217-2 for CO $J = 2 \rightarrow 1$; 353-3 for $J = 3 \rightarrow 2$; and 545-2 for thermal dust emission.) The reference calibration factor is fixed to unity, while the reference offsets are determined such that the resulting component maps obtain physically-meaningful zero-levels. Specifically,

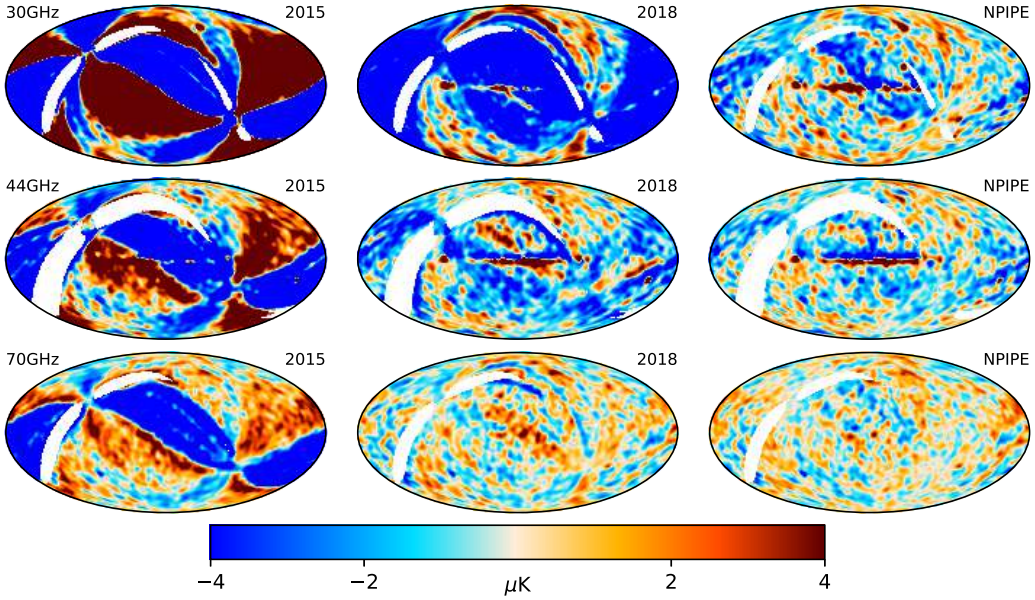


Fig. 45. Odd–even survey intensity differences for LFI smoothed to 5°. These maps reflect the internal consistency achieved, not the total residuals, since the calibration errors are correlated between the surveys. To match the 2015 and 2018 processing, the NPIPE 167 ms baseline offsets for this plot are solved using individual survey data.

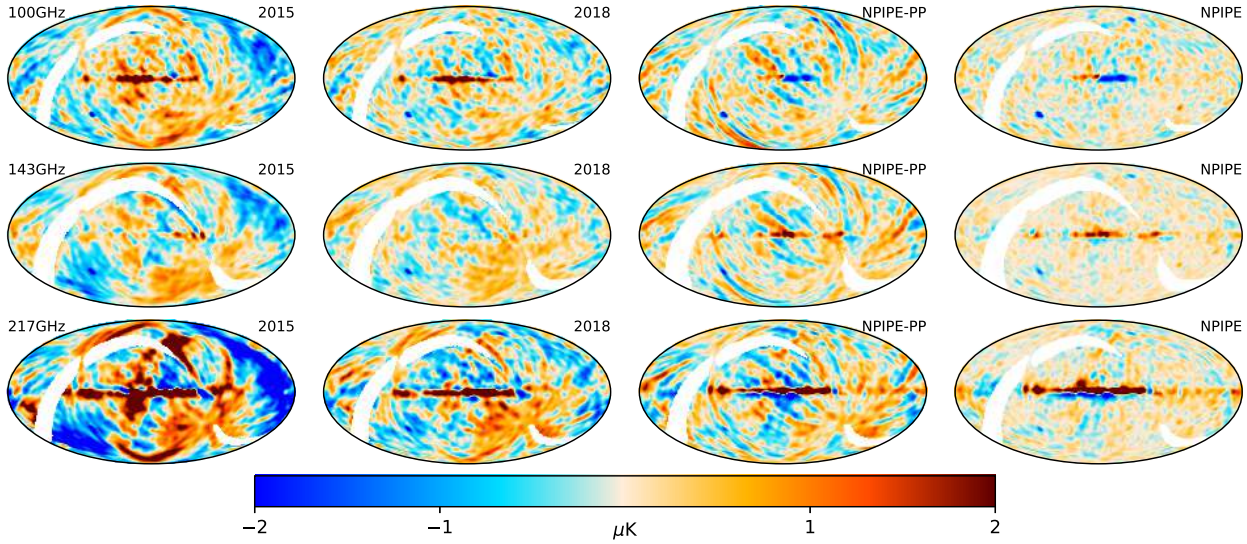


Fig. 46. Odd–even survey intensity differences for 100, 143, and 217 GHz, smoothed to 5°. The 2015 and 2018 maps are the same as in Fig. 12 of Planck Collaboration III (2020). The NPIPE-PP column shows the difference obtained if NPIPE is solved only for pointing-period offsets (like PR3), rather than for the 167 ms baseline offsets. The stripes visible in the NPIPE-PP maps are glitch and ADC nonlinearity-correction residuals that are well captured by the short-baseline solution. The comparison is not perfect, because PR2 (2015) baseline offsets were solved using the survey TOD, while the other versions use full-mission baselines. Three variable radio sources can be identified across the frequencies in the NPIPE maps. These maps reflect the internal consistency achieved, not the total residuals, as the calibration errors are correlated between the surveys.

for 100-1a, 143, 217-2, and 353-3, the reference offsets are set such that the CO and CMB maps have vanishing mean offsets at high Galactic latitudes. For 545-2, the offset is set such that a T–T scatter plot between the derived thermal dust map and an H I survey (Lenz et al. 2017) has vanishing intercept. For 30 GHz, it is set such that the derived low-frequency spectral index map does not correlate strongly with the corresponding amplitude map, which is a typical artefact resulting from incorrectly-set offsets (Wehus et al. 2014). The 30 GHz offset determination is the most uncertain among these offsets, since the spectral index information in the map has relatively low S/N.

We approximate the instrumental noise of each detector as Gaussian, and the likelihood therefore is

$$P(\boldsymbol{d}|\theta) \propto e^{-\frac{1}{2}(\boldsymbol{d}-\boldsymbol{s}(\theta))^{\mathsf{T}}\mathsf{N}^{-1}(\boldsymbol{d}-\boldsymbol{s}(\theta))}, \tag{22}$$

where N is the noise covariance matrix. Since the following analyses are performed at the full angular resolution of the *Planck* instrument, we approximate this matrix by its diagonal, and thereby only take into account the scan-modulated white noise component, not correlated noise features or instrumental systematics. However, since N affects only the relative weighting between channels, the derived marginal map products remain

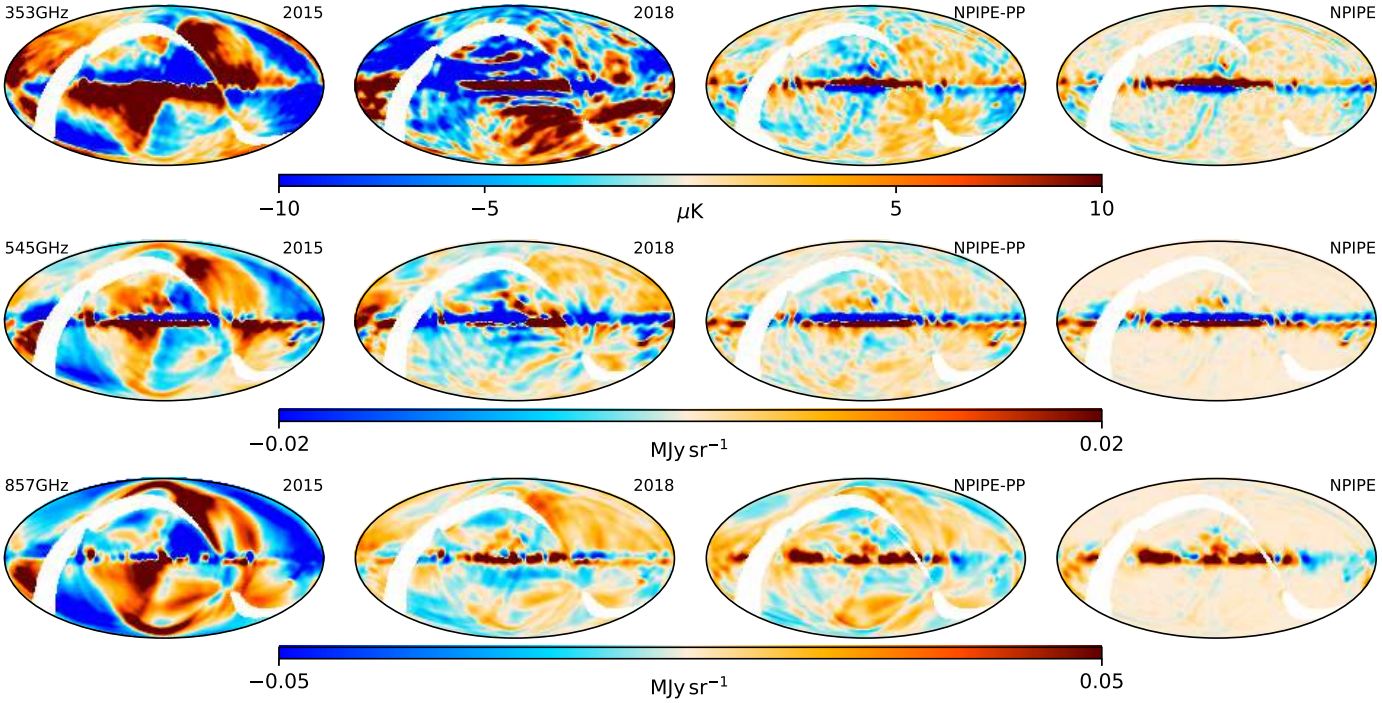


Fig. 47. Same as Fig. 46, but for 353, 545, and 857 GHz. The S-shaped residual along the Ecliptic equator, especially in the 353 GHz NPIPE results, is residual zodiacal emission.

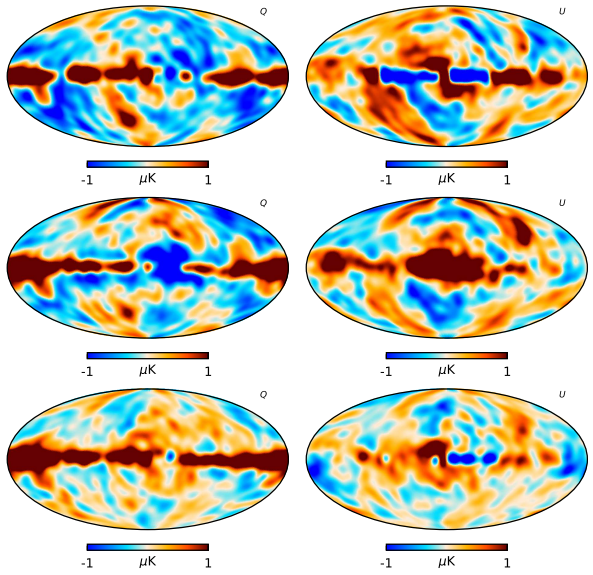


Fig. 48. Internal frequency-difference polarization maps of the form $\boldsymbol{m}_{217} - 0.128\,\boldsymbol{m}_{353}$, where the 353 GHz scaling factor is designed to suppress thermal dust emission at 217 GHz. From top to bottom, the three rows show difference maps based on SRoll, SRoll2, and NPIPE. The left and right columns show Stokes Q and U parameters, respectively. All maps have been smoothed to a common angular resolution of 10°.

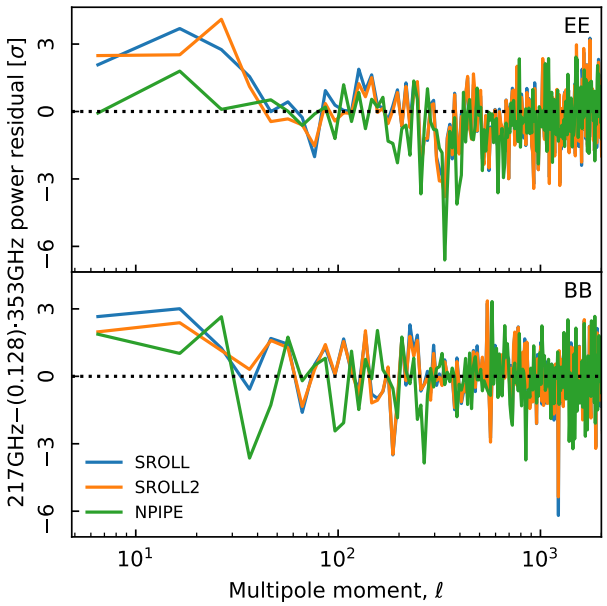


Fig. 49. Angular cross-spectra evaluated from $\boldsymbol{m}_{217}$ to $0.128\,\boldsymbol{m}_{353}$ difference maps. For NPIPE, the cross-spectra are evaluated from detector-split difference maps, while for SRoll and SRoll2 they are evaluated from half-mission split difference maps.

unbiased, albeit slightly sub-optimal in terms of variance. Final uncertainties are assessed with simulations, for which the same effects are present.

7.2. SEVEM methodology

SEVEM (Leach et al. 2008; Fernández-Cobos et al. 2012) is based on internal template cleaning in real space. In previous *Planck* releases, it was one of four approaches used to obtain the CMB signal from frequency maps (Planck Collaboration IV 2020). The internal templates trace foreground emission at the corresponding frequency range, and are constructed as difference maps between two neighbouring *Planck* channels, convolved to the same resolution to facilitate removal of the CMB contribution. A linear combination of these templates is then subtracted from a CMB-dominated frequency map, in such a way that the coefficients of the combination minimize the variance of the

resulting map outside a given mask. Different single-frequency-cleaned maps are produced in the range 70–217 GHz, then a set of them is co-added, in harmonic space, into a single map to produce a final CMB map with higher S/N.

The single-frequency, cleaned maps produced by SEVEM are useful in testing the robustness of results versus the presence of foreground residuals and systematics, for instance for isotropy and statistics estimators (Planck Collaboration VII 2020) or the integrated Sachs-Wolfe stacking analysis (Planck Collaboration XXI 2016). They are also valuable in constructing cross-frequency estimators, which allow one to minimize the impact of certain types of systematic effects, such as possible correlated noise in the data splits.

We used the same pipeline as in Planck Collaboration IV (2020) to extract the CMB signal from the NPIPE frequency maps (unlike Commander, the SEVEM pipeline does not use single-bolometer maps). However, since the NPIPE frequency maps retain the full stationary signal at each frequency, the first step is to subtract the Solar dipole, the frequency-dependent second-order quadrupole contributions, as well as the stationary component of the zodiacal light; the time-dependent component has already been removed during mapmaking. To reduce contamination from point sources, we use specific catalogues derived from the NPIPE maps. As in previous releases, point sources are detected in each frequency map using the Mexican-Hat-Wavelet 2 code in intensity (López-Caniego et al. 2006; Planck Collaboration XXVI 2016) and the Filtered Fusion code in polarization (Argüeso et al. 2009). In general, the number of point sources detected in temperature is similar to that in *Planck* 2018, with some changes close to the sensitivity limit due to the lower noise in the NPIPE maps. However, the number of polarized sources detected in the NPIPE maps is smaller than in the 2018 maps (except at 44 GHz), because better treatment of leakage from temperature to polarization reduces the number of spurious sources.

In addition to the full-survey maps, we also consider the NPIPE A/B detector split maps described in Sect. 4.1.5. These are constructed to be as independent as possible, and the SEVEM coefficients are independently fitted for each case. This is different from earlier *Planck* releases, in which data splits were propagated through the pipeline using the same coefficients as derived from the full-mission data. Because the A/B splits do not have the same effective central frequency, foregrounds do not completely cancel in the AB half-difference map, and independent propagation is as such even more important for this data split[14].

The linear coefficients of the templates used to extract the single-frequency, cleaned maps with SEVEM are shown in Tables 8 (for intensity) and 9 (for polarization). For intensity, the coefficients are quite similar to those from *Planck* 2018, with some differences in the contribution of the low-frequency templates to the cleaned HFI maps, especially 143 GHz. For polarization, the impact in the value of the coefficients is somewhat larger. Given the lower S/N of the polarization data than the temperature data, these values are more affected by the noise in the maps. Therefore, the different processing of the data using NPIPE introduces larger differences in the coefficients. For completeness, the linear coefficients used for the NPIPE splits are also given. Some differences between the coefficients for the NPIPE full-mission and NPIPE split data are seen, mainly due to the larger noise contribution in the A/B splits with respect to the full mission. In most cases, the template fitting tries to minimize the effect of this larger noise in the cleaned CMB map by reducing the absolute value of the corresponding coefficients. In particular, the most pronounced deviations are seen in the coefficients for the LFI templates, specially when cleaning the 100 GHz map. For intensity, we checked that the major effect is actually due to the different noise level in the 44–70 template, which affects to the coefficient of both LFI templates. For polarization, the larger noise levels of the 30–44 templates constructed with the splits also affect the coefficients, although in this case some differences are also introduced by the HFI templates.

7.3. Data selection and goodness of fit

An important feature of the Bayesian parametric analysis framework, as implemented in Commander, is its ability to provide robust and intuitive goodness-of-fit statistics in the form of residual maps, $\boldsymbol{r} = \boldsymbol{d} - \boldsymbol{s}$, where $\boldsymbol{d}$ denotes the observed channel maps and $\boldsymbol{s}$ is the fitted sky model defined in Eqs. (19) and (20), and χ^2 statistics. These statistics are powerful probes of residual systematics in the data, since they highlight any discrepancies between the raw data and the assumed model. By inspecting the morphology of these maps, it is often possible to understand the physical origin of a given model failure, which in turn can suggest improvements either to the data processing or to the fitted model. Indeed, many of the issues discussed in Sect. 2 were discovered precisely through these statistics, which in turn allowed us to improve the overall quality of the NPIPE data processing.

The first detailed analysis of the *Planck* data based on this framework was presented in Planck Collaboration X (2016). These data included single-detector and detector-set temperature sky maps, in addition to co-added full-frequency maps, allowing a much more fine-grained data-inspection and data-selection process than did either the 2013 or the 2018 releases, both of which included only full-frequency maps. This is only relevant for temperature, not polarization, since the *Planck* scan strategy does not allow us to solve for polarized maps from single-detector data. In the following, therefore, we compare the new NPIPE data products to the *Planck* 2015 release in temperature, and to the 2018 data release in polarization.

In the 2015 Commander temperature analysis, 21 single-detector, detector-set, and full-frequency *Planck* maps were considered sufficiently clean from residual instrumental systematic effects to be included in the final analysis, while 10 maps (three detector-set and seven single-detector maps) were excluded due to large unexplained systematic effects. Figure 52 compares *Planck* 2015 and NPIPE residual maps for the seven excluded single-detector sky maps for which a direct head-to-head comparison is possible (the NPIPE data release does not provide detector-set maps in the same form as *Planck* 2015).

Four important effects can be seen in Fig. 52. First and foremost, we see that the overall level of coherent systematic residuals in these channels is greatly reduced in NPIPE, to the level that all except one can be included in the final analysis. The only exception is 857-4, which we find would contribute as much to χ^2 as all other *Planck* channels combined if included in the analysis. Accounting for the fact that NPIPE provides single-detector maps also for polarized channels, a total of 36 *Planck* maps are included in the following Commander temperature analysis. To understand how these improvements are achieved, it is

[14] In contrast to the case in previous *Planck* releases, the beams for the full and split data sets are slightly different, and this is also the case in the cleaned single-frequency maps. However, by construction, the final combined CMB map has the same resolution (Gaussian beam with a FWHM of 5′) for the full-frequency and split-data maps.

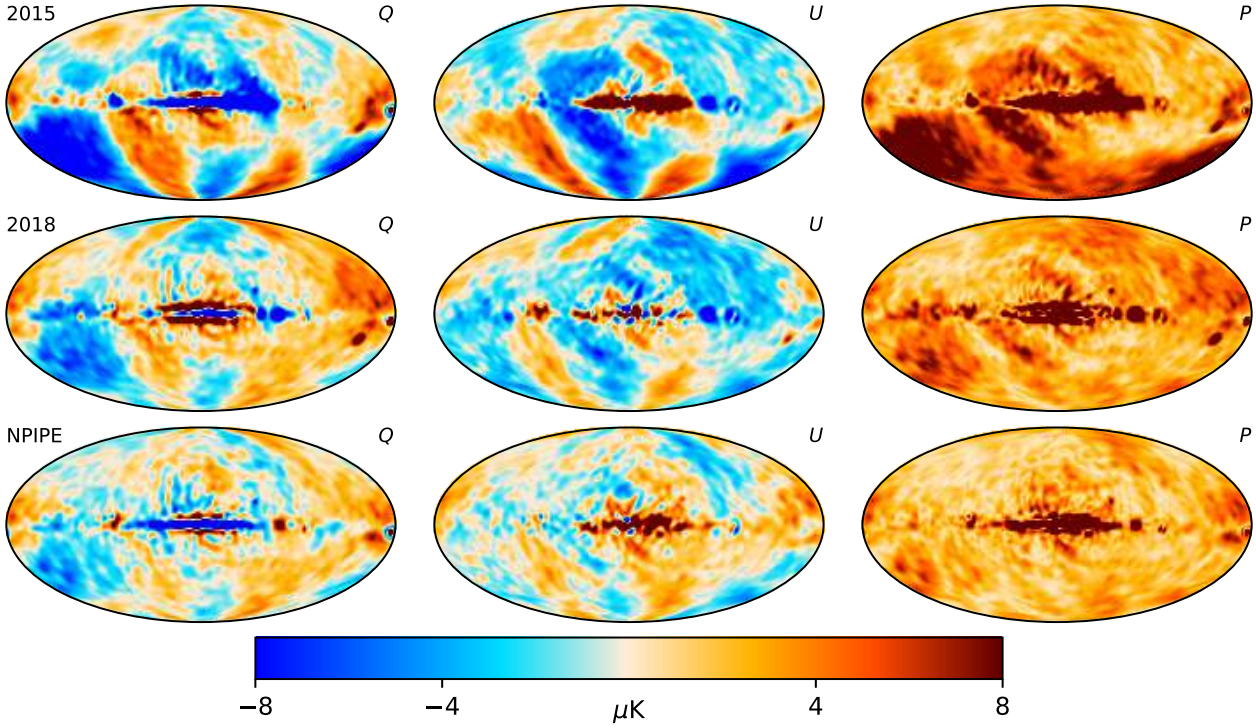


Fig. 50. *Planck* 30 GHz – WMAP *K*-band difference. The *K*-band regression coefficients by row (*top* to *bottom*) are 0.406, 0.451, and 0.462.

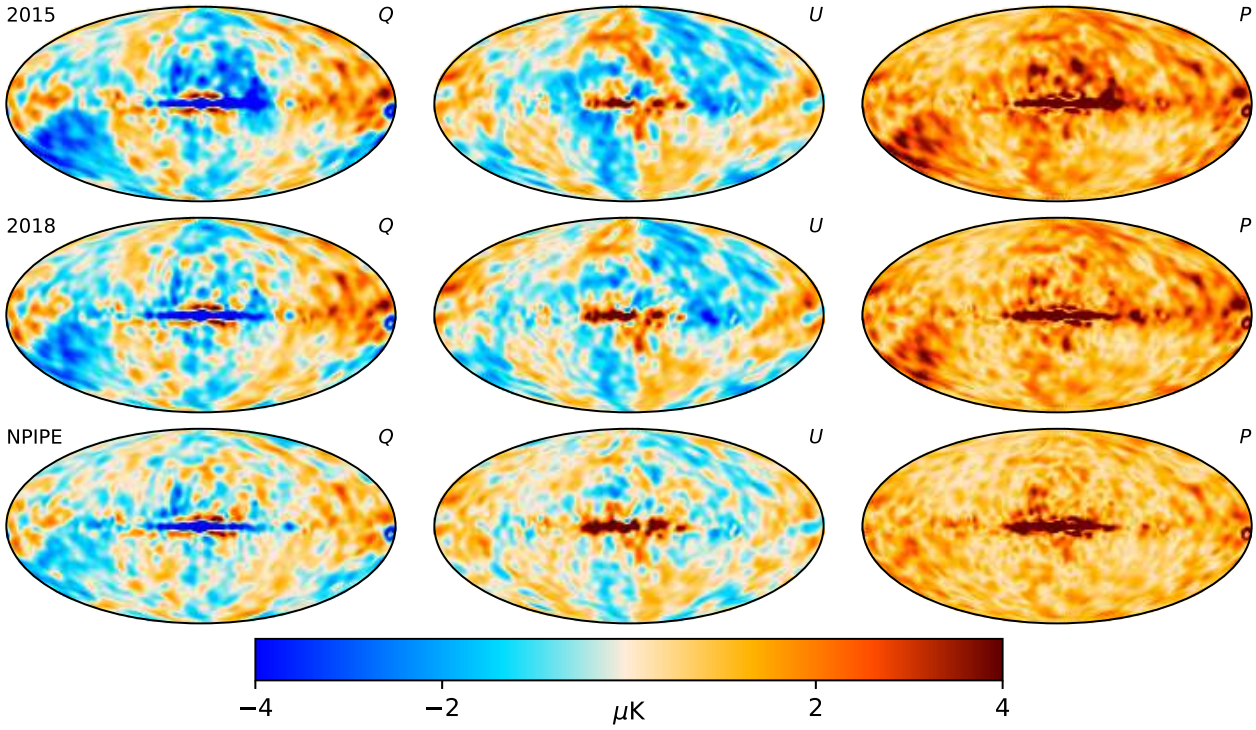


Fig. 51. *Planck* 44 GHz– WMAP *K*-band difference. The *K*-band regression coefficients by row are 0.104, 0.117, and 0.125.

useful to inspect the residual maps in Fig. 52. Starting with 353-8, the dominant feature is a large-scale red and blue pattern at high Galactic latitudes, extending between the ecliptic poles. This pattern arises when multiple detectors with slightly different bandpasses are destriped simultaneously. Specifically, the destriping (or mapmaking) process intrinsically assumes that there is only one true value within a single sky pixel, and that any deviation from this must therefore be due to noise in the detectors. However, because different detectors have different bandpasses, they also see a slight difference in foreground sky signal. The destriper therefore attempts to suppress this residual sky signal in the same way as actual correlated noise, and effectively "drags" the signal along the scan path of the instrument, resulting in the large-scale features seen in Fig. 52. To solve this

Table 8. Linear coefficients of the templates used to clean individual frequency maps with SEVEM for temperature.

Template	SEVEM coefficients			
	70 GHz	100 GHz	143 GHz	217 GHz
Full mission				
30–44	1.64×10^{-1}	-8.34×10^{-2}	1.70×10^{-2}	-1.49×10^{-1}
44–70	…	3.90×10^{-1}	1.56×10^{-1}	3.95×10^{-1}
353–143	6.85×10^{-3}	…	…	…
545–353	…	4.12×10^{-3}	6.28×10^{-3}	1.71×10^{-2}
857	…	-3.18×10^{-5}	-5.09×10^{-5}	-1.05×10^{-4}
A split				
30–44	1.66×10^{-1}	-4.58×10^{-2}	2.30×10^{-2}	-1.22×10^{-1}
44–70	…	2.92×10^{-1}	1.40×10^{-1}	3.14×10^{-1}
353–143	6.82×10^{-3}	…	…	…
545–353	…	4.00×10^{-3}	6.15×10^{-3}	1.67×10^{-2}
857	…	-3.21×10^{-5}	-5.14×10^{-5}	-1.05×10^{-4}
B split				
30–44	1.62×10^{-1}	-4.80×10^{-2}	1.91×10^{-2}	-1.05×10^{-1}
44–70	…	2.92×10^{-1}	1.43×10^{-1}	2.84×10^{-1}
353–143	6.96×10^{-3}	…	…	…
545–353	…	4.15×10^{-3}	6.33×10^{-3}	1.73×10^{-2}
857	…	-3.20×10^{-5}	-5.06×10^{-5}	-1.06×10^{-4}

Notes. The 353−143 template has been produced at the same resolution as the 70 GHz frequency channel, while the 857 GHz map has been convolved with the 545 GHz beam. The rest of the templates are constructed such that the first map in the subtraction is convolved with the beam of the second map, and vice versa.

Table 9. Linear coefficients for each of the templates used to clean individual frequency maps with SEVEM for polarization.

Template	SEVEM coefficients							
	70 GHz Q	70 GHz U	100 GHz Q	100 GHz U	143 GHz Q	143 GHz U	217 GHz Q	217 GHz U
Full mission								
30–44	3.07×10^{-2}	3.60×10^{-2}	1.10×10^{-2}	1.32×10^{-2}	4.74×10^{-3}	8.43×10^{-3}	1.11×10^{-2}	1.37×10^{-2}
143–100	…	…	…	…	…	…	0.79	0.62
217–100	…	…	…	…	1.29×10^{-1}	1.20×10^{-1}	…	…
217–143	…	…	0.87×10^{-1}	0.84×10^{-1}	…	…	…	…
353–143	1.18×10^{-2}	0.98×10^{-2}	…	…	…	…	1.18×10^{-1}	1.17×10^{-1}
353–217	…	…	1.38×10^{-2}	1.32×10^{-2}	3.15×10^{-2}	3.07×10^{-2}	…	…
A split								
30–44	2.03×10^{-2}	1.77×10^{-2}	0.67×10^{-2}	0.68×10^{-2}	1.48×10^{-3}	3.65×10^{-3}	1.45×10^{-2}	1.31×10^{-2}
143–100	…	…	…	…	…	…	0.90	0.68
217–100	…	…	…	…	1.75×10^{-1}	1.64×10^{-1}	…	…
217–143	…	…	1.12×10^{-1}	1.04×10^{-1}	…	…	…	…
353–143	1.18×10^{-2}	0.95×10^{-2}	…	…	…	…	1.17×10^{-1}	1.13×10^{-1}
353–217	…	…	1.13×10^{-2}	1.12×10^{-2}	2.58×10^{-2}	2.55×10^{-2}	…	…
B split								
30–44	2.00×10^{-2}	2.48×10^{-2}	0.64×10^{-2}	0.88×10^{-2}	3.62×10^{-3}	6.48×10^{-3}	1.61×10^{-2}	1.62×10^{-2}
143–100	…	…	…	…	…	…	1.31	1.12
217–100	…	…	…	…	2.09×10^{-1}	2.13×10^{-1}	…	…
217–143	…	…	1.31×10^{-1}	1.29×10^{-1}	…	…	…	…
353–143	1.19×10^{-2}	0.94×10^{-2}	…	…	…	…	1.03×10^{-1}	0.96×10^{-1}
353–217	…	…	0.94×10^{-2}	0.86×10^{-2}	2.11×10^{-2}	1.90×10^{-2}	…	…

Notes. Here, the 30−44 template is constructed such that the 30 GHz map is smoothed with the 44 GHz beam and vice versa. The 217−143, 217−100 and 143−100 templates are produced with 1° resolution, and 353−217 and 353−143 with 10′. In addition, this last template is also constructed at the resolution of the 70 GHz beam, in order to clean that channel.

problem, one can either destripe each detector independently (after removing a polarization template in the time domain), or fit a set of residual foreground templates in the time domain jointly with the destriping offsets. NPIPE implements both solutions, the former for the single-detector temperature maps, and the latter for the polarized full-frequency and detector-set maps.

A second, and visually striking, effect seen in Fig. 52 is the result of sub-optimal instrumental polarization parameters, in the form of the assumed polarization angle and efficiency of each detector. This is most clearly seen in 353-2, in the form of an alternating blue and red Galactic plane, with amplitudes following the overall polarization amplitude. When either

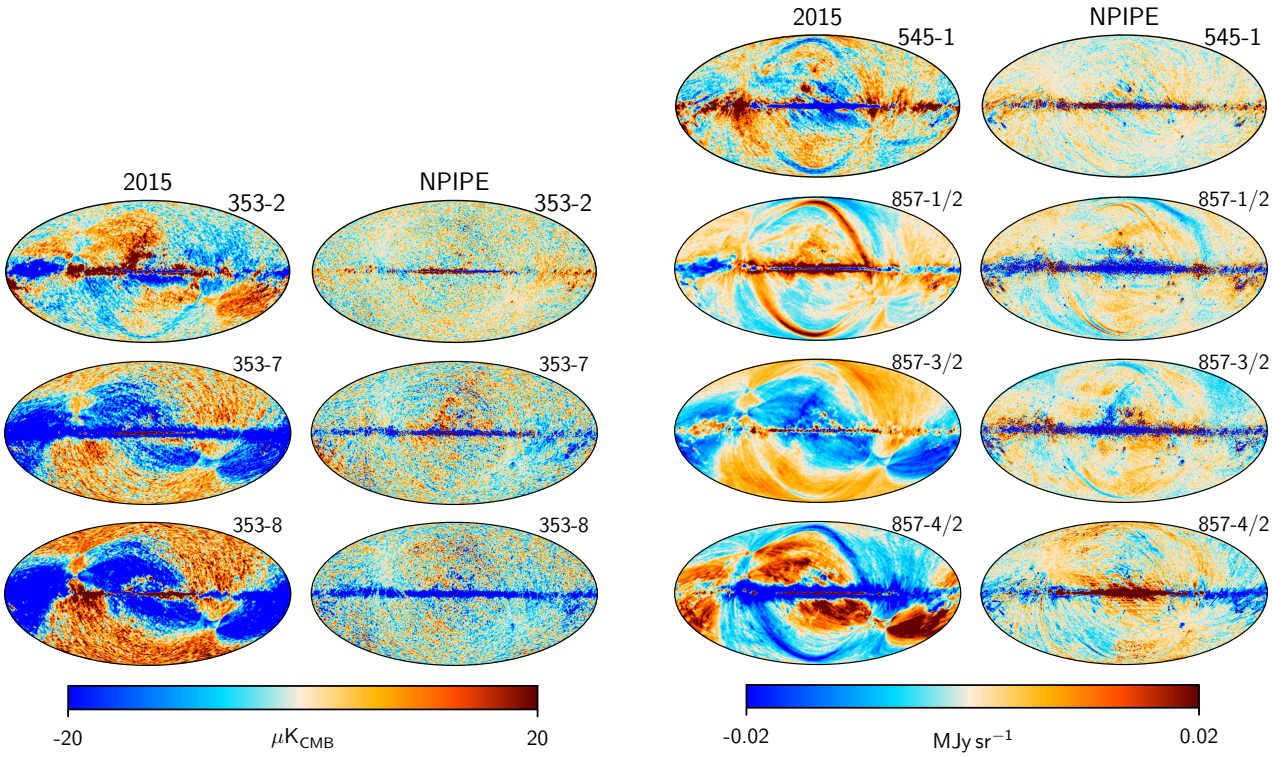


Fig. 52. *Planck* 2015 residual maps of rejected single bolometers (see Planck Collaboration X 2016) and the corresponding NPIPE residual maps. All maps have a smoothing of 60′ FWHM. The fraction in the label of the 857 GHz detectors indicates that the maps are divided by 2 with respect to the colour bar. The overall level of coherent systematic residuals in these maps is dramatically reduced in NPIPE, to the level that all except 857-4 (see text) can now be included in the final analysis.

polarization parameter is incorrectly set prior to mapmaking, some of the polarization signal will be interpreted as a temperature signal and vice versa. These residuals are precisely what allows us to perform in-flight determination of the *Planck* polarization parameters, as described in Sect. 2.

Third, for 857-1 in Fig. 52 we clearly see the imprint of residual sidelobe contamination in the form of two red arcs at high Galactic latitudes. These are greatly suppressed (albeit still visually clear) in NPIPE, due to better sidelobe fitting.

Fourth, the effect of improved transfer-function estimation is seen for 857-4 in Fig. 52. The thick blue region around the Galactic plane is caused by sub-optimal detector transfer functions, which in effect smear out the very bright Galactic signal. Overall, transfer-function estimation is greatly improved in the NPIPE processing, even though this channel has proved particularly challenging, and is still not deemed sufficiently clean for detailed analysis.

Figure 53 shows data-minus-signal residual maps for the 36 NPIPE maps included in the Commander temperature analysis. Even though some coherent structures are visible, for frequencies below 545 GHz these are visually dominated by white noise at high Galactic latitudes. Starting with the 44 GHz channel, the most striking features are patterns consistent with Galactic residuals, some bright red with free-free-like morphology and some faint blue with the morphology of dust emission. These are due to the fact that only a single low-frequency component is fitted to these *Planck*-only data, and a single component cannot account separately for synchrotron, free-free, and spinning dust emission. For 100-1a, the dominant feature is a roughly 1 μK negative monopole, resulting from the fact that the offset for this particular channel is not fitted freely in the analysis, but rather determined by enforcing vanishing CMB and CO zero-level at high Galactic latitudes.

For 100-1b, the most notable features are a slight red CO residual in the central Galactic plane, countered by a faint blue residual with thermal-dust-like morphology extending into the Galactic wings. These structures suggest that CO is not perfectly modelled in the current setup. One possible explanation may be the presence of multiple CO isotopes (i.e., ^{12}CO and ^{13}CO) in this band, while we only fit for one overall component in our analysis. Extensions of this model will be considered in the future. A second possibility is residual thermal dust temperature uncertainties. In any case, this model failure appears to bias the bandpass corrections very slightly, resulting in a slightly negative thermal-dust imprint outside the Galactic centre. This particular case illustrates very directly the importance of joint global component separation and instrumental characterization, as implemented in NPIPE, as well as the importance of employing a complete physical model for this task.

A third model failure of similar type is seen in the 100-2b map, in the form of sharp blue and red residuals on either side of the Galactic centre. These are due to the rotation of the Milky Way, resulting in a relative blue- or red-shift of the CO lines. Since the derivative of the detector bandpasses at the CO centre frequency is not zero, an effective shift in the apparent line ratio results. This effect was already noted in the *Planck* 2015 data set (Planck Collaboration X 2016), but with the improved NPIPE processing, it is now the dominant signal for several channels.

For channels below 217 GHz, sky-model failures dominate over instrumental effects, whether they come in the form of residual free-free, spinning dust, or CO features. Resolving these will require both a more fine-grained astrophysical model and

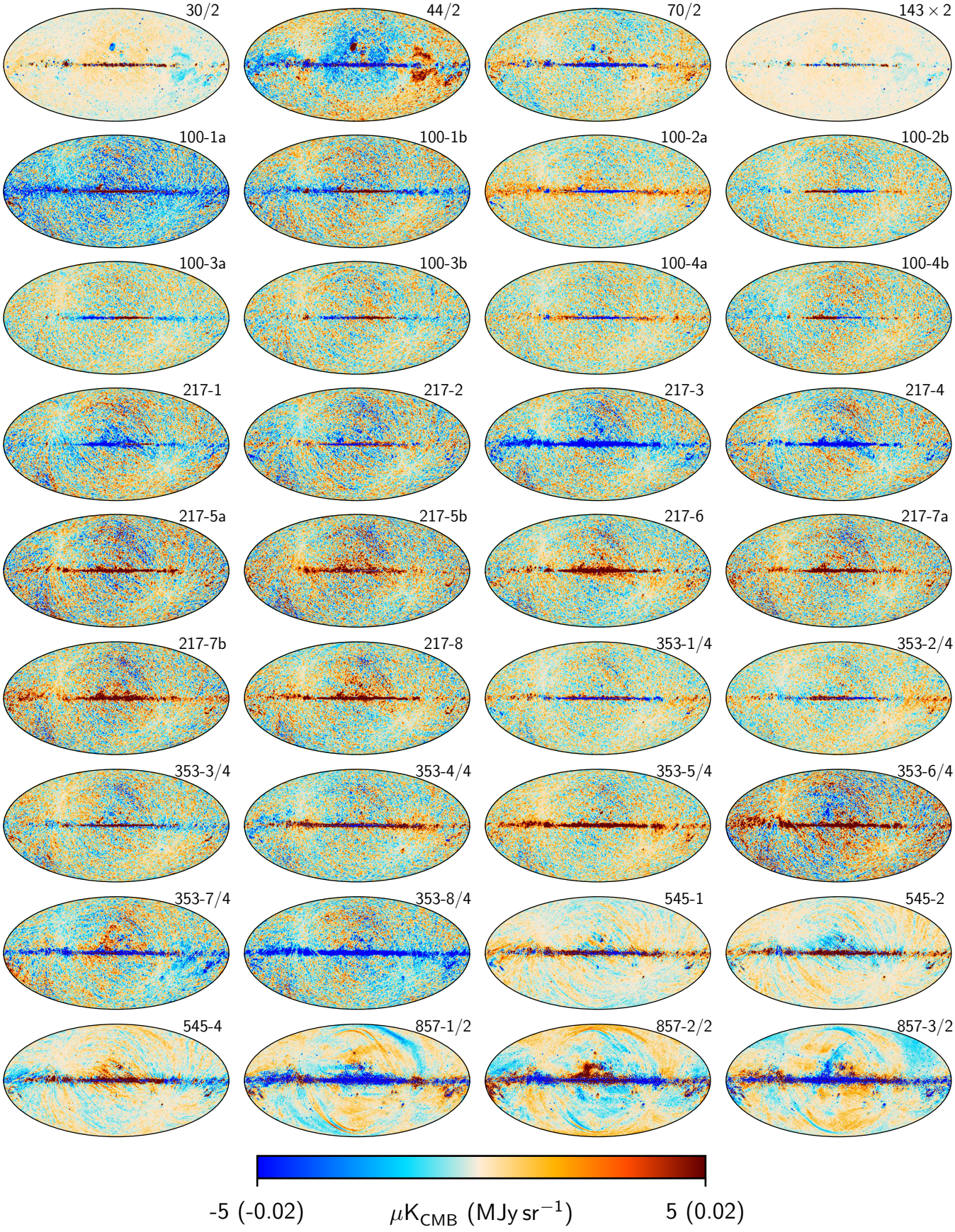


Fig. 53. *Planck*-only NPIPE residual maps for frequency bands and individual detectors. Maps up to 353 GHz are plotted in μK_{CMB}, while maps at 545 and 857 GHz are plotted in MJy sr^{-1}. Maps with a fraction in the label have been divided with respect to the colour bar, while the 143 GHz band has been multiplied by a factor of 2. All maps are smoothed to 60′ FWHM.

combination with external data sets. This is beyond the scope of the current paper, and will be addressed in future publications.

At frequencies above 143 GHz, the picture is more complicated. Here we see many features that have a clear instrumental signature, most of which have already been discussed qualitatively. Thus, even though the NPIPE data set exhibits lower systematic uncertainties than previous *Planck* releases, the data are clearly still not consistent with white noise. It is important to bear these residuals in mind in subsequent analyses, and analysing realistic simulations is important to quantify the resulting uncertainties. On the other hand, in terms of their potential effect on cosmology, one should bear in mind that these residuals are at a low level compared to the CMB signal, especially since they will average to an even lower level in the combination of maps

Figure 54 shows similar residual maps from the polarization analysis, compared with the corresponding *Planck* 2018 residuals. The single most striking difference between these two data sets is the lower level of coherent and scan-aligned large-scale features at high Galactic latitudes. This is primarily due to improved calibration in NPIPE, resulting in lower levels of dipole leakage into the polarization sector.

7.4. Global parameters

We start our discussion of the derived sky model by considering global parameters computed by Commander and listed in Table 10. For each sky map included in the analysis, three values are listed: monopole (or offset); calibration factor; and bandpass shift. Values that are kept fixed during a Commander run are marked with a superscript "a". Note, however, that the fixed monopoles are adjusted between preliminary and final runs, as described in Sect. 7.1.

Starting with the monopoles, the reported values are numerically large, and do not correspond directly to typical values reported for previous data sets. In this respect, we note that *Planck* is not meaningfully sensitive to the zero-level of any frequency channel. Accordingly, the NPIPE analysis pipeline makes no attempt to adjust the raw outputs from the Madam mapmaker (from which the monopoles are spurious), but rather leaves this task to the component-separation stage, which is far better equipped to solve this problem, exploiting both physical priors and correlations between frequency channels. We therefore recommend subtraction of the values reported in Table 10 prior to any application that is sensitive to the absolute zero-level of the maps.

Regarding the calibration coefficients, the CMB Solar dipole is, for the first time in *Planck* data, retained in all sky maps. This provides a very bright relative calibration target that allows for high-precision calibration and component separation. The 143 GHz calibration factor is fixed to unity in these analyses, and all other values are therefore in effect defined relative to this channel. For the 30 and 44 GHz data, we find calibration corrections of −0.75 and −0.40%, respectively. These channels have relatively bright foregrounds combined with lower signal-to-noise ratios, and are therefore more susceptible to potential foreground residuals during calibration. In particular, we repeat that the foreground model used in this *Planck*-only analysis relies on a single joint low-frequency component, as opposed to separate synchrotron, free-free, and spinning dust components.

We observe excellent relative calibration among the 70 and 100 GHz sky maps, with an overall shift of about −0.08% relative to 143 GHz, and internal variations smaller than 0.02%. The same holds true at 217 GHz, for which the overall shift is 0.02%, and the scatter is about the same. For the 353 GHz channels, where the thermal dust foregrounds become significantly brighter, we find an overall shift of about 0.25%, with a scatter of about 0.05%.

NPIPE represents the first *Planck* processing for which the 545 GHz channel is calibrated with the CMB orbital dipole along with the lower-frequency channels. After component separation, these estimates appear accurate with a precision of about 2%.

We obtain very large correction factors for the 857 GHz channels, with values up to 16%. The reason for this is that the NPIPE processing natively adopts the same planet-based calibration procedure as the 2018 DPC processing, defined in units of MJy sr^{-1}, but subsequently converts these to thermodynamic units of $K_{\rm CMB}$. The primary advantage of this conversion is that all maps then have the same units. The main disadvantage is that the conversion factor between flux density and thermodynamic units is highly sensitive to small variations and uncertainties in the shape of the bandpass. This is particularly striking for the 857-2 channel, with its total correction factor of 16.1%.

The bandpass corrections listed in the rightmost column of Table 10 are generally small. The largest numerical values are observed within the 100 GHz channel, for which a mean negative shift of about 1 GHz is seen. This corresponds to a reduction in the effective thermal-dust-emission amplitude of about 1.5% in this channel, and an increase in the free-free emission of about 2%, compared to the nominal bandpasses. Both effects are relatively small in an absolute sense, but still statistically significant given the very high signal-to-noise ratio of the *Planck* observations.

For the 70 GHz channel, we observe an average bandpass shift of about (-1.1 ± 1.0) GHz. For comparison, a value of (0.4 ± 1.0) GHz was found in the corresponding *Planck*+WMAP analysis presented in Planck Collaboration Int. XLVI (2016). The *Planck* 70 GHz bandpass shifts are associated with a large uncertainty for three main reasons. First, the 70 GHz bandpasses were measured with lower accuracy on the ground than other detectors, and significant power is missing from the tails of the bandpass profiles (see Fig. 18 of Villa et al. 2010). Second, the foreground minimum occurs close to 70 GHz (Planck Collaboration Int. XLVI 2016), leaving a relatively faint calibration target for the bandpass shift parameters. Third, the strongest foreground components – thermal dust, synchrotron, and free-free – are all about equally strong around 70 GHz. Modelling errors and internal degeneracies are therefore significant for this channel. The net result of these three effects is a large overall uncertainty, in which the modelling aspects dominate the error budget.

We conclude this section with a brief discussion of uncertainties. Because of *Planck*'s high S/N, these are dominated by systematic contributions. For example, the conditional statistical uncertainty on the 100-1a monopole from instrumental noise alone is about 10 nK. This contribution is completely negligible compared to foreground modelling uncertainties and instrumental systematics. Deriving a statistically rigorous estimate for these uncertainties is therefore complicated. However, as a useful approximation, we exploit the fact that the NPIPE data set has undergone hundreds of larger or smaller variations during its development phase, where each variation has included different processing features and foreground modelling. The uncertainties quoted in Table 10 represent the typical variations observed among the two most mature NPIPE versions. The uncertainty in the 545-2 monopole is a special case however, being derived from a linear correlation with HI observations.

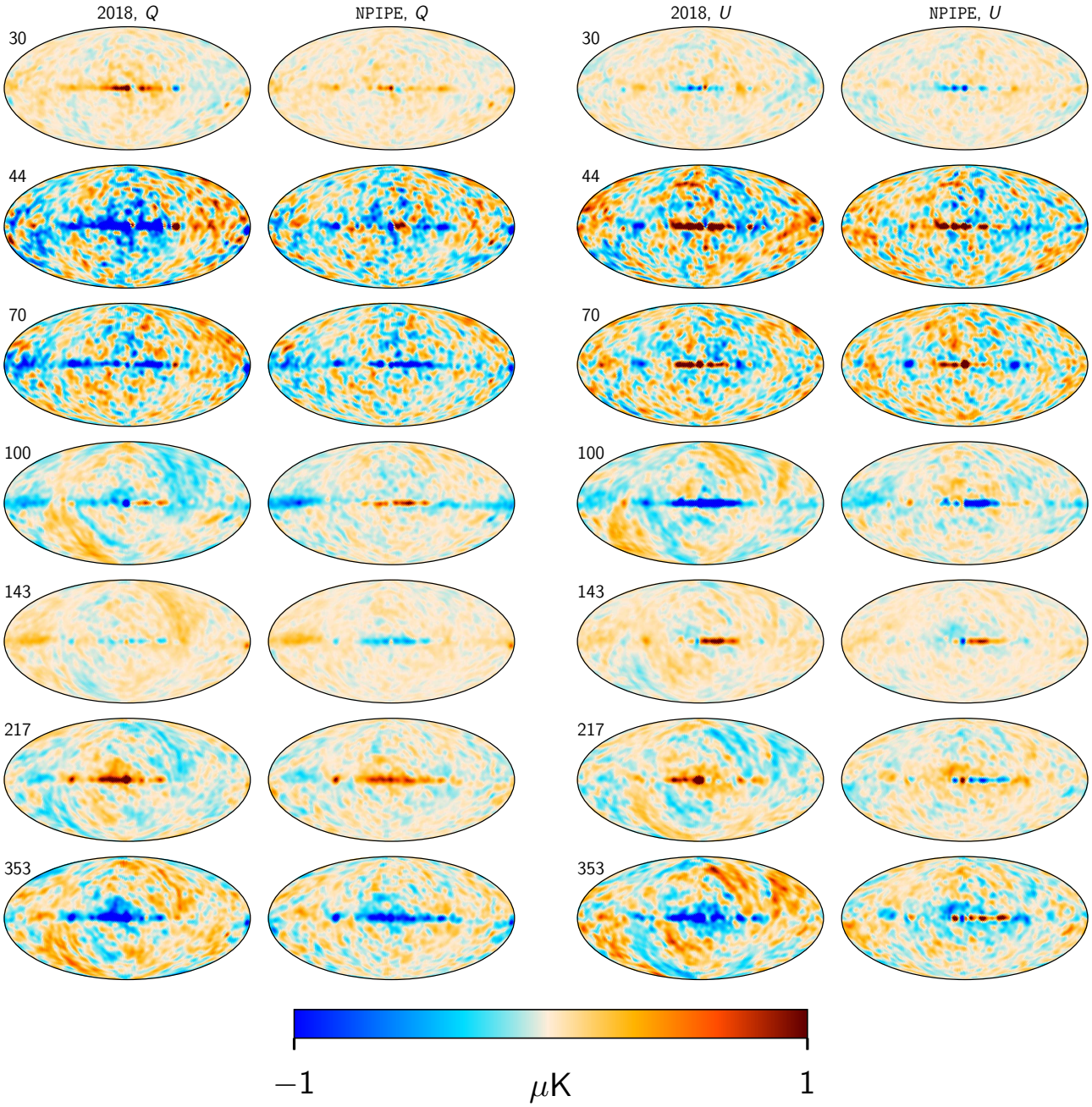


Fig. 54. Comparison between Commander residual maps, $\boldsymbol{d}_\nu - \boldsymbol{s}_\nu$, for each polarized *Planck* frequency map between (from *top* to *bottom*) 30 and 353 GHz. The *two leftmost columns* show residual maps for Stokes Q for *Planck* 2018 and NPIPE, while the *two rightmost columns* show the same for Stokes U. All maps have been smoothed to a common angular resolution of 3° FWHM.

7.5. CMB maps and power spectra

Next we consider the cosmological CMB signal as extracted with both Commander and SEVEM. Figure 55 shows three versions of the NPIPE temperature CMB map. The top panel shows the full CMB temperature map as estimated with Commander. This map is the first component-separation-based product that allows direct estimation of the Solar dipole across the entire *Planck* frequency range, as discussed in detail in Sect. 8. The middle panel of Fig. 55 shows the same CMB temperature map after subtracting the best-fit NPIPE dipole, while the bottom panel shows the same, but estimated with SEVEM. Figure 56 shows the corresponding maps of Stokes Q and U from both Commander (top row) and SEVEM (bottom row).

While the overall morphology of the CMB I, Q, and U maps at high Galactic latitudes looks generally consistent with expectations based on a Gaussian and isotropic ΛCDM universe, there are clear visual indications of significant foreground-induced residuals at low Galactic latitudes. Consequently, proper masking is required before subjecting these maps to detailed scientific analysis. For temperature, we construct an analysis mask in a manner analogous to the approach taken in Planck Collaboration IV (2020), where a smoothed standard-deviation map evaluated between four different CMB component-separation estimates was thresholded at a given value. In this paper, we instead threshold the standard-deviation map evaluated from the Commander CMB temperature maps derived from the three latest *Planck* data releases, namely the

Table 10. Monopoles, calibration factors, and bandpass corrections derived within the baseline temperature model.

Frequency [GHz]	Detector label	Monopole [μK]	Calibration [%]	Bandpass shift [GHz]
30	. . .	773 ± 5 [a]	−0.75 ± 0.07	0.4 ± 0.1
44	. . .	1187.3 ± 0.3	−0.40 ± 0.08	−0.6 ± 0.1
70	. . .	70.8 ± 0.3	−0.09 ± 0.01	−1.1 ± 1.0
100	1a	−68.9 ± 1 [a]	−0.08 ± 0.03	0.1 ± 0.8
	1b	−70.0 ± 0.1	−0.07 ± 0.02	−0.4 ± 0.8
	2a	−69.4 ± 0.1	−0.07 ± 0.02	−2.1 ± 0.8
	2b	−69.5 ± 0.1	−0.08 ± 0.02	−1.2 ± 0.3
	3a	−69.7 ± 0.1	−0.07 ± 0.02	−1.7 ± 0.4
	3b	−69.9 ± 0.1	−0.06 ± 0.02	−0.7 ± 0.2
	4a	−69.3 ± 0.1	−0.08 ± 0.02	−2.2 ± 0.4
	4b	−69.6 ± 0.1	−0.08 ± 0.02	−1.4 ± 0.3
143	. . .	−81.1 ± 2 [a]	0 [a]	0.4 ± 0.1
217	1	−182.6 ± 0.2	−0.01 ± 0.01	0.0 [a]
	2	−182.4 ± 2 [a]	−0.02 ± 0.01	−0.3 ± 0.1
	3	−182.8 ± 0.2	−0.03 ± 0.01	0.4 ± 0.1
	4	−182.6 ± 0.2	−0.02 ± 0.01	0.1 ± 0.1
	5a	−182.2 ± 0.2	−0.03 ± 0.01	0.3 ± 0.1
	5b	−182.2 ± 0.2	−0.04 ± 0.01	0.4 ± 0.1
	6	−182.1 ± 0.2	−0.01 ± 0.01	0.6 ± 0.1
	7a	−182.1 ± 0.2	−0.02 ± 0.01	−0.2 ± 0.1
	7b	−182.2 ± 0.2	−0.03 ± 0.01	0.2 ± 0.1
	8	−182.1 ± 0.2	−0.02 ± 0.01	0.0 ± 0.1
353	1	395.2 ± 1.0	0.27 ± 0.08	0.2 ± 0.1
	2	394.6 ± 1.0	0.27 ± 0.08	−0.4 ± 0.1
	3	394 ± 4 [a]	0.23 ± 0.08	0.0 ± 0.1
	4	394.7 ± 1.0	0.24 ± 0.08	0.3 ± 0.1
	5	397.8 ± 1.0	0.24 ± 0.08	−0.2 ± 0.1
	6	398.2 ± 1.0	0.20 ± 0.08	−0.1 ± 0.1
	7	386.1 ± 1.0	0.26 ± 0.08	−0.5 ± 0.1
	8	383.8 ± 1.0	0.26 ± 0.08	−0.2 ± 0.1
545	1	−5003 ± 30	1.5 ± 0.3	−0.4 ± 0.1
	2	−4891 ± 852 [a]	1.5 ± 0.3	0.4 ± 0.1
	4	−4951 ± 30	2.4 ± 0.3	0.0 ± 0.1
857	1	−0.72 ± 0.01 [b]	4.7 ± 0.5	0.0 [c]
	2	−0.72 ± 0.01 [b]	16.1 ± 0.6	0.0 [c]
	3	−0.72 ± 0.01 [b]	7.9 ± 0.5	0.0 [c]

Notes. [a]Fixed at reference value. [b]The units for 857 GHz are kelvin. [c]Fixed to the input value due to strong degeneracies with calibration.

Planck 2015 and 2018 data sets and the new NPIPE data set. This is a conservative approach, since it only admits the parts of the sky that all three of the latest *Planck* processing pipelines agree on, and that are robust with respect to the very different foreground models adopted for the three data sets. Specifically, for *Planck* 2015 we combined the *Planck* data with Haslam 408-MHz and WMAP maps, and employed a detailed model including synchrotron, free-free, and spinning dust, while for *Planck* 2018 and NPIPE, we included only a single combined low-frequency component. Similarly, for *Planck* 2015 and NPIPE, we include three independent CO-line components in the model, while for *Planck* 2018, only one combined CO line was included. For a pixel to be accepted in the new mask, all three versions must agree to a precision better than 3 μK[15]. In addition, all pixels with a reduced χ^2 in NPIPE greater than 2 are excluded (as it happens, most of these pixels are already excluded by the standard-deviation cut). The resulting mask is shown in Fig. 57, and retains 76% of the sky.

[15] The precise numerical value for the standard-deviation cut-off is somewhat arbitrary, but also of little importance, as the gradient in the map is very sharp near the Galactic plane. Any value between 2 and 5 μK results in very similar masks, whereas values below 2 μK start picking up noise fluctuations.

The top panel of Fig. 58 shows the difference between the Commander-derived NPIPE CMB temperature map and the *Planck* 2015 CMB map smoothed to 1° FWHM. The middle panel shows the corresponding difference map with respect to the *Planck* 2018 map. The difference between the NPIPE and *Planck* 2018 SEVEM maps is shown in the bottom panel. The grey regions show the confidence mask described above. We see that the various CMB estimates outside this mask typically agree to within 1 or 2 μK. The Commander difference map evaluated with respect to the 2015 solution exhibits slightly lower residuals at high latitudes than the corresponding 2018 solution. In this respect, the NPIPE and 2015 analyses consider single-detector maps, while 2018 only considers full-frequency maps. As already noted, a more fine-grained data set allows for better foreground modelling and more selective removal of obviously bad channels. Comparing the two lower panels, we see that the morphologies are very similar between Commander and SEVEM.

For polarization, we adopt the same common analysis mask as discussed in Planck Collaboration IV (2020). This choice is motivated by the fact that in the following we will compare the NPIPE polarization products with corresponding *Planck* 2018 products, and we note that the NPIPE maps generally have smaller systematics and foreground residuals than *Planck* 2018.

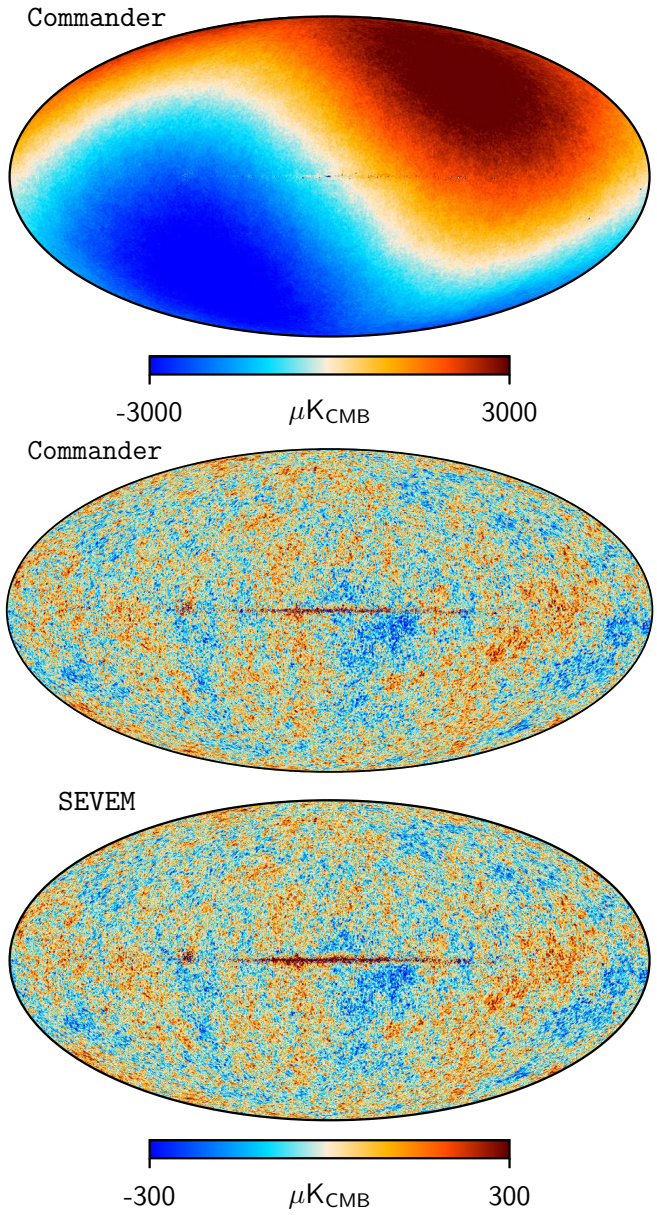


Fig. 55. *Top*: CMB I map derived from the NPIPE data set with Commander, plotted with an angular resolution of 5′ FWHM. The Solar dipole is retained in the NPIPE data set, and in this map. *Middle*: same as above, but after subtracting the best-fit CMB Solar dipole described in Sect. 8. *Bottom*: dipole-subtracted CMB temperature map derived with SEVEM.

Figure 59 compares the Commander Stokes Q and U polarization maps from the *Planck* 2015, *Planck* 2018, and NPIPE data sets, as well as the SEVEM polarization maps from *Planck* 2018 and NPIPE. All maps are smoothed to a common angular resolution of 3° FWHM, and the grey regions show the *Planck* 2018 common confidence mask. Starting from the top, we see that the Commander 2015 polarization map is dominated by large-scale systematic features. Due to this large systematic contribution, the map was never publicly released in its raw form, but only in the form of a high-pass-filtered map, from which multipoles below $\ell \leq 20$ had been removed.

Much of the analysis effort of the *Planck* team between 2015 and 2018 concentrated on understanding and mitigating these residuals. As seen in the second and third rows of Fig. 59, this work was highly successful. The large-scale polarization systematics were reduced by an order of magnitude. Furthermore, this process has continued beyond the final *Planck* 2018 release within the NPIPE framework, as shown in the bottom two rows of Fig. 59, which exhibits even lower residuals than *Planck* 2018.

Figure 60 shows cleaned, single-frequency polarization maps derived with SEVEM. Large-scale systematic features are significantly mitigated in the NPIPE data.

To quantify the differences between the *Planck* 2018 and NPIPE polarization maps further, Fig. 61 shows the angular EE and BB cross-spectra evaluated from half-mission (for *Planck* 2018) and detector-set (for NPIPE) splits, evaluated outside the 2018 common confidence mask. These two splits represent the two most independent data subsets available within each data set. The most notable feature in these plots is the fact that the blue curves (corresponding to the Commander 2018 spectrum) generally encompass the red and orange curves (corresponding to the Commander and SEVEM NPIPE spectra). This is a direct reflection of the fact that the overall noise level in the NPIPE data set is about 15% lower than in *Planck* 2018, which results in a lower level of BB power at all scales. Other than this noise difference, the two data sets appear statistically quite similar.

Figure 62 shows the same spectra, but only for $\ell \leq 15$. Here we see that the *Planck* 2018 cross-spectra exhibit a statistically significant power excess on large scales. This was already noted in Planck Collaboration IV (2020), where it was shown that the same excess was present in realistic end-to-end simulations. The excess was found to be caused by joint processing of the two halves of the data split, which led to cross-split correlations.

In contrast, the two halves of the NPIPE detector split are processed independently, and, as a result, the cross-spectrum appears consistent with the ΛCDM prediction, without the need for further simulation-based interpretation. Indeed, NPIPE represents the first processing of the *Planck* data for which the large-scale reionization peak is visually apparent in the raw power spectrum, and not only statistically detected through high-level likelihood analysis. Likewise, the large-scale NPIPE BB spectrum is visually consistent with zero.

7.6. Astrophysical foreground maps

We now turn our attention to the astrophysical foreground maps resulting from the Commander analysis applied to the NPIPE data, starting with the temperature components. Figure 63 shows the various temperature amplitude maps included in the analysis. From top to bottom, these are the (1) the combined low-frequency power-law component evaluated at 30 GHz; (2) the thermal dust emission component evaluated at 545 GHz; and (3)–(5) CO $J = 1 \rightarrow 0$, $J = 2 \rightarrow 1$, and $J = 3 \rightarrow 2$ line emission. Figure 64 compares the CO $J = 1 \rightarrow 0$ and $J = 2 \rightarrow 1$ maps with the corresponding Dame et al. $J = 1 \rightarrow 0$ (Dame et al. 2001) and the *Planck* 2015 $J = 2 \rightarrow 1$ (Planck Collaboration X 2016) maps. Overall, NPIPE maps are in good agreement with previous results.

Figure 65 shows difference maps between the NPIPE and (1) the Commander 2015 and (2) the Generalized Internal Linear Combination (GNILC; Remazeilles et al. 2011) 2018 thermal dust amplitude maps. The latter provides an algorithmically independent and recent estimate of thermal dust emission, as discussed in Planck Collaboration IV (2020). In both cases, best-fit relative slopes and offsets have been accounted for. We see here that all three maps agree to high precision, with most residuals smaller than 0.01 MJy sr^{-1} at high Galactic latitudes and smaller than 1 MJy sr^{-1} in the Galactic plane. We further see that the

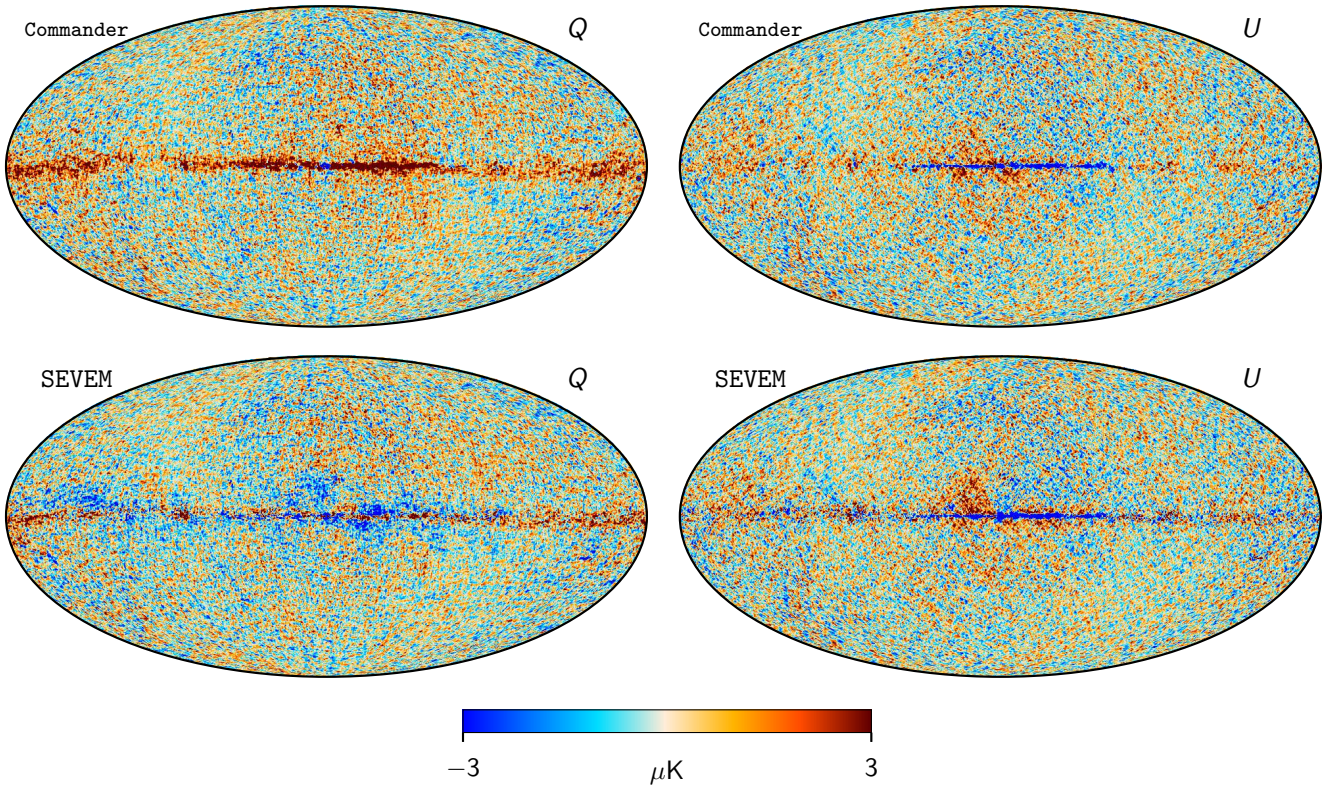


Fig. 56. CMB Q and U maps derived from the NPIPE data set with Commander (*top row*) and SEVEM (*bottom row*). All maps are plotted with an angular resolution of 40′ FWHM.

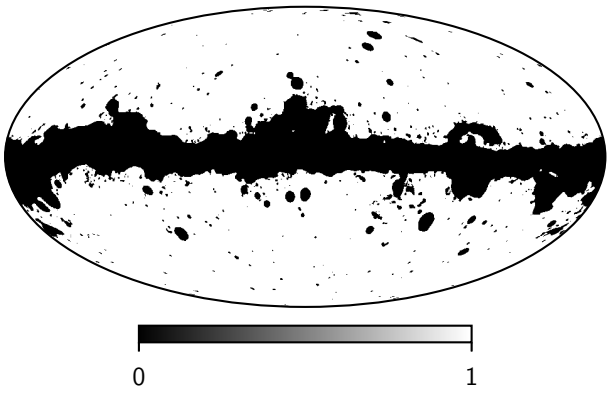


Fig. 57. NPIPE CMB temperature analysis mask. This mask is derived by thresholding the standard-deviation map evaluated among the Commander CMB temperature maps resulting from three independent *Planck* processings (*Planck* 2015, *Planck* 2018, and NPIPE), multiplied by another mask, which is thresholded on an overall χ^2 cut corresponding to the Commander NPIPE analysis.

difference between the NPIPE and *Planck* 2015 data sets is dominated by the bandpass leakage effect discussed in Sect. 7.3. For both *Planck* 2015 and GNILC, we see a weak imprint of zodiacal light aligned with the ecliptic plane, and for GNILC we additionally note a negative Galactic residual, consistent with the morphology of CO line emission.

Figure 66 shows T–T scatter plots between the NPIPE thermal dust amplitude map and three alternative thermal dust tracers. The top panel shows the correlation with respect to the HI4PI survey (Lenz et al. 2017) at low column densities, which provides the statistically most independent test of the derived amplitude map. This correlation is also used to determine the zero-level of the NPIPE 857-1 detector map (including all pixels with column densities up to $4 \times 10^{20}\,\mathrm{cm}^{-2}$ (Lenz et al. 2017), and thereby also in effect the overall zero-level of the NPIPE thermal dust component. As seen in Fig. 66, the relative residual offset between these two maps after final processing is $0.0079\,\mathrm{MJy\,sr}^{-1}$ according to a linear fit, which is negligible compared to intrinsic systematic uncertainties. For comparison, a quadratic fit (indicated by a red dashed curve) results in a relative offset of $0.06\,\mathrm{MJy\,sr}^{-1}$. The middle and bottom panels show similar correlation plots with respect to the *Planck* 2015 Commander and the *Planck* 2018 GNILC thermal dust amplitude maps. In both cases we observe very tight correlations. For the *Planck* 2015 map, the relative slope is 0.988, indicating that the high-frequency calibration of NPIPE and *Planck* 2015 agree to about 1%. This is reassuring, considering that NPIPE calibrates the 545 GHz channel on the orbital dipole and also adopts thermodynamic units of K_{CMB} at both 545 and 857 GHz, while *Planck* 2015 calibrated the 545 GHz channel with planets, and used flux density units of $\mathrm{MJy\,sr}^{-1}$ for the two highest frequency channels. For GNILC, we observe a relative slope of 0.887, which is simply due to the fact that no colour corrections were applied to the GNILC map, whereas the Commander maps are all measured relative to a sharp reference frequency of 545 GHz. Odegard et al. (2019) performed a re-calibration analysis of the *Planck* 2015 HFI maps based on COBE-FIRAS, obtaining results similar to those presented here. For instance, they derive a slope between *Planck* 545 GHz and HI4PI of 0.144 when adopting a threshold of $2.5 \times 10^{20}\,\mathrm{cm}^{-2}$ and (crucially) defining the *Planck* data in units of $\mathrm{MJy\,sr}^{-1}$, while for NPIPE we find a slope of 0.132. The resulting relative difference of 5%

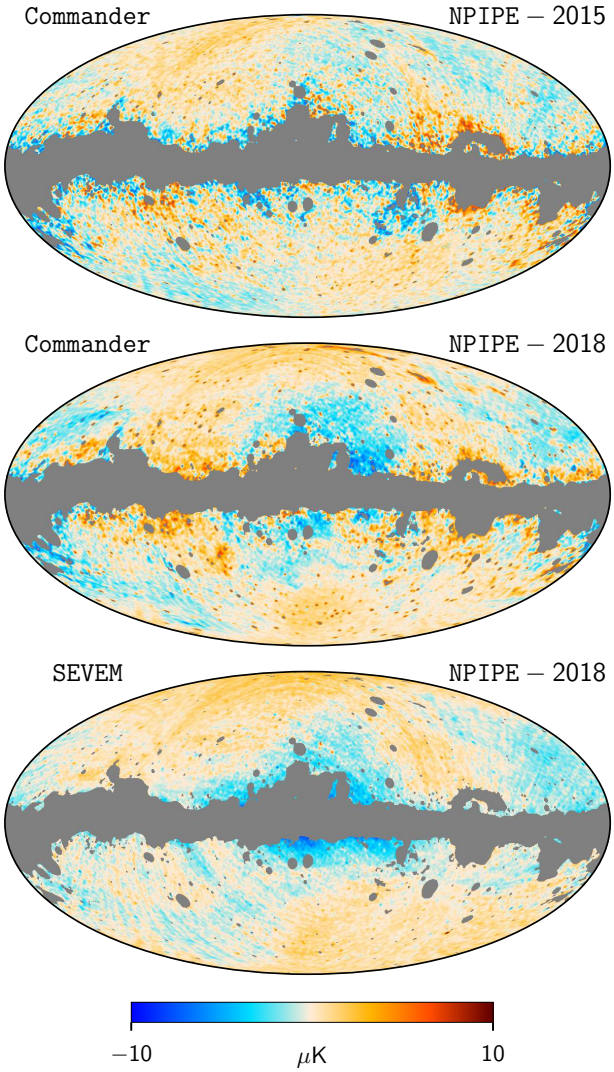


Fig. 58. *Top*: Stokes I difference map between the *Planck*-only NPIPE and the *Planck* 2015 (Planck Collaboration X 2016) Commander CMB maps. Both maps have been smoothed to 1° FWHM and masked with the mask shown in Fig. 57. *Middle*: same, but with the 2018 Commander CMB map (Planck Collaboration IV 2020). *Bottom*: same as middle panel, but evaluated for SEVEM instead of Commander.

is likely dominated by uncertainties in the 545 GHz bandpass profile, given that NPIPE 545 GHz is calibrated in thermodynamic units while Odegard et al. (2019) calibrate in flux density units.

Figure 67 compares T–T scatter plots between the three NPIPE and *Planck* 2015 CO line maps and the Dame et al. maps. Qualitatively, all estimates agree very well with each other, both visually in the maps and in terms of scatter plots. However, the NPIPE maps exhibit generally stronger correlations with respect to the Dame et al. survey than do the *Planck* 2015 maps. Specifically, the Pearson's correlation coefficients for *Planck* 2015 (for $1 \rightarrow 0$, $2 \rightarrow 1$, and $3 \rightarrow 2$) are $r = 0.988$, 0.981, and 0.905, respectively, while for NPIPE they are $r = 0.997$, 0.993, and 0.945. These improvements are due to a more fine-grained set of detector maps available in NPIPE than in *Planck* 2015, better control of instrumental systematic effects in high S/N regions, and better transfer-function and ADC corrections.

Next we consider the reconstructed polarized foreground emission. Figure 68 shows NPIPE synchrotron and thermal dust polarization amplitude maps, as well as corresponding difference maps with respect to *Planck* 2018, all plotted in brightness temperature units. These two data sets agree very well in terms of thermal dust emission at high Galactic latitudes, with most of the sky exhibiting differences smaller than $0.2\,\mu K_{RJ}$. The Galactic plane shows larger differences, with morphology similar to the absolute level of thermal dust emission, but with alternating sign along the plane. Such features typically arise from spatial variations in the thermal dust temperature or from different instrumental parameters in the form of detector polarization efficiency and angle. For synchrotron, larger relative differences are observed, both at low and high Galactic latitudes. NPIPE uses active polarization priors for the LFI frequencies, as was done in the HFI 2018 DPC processing, but different from what was done in the LFI 2018 processing. This approach has increased the overall level of polarization in the LFI sky maps, bringing them into better internal agreement among themselves, and also in better agreement with the WMAP observations (see Figs. 50 and 51).

Figure 69 compares the thermal dust polarization fraction as estimated from the NPIPE data set with Commander (top panel) and as estimated from the *Planck* 2018 data set with GNILC (middle panel). The bottom panel shows their difference. Overall, we see that these data sets agree exceedingly well in terms of polarization fractions, with low latitude differences being much smaller than 1%, and high-latitude differences being dominated by noise-like features.

7.7. Assessment of uncertainties

We conclude this section with an assessment of uncertainties in the derived products, for which we adopt two fundamentally different types of estimates. The first type of estimate is based on simulations. As described in Sect. 5, the NPIPE data set is accompanied with a set of 600 complete end-to-end-simulations, and each of these is propagated through the Commander analysis described above. To ensure that the resulting noise levels match the true data set, spectral indices (i.e., mixing matrices) are fixed at the values derived from the real observations, and only amplitudes are fitted freely; this is the same approach as used for the *Planck* 2013, 2015, and 2018 data releases (Planck Collaboration XII 2014; Planck Collaboration IX 2016; Planck Collaboration IV 2020).

Figure 70 shows the difference between a single foreground-cleaned Commander CMB temperature simulation and the corresponding true input CMB realization, smoothed to a resolution of 1° FWHM. The grey region indicates the NPIPE temperature-analysis mask defined in Fig. 57. For most of the sky, we see that the reconstruction error is less than about $3\,\mu K$, with slightly higher values near the edge of the mask, in particular close to regions with bright free-free emission. For most cosmological analyses, these errors are very small compared to the CMB fluctuations, which have a standard deviation of about $70\,\mu K$ on these angular scales. The Commander temperature analysis for NPIPE relies on a fine-grained detector-map analysis, and the computational cost of producing such simulations is very high. At the time of publication of this paper, 100 such simulations are available.

Figure 71 shows a similar comparison for Stokes Q and U. The top panels show the reconstructed, foreground-cleaned, CMB Q and U maps. The middle panels show the input CMB maps. The bottom panels show the differences. At a visual level, the simulated process of measuring the sky, processing with NPIPE and fitting and removing foregrounds, provides a noisy

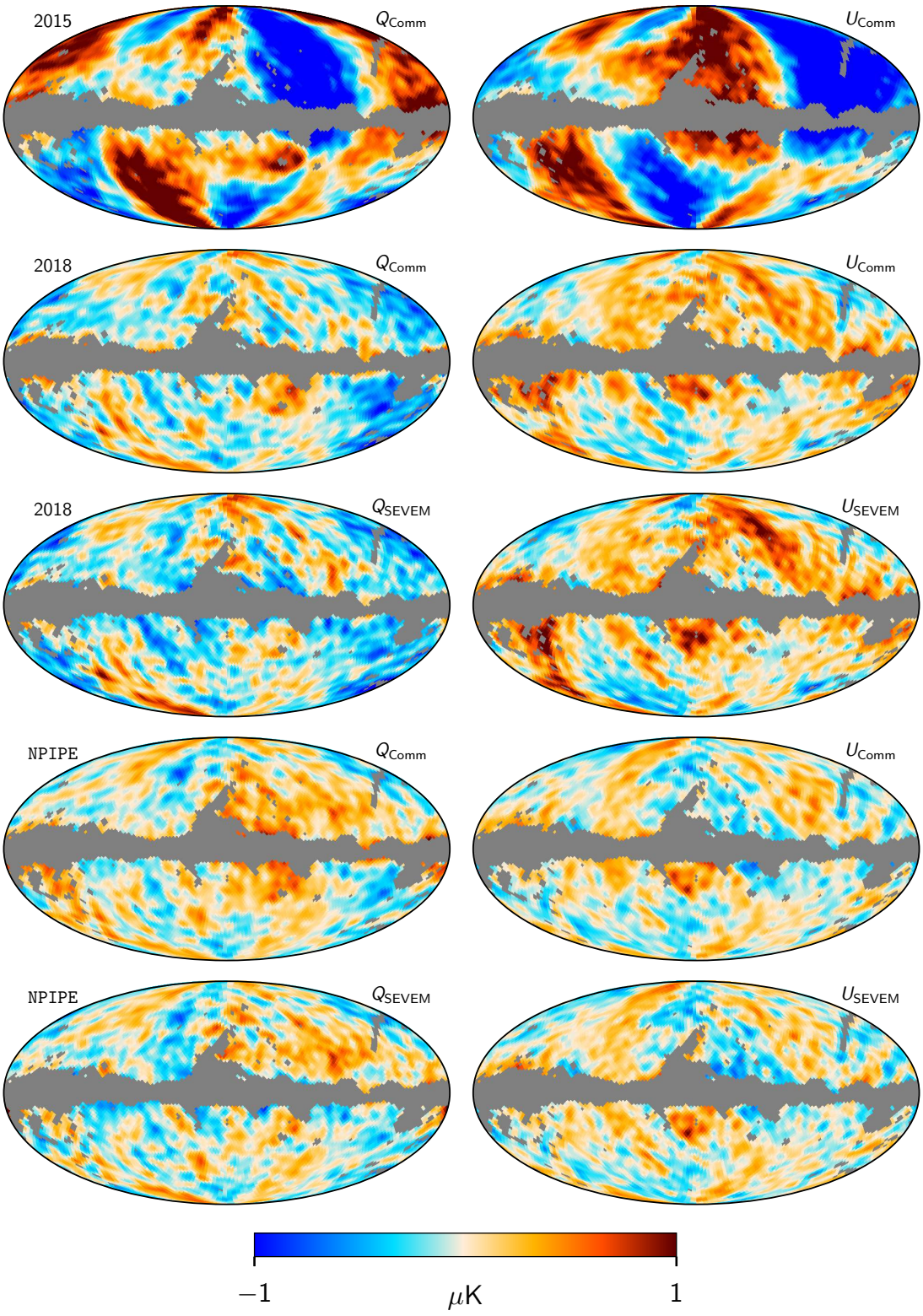


Fig. 59. Comparison of large-scale CMB Q and U maps from, *top to bottom*: Commander *Planck* 2015; Commander *Planck* 2018; SEVEM *Planck* 2018; Commander NPIPE; and SEVEM NPIPE. Note that the large-scale *Planck* 2015 CMB map in the *top row* was never publicly released, due to the high level of residual systematic effects. The grey region corresponds to the *Planck* 2018 common component-separation mask (Planck Collaboration IV 2020). All maps are smoothed to a common angular resolution of 5° FWHM.

image of the CMB polarization sky, with significant systematic errors present both at low and high Galactic latitudes. However, it is evident, even at a visual level, that the NPIPE maps provide a clear tracer of true large-scale CMB features.

Figure 72 shows the standard deviation of the foreground-cleaned Commander CMB polarization simulations at 1° FWHM resolution. We see that the overall noise standard deviation varies between 0.2 and 0.6 μK at high Galactic latitudes, with a spatial

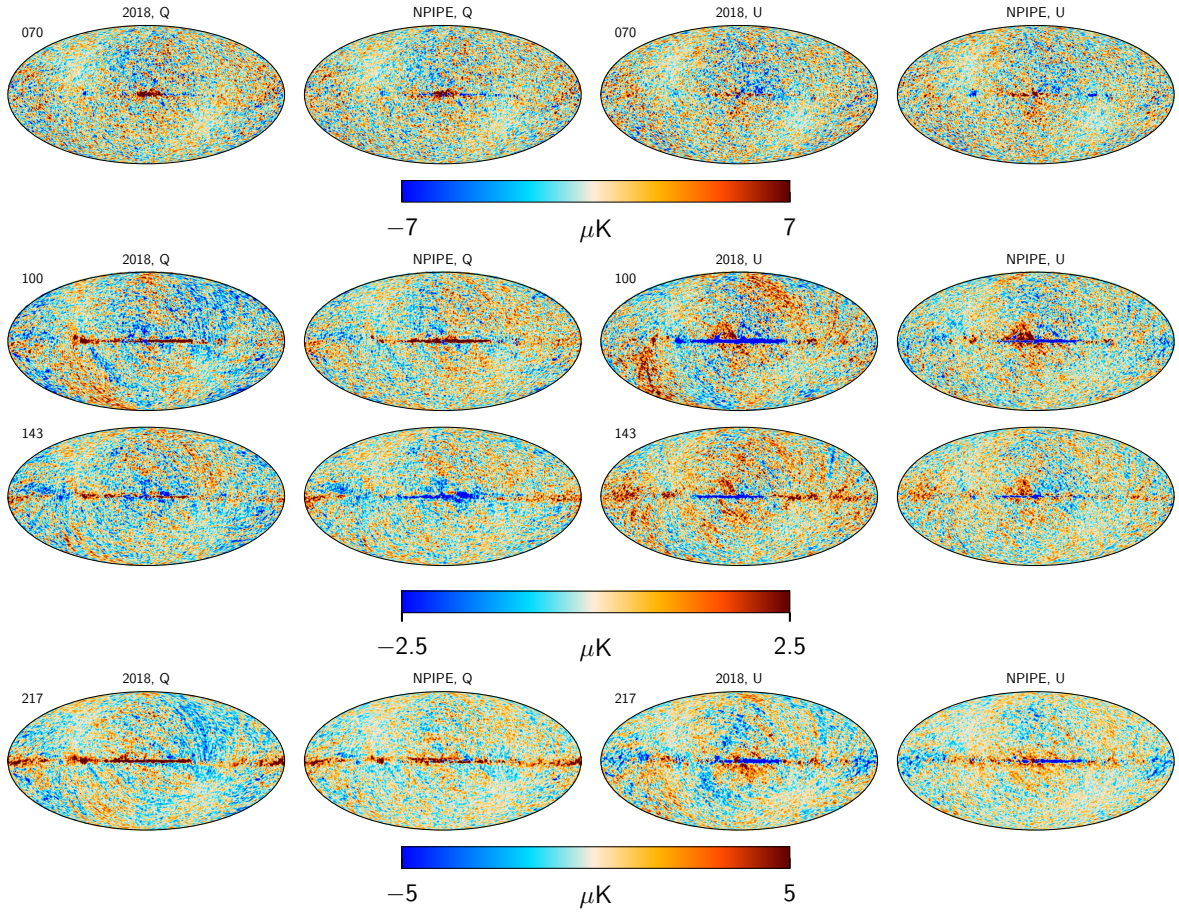


Fig. 60. Comparison of single-frequency polarization maps derived with SEVEM from *Planck* 2018 and from NPIPE. All maps are smoothed to a common angular resolution of 80′ FWHM.

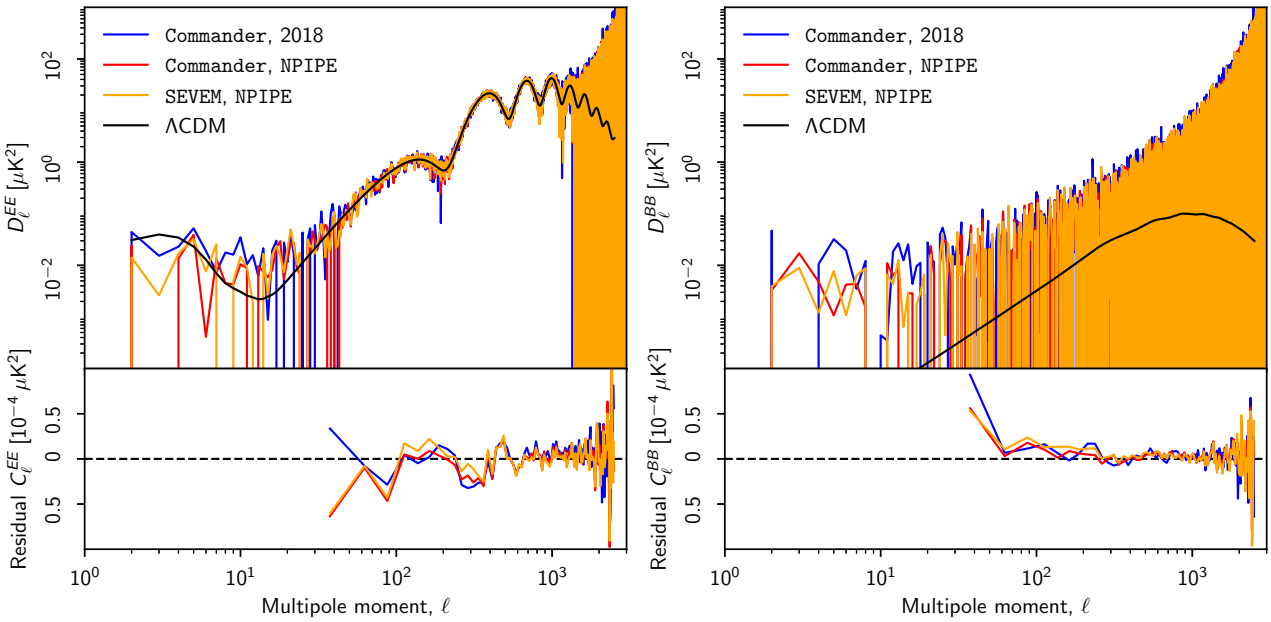


Fig. 61. Angular CMB polarization cross-power spectra evaluated from the *Planck* 2018 (blue curves) and NPIPE (red curves for Commander; orange curves for SEVEM) data sets. *EE* and *BB* spectra are shown in the left and right panels, respectively. Within each main panel, the full spectrum is shown in the top sub-panel, while the residuals with respect to the *Planck* 2018 best-fit ΛCDM spectrum (black curves) are shown in the bottom sub-panels. The latter have been binned with $\Delta\ell = 25$. The cross-spectra are evaluated from the most independent data split that is available for each data set, corresponding to the A/B detector split for NPIPE and the half-mission split for *Planck* 2018.

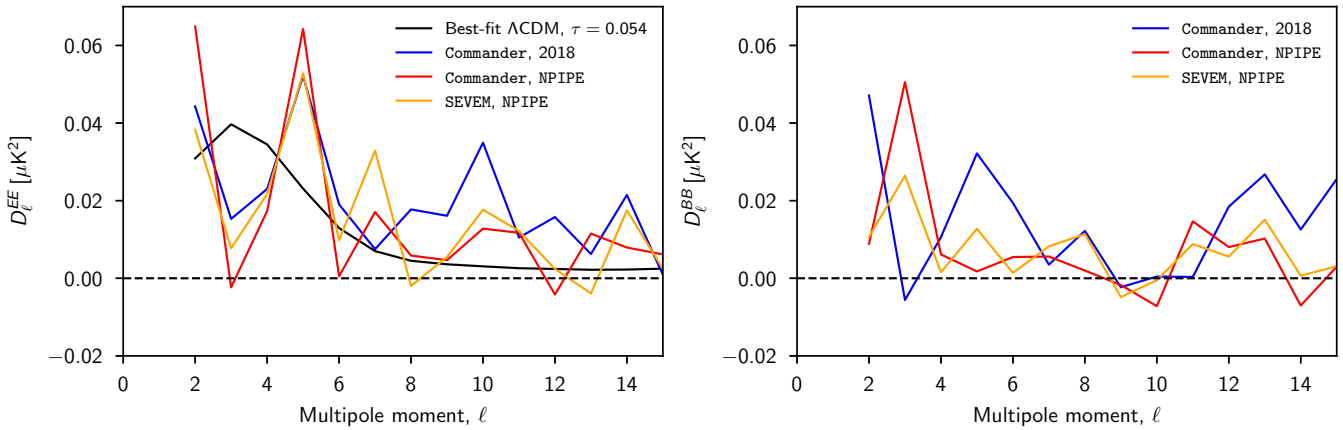


Fig. 62. Same as Fig. 61, except showing only the lowest multipoles. Corrections for the low-ℓ NPIPE transfer function have been applied to both sets. Note that neither of the spectra has been corrected for potential biases from correlations arising due to joint processing of the two data splits. As shown in Planck Collaboration III (2020) and Planck Collaboration XI (2019), these correlations are significant for the 2018 half-mission data split, and reproducible in end-to-end simulations.

distribution dominated by the scanning pattern of the *Planck* telescope. At low Galactic latitudes, higher values are observed, due to the combination of additional foreground uncertainties and instrumental systematic effects, particularly in the form of intensity-to-polarization leakage.

Figure 73 shows the fractional difference between angular power spectra computed from the observed foreground-cleaned Commander and SEVEM CMB polarization maps, and the mean of the simulations[16], specifically

$$\eta_\ell \equiv \frac{D_\ell^{\rm data} - \left\langle D_\ell^{\rm sim} \right\rangle}{\left\langle D_\ell^{\rm sim} \right\rangle}, \tag{23}$$

as computed with PolSpice (Chon et al. 2004) over the *Planck* 2018 common polarization mask. Blue and red curves show NPIPE results with and without noise alignment (i.e., rescaling of the noise), respectively, while grey curves show similar results for *Planck* 2018, also with noise alignment, as reproduced from Fig. 12 in Planck Collaboration IV (2020).

Overall, we see that the NPIPE polarization simulations without noise adjustment (red curves) agree with the observed data to about 1% in power, which is better than the precision of the *Planck* 2018 simulations *after* noise adjustment (grey curves). However, it is important to note that the raw NPIPE simulations *under-estimate* the total power in the real data, as opposed to the *Planck* 2018 simulations, which *over-estimate* the total power level. This difference is important, because it allows a straightforward path to accurate noise readjustment, simply by adding slightly more noise. Accordingly, a second set of NPIPE simulations is also provided, for which the power deficit has been corrected by the addition of scale-dependent noise. The blue curves in Fig. 73 show the corresponding power consistency after accounting for this extra power. Following this noise addition, we find excellent agreement between the observed NPIPE data and simulations.

Figure 74 shows a simulation-based comparison for the thermal dust polarization amplitude. In this case, we estimate the square of the polarization amplitude by cross-correlating the reconstructed A- and B-split thermal dust amplitude maps at 353 GHz, smooth this to an effective angular resolution of 3° FWHM, and divide by the corresponding square of the simulation input thermal dust amplitude. To reduce noise, we average over 300 simulations, before finally computing the square root to obtain an estimate of the linear polarization amplitude. The grey lines in Fig. 74 indicate a polarization amplitude of $3\,\mu{\rm K_{RJ}}$ at 353 GHz, which corresponds roughly to a signal-to-noise ratio of unity. Overall, we see that wherever the signal-to-noise ratio is significant, the reconstructed thermal dust amplitude is unbiased to $\lesssim$1%, while where the signal-to-noise ratio is less than unity, biases up to 5–10% may be observed, corresponding to an absolute error of $\lesssim 0.3\,\mu{\rm K_{RJ}}$. This value thus defines a systematic uncertainty for the NPIPE thermal dust polarization map at 353 GHz, and takes into account the full end-to-end processing, including calibration, mapmaking and component separation.

The second type of estimate that we use to assess uncertainties is based on detector-split differences. Specifically, each half of the NPIPE detector-split is processed through the Commander analysis framework, establishing two independent estimates of the same quantities. The resulting half-difference amplitude maps are shown in Fig. 75. These provide a direct view of the overall instrumental noise level in each component map, and residual systematics. For instance, while the low-frequency component is clearly dominated by instrumental noise at high Galactic latitudes, the thermal dust amplitude is equally clearly dominated by systematic effects, particularly in the form of calibration uncertainties. Most systematic features seen in these maps can be matched to the instrumental effects discussed in Sect. 7.3.

8. CMB Solar dipole

As discussed in Sects. 2 and 7, the NPIPE maps retain the Solar dipole (only the orbital CMB dipole is removed prior to mapmaking). This makes it possible to include a dipole in the CMB term (Eq. (19)) of the temperature model in Sect. 7.1, and therefore to fit the Solar dipole simultaneously with the global monopole, calibration, and bandpass shift factors given in Table 10 and the foreground and CMB maps, as discussed in Sect. 7. Thus, for the first time, a *Planck* Solar dipole based on data from both LFI and HFI can be determined, with uniform determination of uncertainties.

[16] Note that there are differences between the Commander and SEVEM simulations. While the Commander simulations are a foreground-cleaned version of each realization, the SEVEM ones are generated by propagating CMB plus noise simulations through the pipeline.

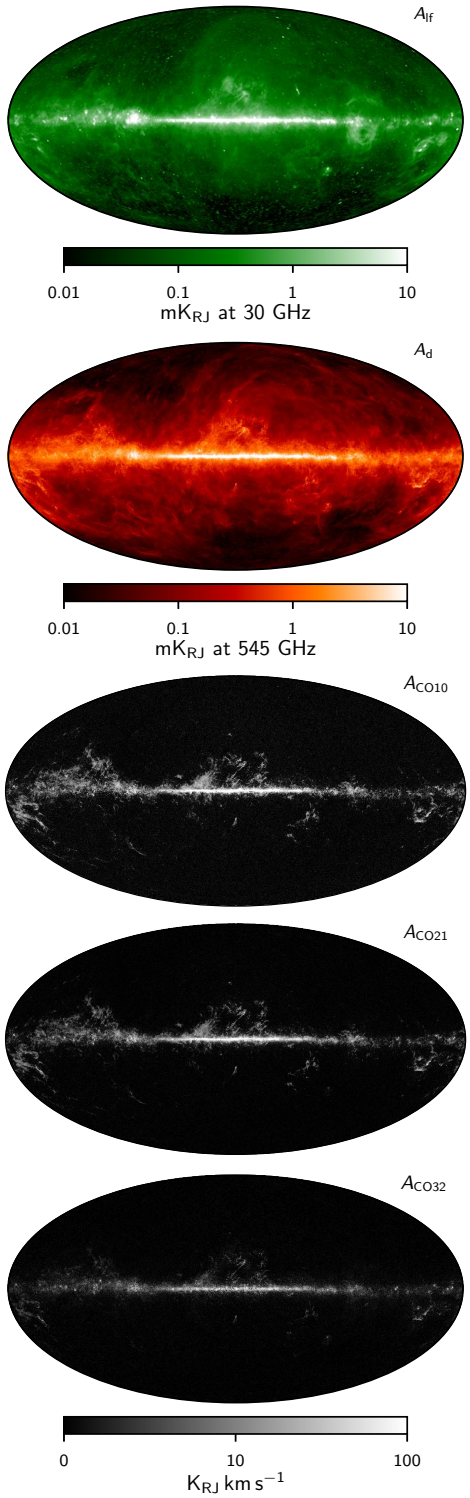


Fig. 63. *Top to bottom*: low-frequency (evaluated at 30 GHz, smoothed to 40′ FWHM), thermal dust (evaluated at 545 GHz, smoothed to 14′ FWHM), and CO intensity maps.

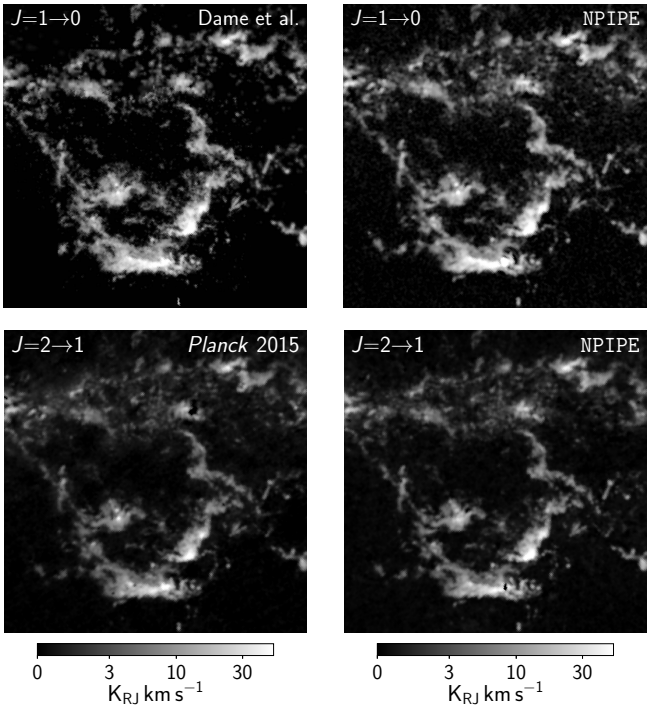


Fig. 64. 30° × 30° expansion of various CO emission line maps. All maps are smoothed to 14′ FWHM and are centred on the Orion region, with Galactic coordinates (l, b) = (210°, −9°).

The task of determining the CMB Solar dipole parameters is equivalent to fitting a dipole to the CMB map shown in the top panel of Fig. 55. As usual, regions of brightest Galactic emission must be excluded from the fit. And as with any such fit on a partial sky, care must be taken to avoid confusion from higher-order CMB moments. Our approach to this problem was described in general form by Jewell et al. (2004), Wandelt et al. (2004), and Eriksen et al. (2004), and later implemented by Hinshaw et al. (2009) for WMAP. The algorithm has been explored in detail by Thommesen et al. (in prep.) for application to the *Planck* observations, and the following description and results are based on the Thommesen et al. (in prep.) implementation.

The algorithm is a variation of the Gibbs-sampling method described in Sect. 7 and implemented in `Commander`. In addition to the parametric setup described in Sect. 7, we assume that the CMB fluctuations constitute an isotropic, Gaussian, random field. Under this mild assumption, whose validity to high precision is confirmed by the entire body of *Planck* results, the CMB field is described by an angular power spectrum, C_ℓ. Combined with the observed phase information of the non-masked CMB fluctuations, this power spectrum then acts as an informative prior for masked regions. Specifically, the large-scale structures can be partially reconstructed from the observed structures outside the mask, given the requirement that the total field must be Gaussian and isotropic. This intuitive idea may be formulated quantitatively in terms of the following equation for Gaussian constrained realizations,

$$(\mathsf{S}^{-1} + \mathsf{N}^{-1})\boldsymbol{x} = \mathsf{N}^{-1}\boldsymbol{d} + \mathsf{S}^{-1/2}\omega_1 + \mathsf{N}^{-1/2}\omega_2, \tag{24}$$

where $\mathsf{S} = \mathsf{S}(C_\ell)$ is the signal covariance matrix defined by the angular power spectrum, N is the noise covariance matrix, in which the variance of masked pixels is set to infinity, $\boldsymbol{d}$ is the (foreground-cleaned, but noisy) CMB map, and ω_1 and ω_2 are two random Gaussian fields, with zero mean and unit variance. The solution $\boldsymbol{x}$ is known as a "constrained realization", and typically a large number of such samples is produced in order to quantify the uncertainties due to the sky cut and noise. Furthermore, the power spectrum C_ℓ is in principle unknown, and in practice one therefore computes a Gibbs chain to produce final results, iterating between determining constrained realizations

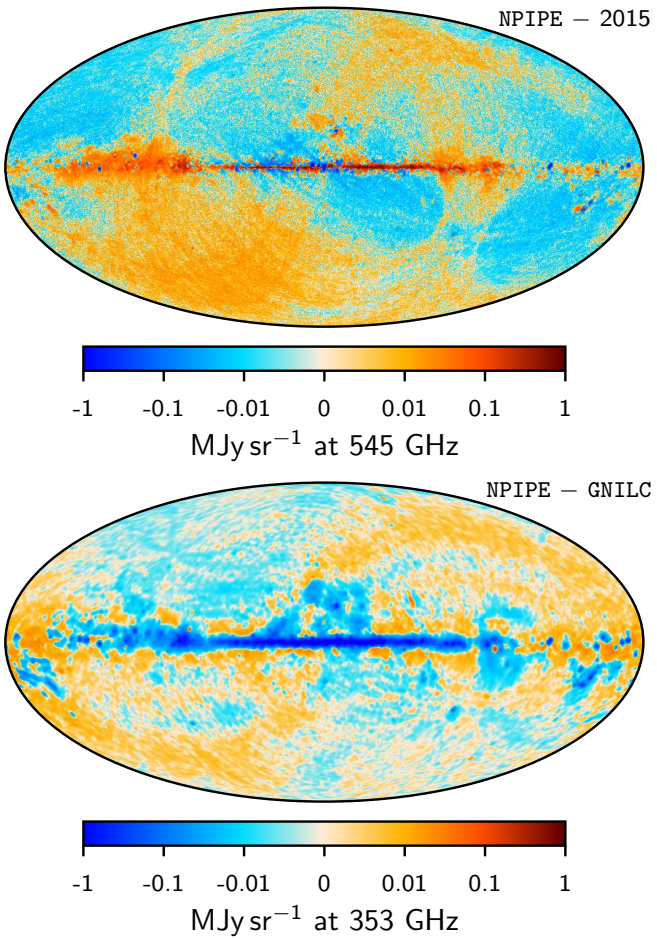


Fig. 65. *Top:* Stokes I difference map between the NPIPE and the *Planck* 2015 (Planck Collaboration X 2016) dust amplitudes, adjusted for the scaling and offset seen in Fig. 66. Both dust amplitude maps are smoothed to 1° FWHM and pixelized with a HEALPix resolution parameter $N_{\rm side} = 256$. *Bottom*: similar difference plot between the NPIPE dust amplitude map and the GNILC (Planck Collaboration IV 2020) dust amplitudes, adjusted for the scaling and offset seen in Fig. 66. Both dust amplitude maps were smoothed to 80′ FWHM and pixelized with a HEALPix resolution parameter $N_{\rm side} = 256$.

and angular power spectra. For further details, see Eriksen et al. (2004) and Thommesen et al. (in prep.).

Once the masked region is filled in by this constrained realization, we have a *full-sky* CMB map that is consistent with the observed data, from which the dipole can be computed directly without reference to a sky mask. The statistical uncertainty introduced by noise and this process of dealing with a masked sky is well-represented by the standard deviation measured from the ensemble of constrained realizations. In the following, we report values averaged over 100 Gibbs samples per mask.

In addition to the statistical uncertainties just described, there are three main sources of systematic uncertainty that we now estimate. The first is due to uncertainties in modelling component separation. We estimate this using the same approach used for other global parameters described in Sect. 7.4. Specifically, we quantify the typical scatter seen among various internal NPIPE versions, analysed with different foreground models. We estimate the 1 σ systematic uncertainties to be 1 μK for the CMB dipole amplitude, 1.′8 for the Galactic latitude, and 1.′3 for the Galactic latitude.

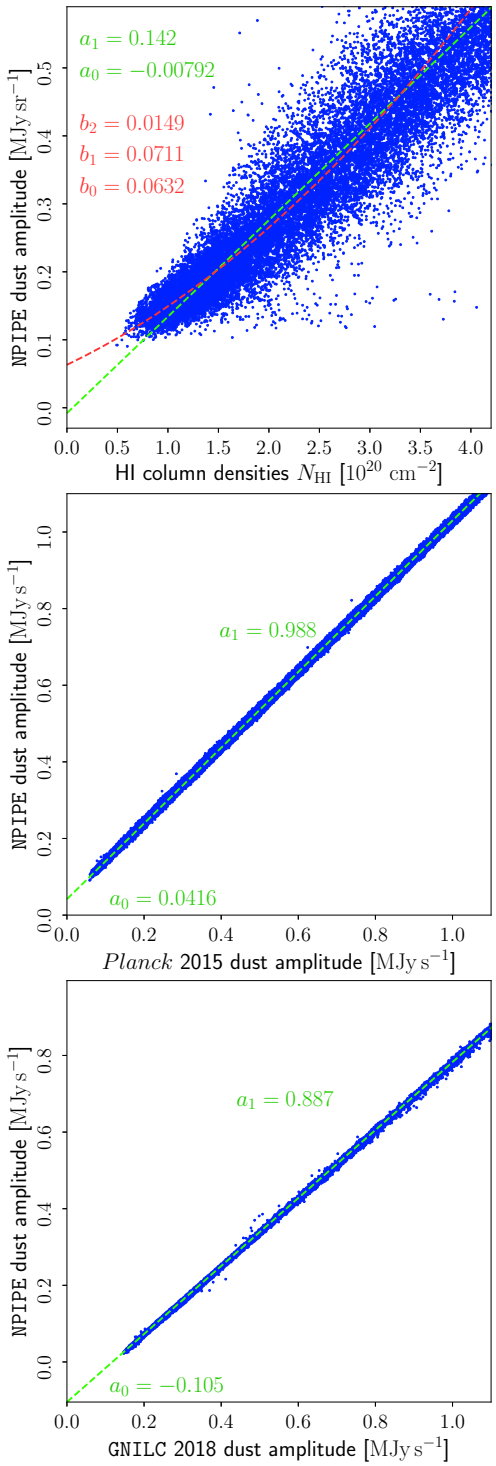


Fig. 66. T–T scatter plots between the NPIPE dust amplitude map at 545 GHz and three alternative thermal dust estimates. The panels show, *from top to bottom*, correlations with respect to: the HI4PI survey (Lenz et al. 2017); the *Planck* 2015 thermal dust amplitude map; and the GNILC 2018 thermal dust amplitude map. The *top and middle panels* show maps smoothed to a common resolution of 60′ FWHM, while the *bottom panel* employs a smoothing scale of 80′ FWHM, determined by the resolution of the GNILC map. The numbers marked by b_2, b_1, and b_0 (first plot) correspond to the best-fit polynomial parameters of a quadratic model (red dashed line, b_0 being the intercept), while a_1 and a_0 (all plots) correspond to the slope and intercept of the best fit line (green dashed line) through each distribution of points.

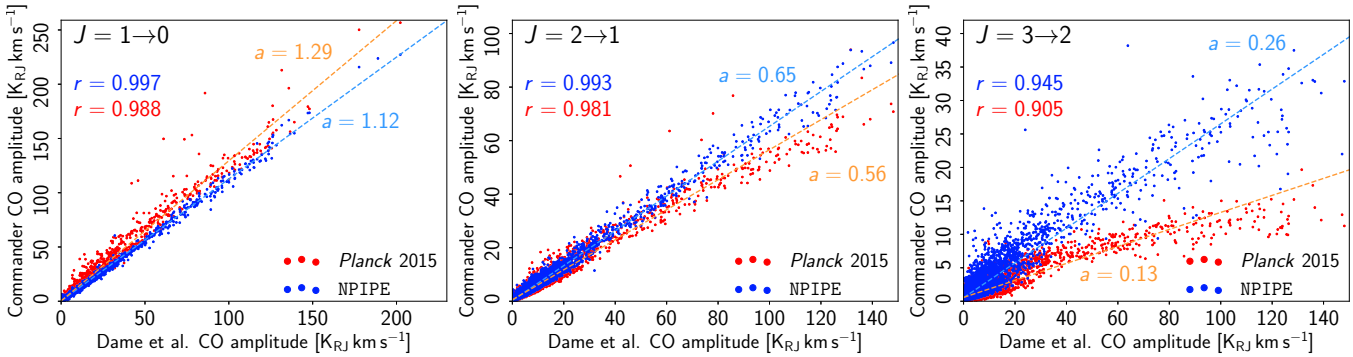


Fig. 67. Comparison of T–T scatter plots evaluated between the Dame et al. CO map (Dame et al. 2001) and the Commander NPIPE (blue) and *Planck* 2015 (red) CO maps. Panels show, *from left to right*, results for the CO $J = 1 \rightarrow 0$, $J = 2 \rightarrow 1$, and $J = 3 \rightarrow 2$ line maps, respectively. All maps have been smoothed to 1° FWHM, and pixelized with $N_{\rm side} = 64$. The parameter marked by "a" is the best-fit linear slope of the scatter plot including values between 0.01 and 150 $\mathrm{K_{RJ}\,km\,s^{-1}}$ of the Dame et al. $J = 1 \rightarrow 0$ map. The parameter r is the Pearson's correlation coefficient between Dame et al. and the respective CO amplitudes.

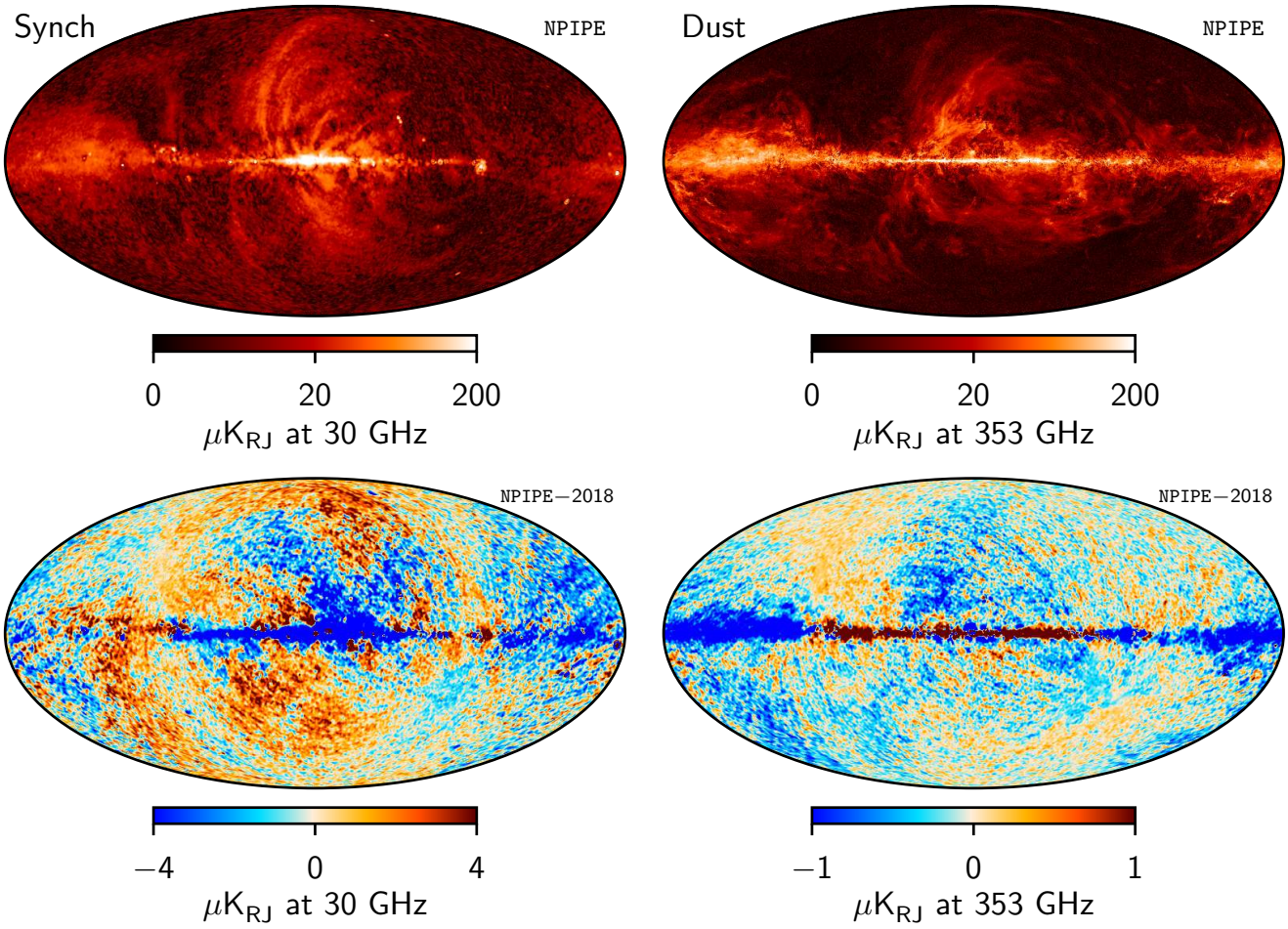


Fig. 68. *Top*: synchrotron and thermal dust polarization-amplitude ($P = \sqrt{Q^2 + U^2}$) maps derived from the NPIPE data set. The synchrotron map and thermal dust maps are smoothed to angular resolutions of 40′ and 5′ FWHM, respectively. *Bottom*: corresponding polarization-amplitude difference maps taken between the NPIPE and *Planck* 2018 component maps. Both maps are smoothed to a common resolution of 60′ FWHM. The top panels use the nonlinear *Planck* colour scale, while the bottom panels use linear colour scales.

The second source of systematic uncertainty is due to the uncertainty in absolute calibration between channels. Specifically, the Commander analysis presented above adopts the 143 GHz channel as its calibration reference channel, based on the fact that it has the lowest absolute noise of any *Planck* frequency and no CO contamination; as a result, component separation is generally more robust for this channel. An equally valid choice, however, might have been the 100 GHz channel, and this channel was indeed adopted for the same purpose by the corresponding HFI DPC analysis. However, as shown in Table 10, there is an overall 0.07% relative calibration difference between the 100 and 143 GHz channels. Given an overall

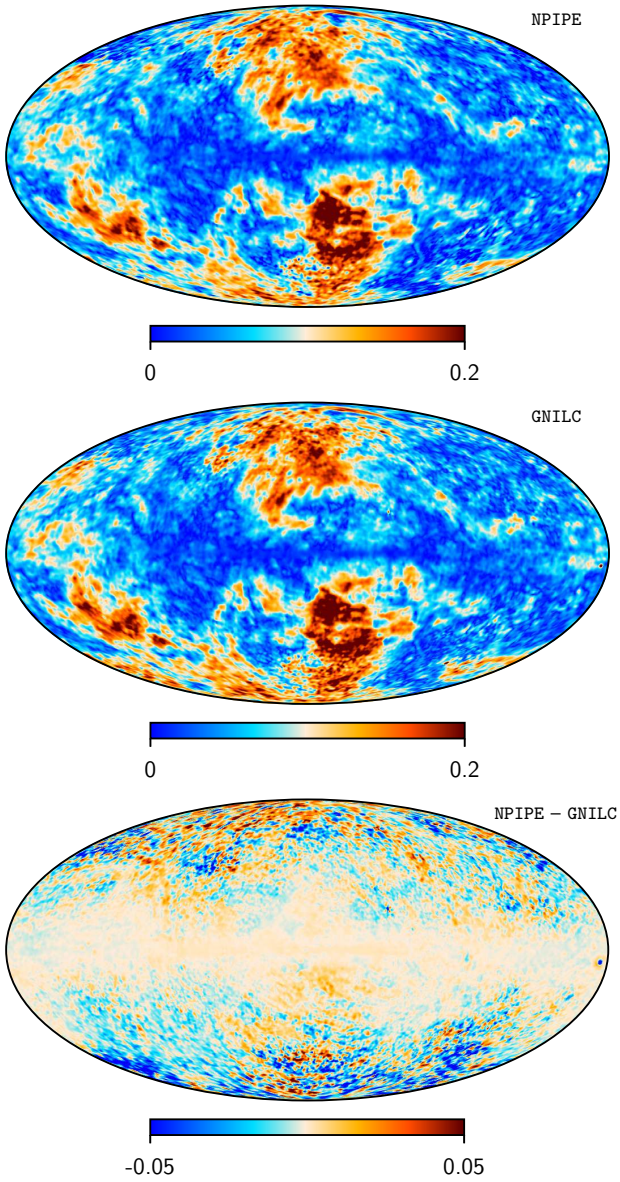


Fig. 69. Comparison of thermal dust polarization-fraction maps, as evaluated from the NPIPE (*top*) and GNILC 2018 (*middle*) data sets, and a difference map (*bottom*) between the two. All maps are smoothed to 80′ FWHM. We have subtracted a monopole of 389 μK from the GNILC intensity map (Planck Collaboration XI 2019).

dipole amplitude of about 3360 μK, this translates directly into a shift of 2.4 μK in the overall dipole amplitude, depending on which channels is chosen as the reference. We do not have any strong reason to prefer one channel over the other, so we adopt this value as another systematic error on the overall dipole amplitude.

The third term is due to the uncertainty in the CMB monopole value, for which we adopt $T_0 = 2.72548 \pm 0.00057$ K (Fixsen 2009). For a dipole value of 3.3 mK, this monopole uncertainty translates into a dipole amplitude uncertainty of 0.7 μK. Taking all four terms together, we therefore end up with a total dipole amplitude uncertainty of $(0.25^2 + 1^2 + 2.4^2 + 0.7^2)^{1/2}\,\mu\mathrm{K} = 2.7\,\mu\mathrm{K}$. The relative calibration uncertainty between 100 and 143 GHz turns out to be the dominant uncertainty in the dipole amplitude.

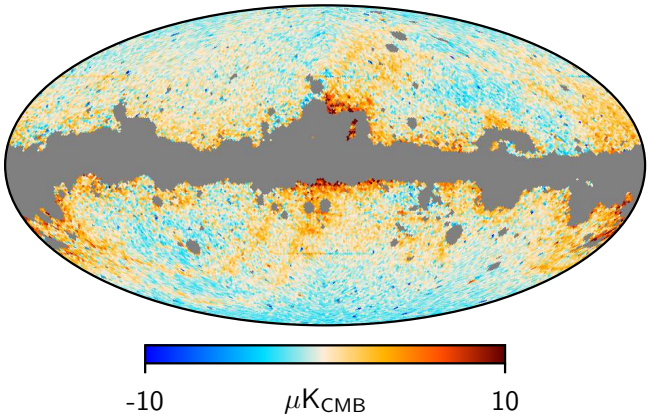


Fig. 70. Stokes I difference map between the Commander NPIPE CMB map of simulation No. 200 and the corresponding input CMB map of the simulation. Both maps are pixelized with a HEALPix resolution parameter $N_{side} = 64$. The map has been masked with the mask shown in Fig. 57.

Figure 76 shows results obtained from applying this algorithm to the Commander-based NPIPE map shown in Fig. 55 as a function of sky fraction. The series of masks adopted for this analysis is the same as was employed by the HFI-DPC analyses presented in Planck Collaboration III (2020). The NPIPE posterior mean values are shown as solid black lines, while the 1 σ posterior confidence regions are shown as grey regions. For comparison we also plot previously published results as coloured points[17].

The NPIPE confidence bands show good stability on sky fractions ranging from $f_{sky} = 0.20$–0.95, suggesting that the foreground removal process has been successful. We also observe good qualitative agreement between the estimates derived from different experiments and data sets. The agreement is particularly striking between the LFI and the NPIPE results, with less than 1 σ shifts in any of the three parameters, despite the fact that the NPIPE CMB map is strongly dominated by HFI observations. The WMAP data also agree well with both of these, although exhibiting a slightly lower (1.4 σ in units of the WMAP uncertainty) CMB dipole amplitude.

As final NPIPE dipole estimates, we adopt the values derived for a sky fraction of 81%. As seen in Fig. 76, the error bars do not decrease further for higher sky fractions, since the fixed systematic uncertainty contribution starts to dominate the error budget. For convenient reference and comparison, these values are tabulated in Table 11, together with the previously published estimates shown in Fig. 76.

Statistically speaking, the most striking feature seen in these plots is the apparent qualitative disagreement between the NPIPE and HFI DPC uncertainties. In particular, the HFI directional uncertainties are nominally between 3 and 10 times smaller

[17] The HFI measurements are formally assigned a high sky fraction of about 96%, but a precise specification of this sky fraction is difficult to make. HFI adopted an approach in which foreground-reduced *Planck* CMB fluctuation maps (Planck Collaboration IV 2020) were subtracted from each frequency map prior to the dipole estimation process. This procedure is inherently somewhat circular, since the dipole of the CMB maps itself must be adjusted prior to subtraction; however, the small adjustment to the dipole can be assessed using a smaller mask. Nevertheless, this approach is fundamentally different from approaches that apply hard masks before the dipole analysis (e.g., Lineweaver et al. 1996). Due to the use of the strong CMB prior, sky fractions given in Fig. 76 are suggestive, not definitive.

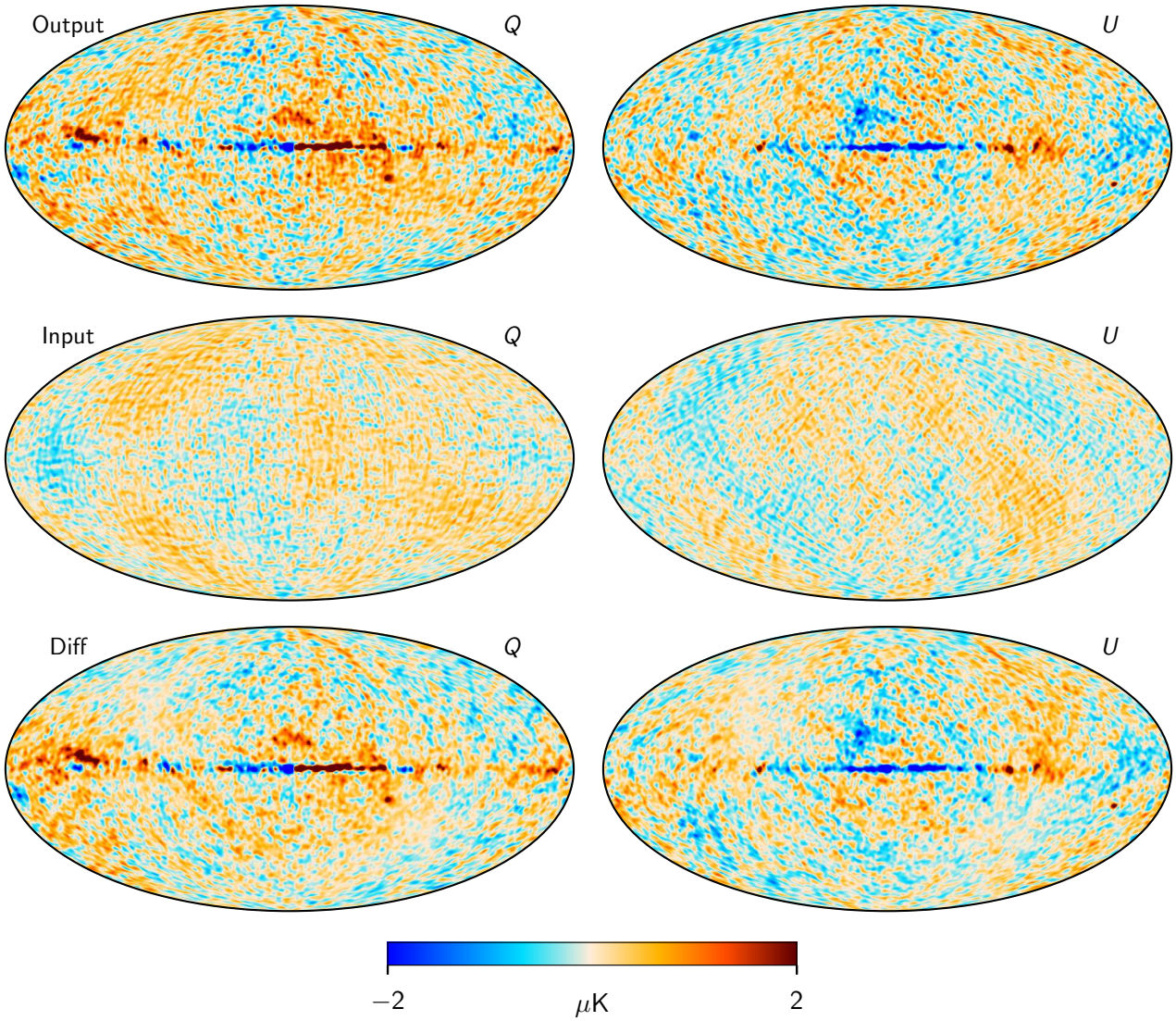


Fig. 71. Comparison of end-to-end reconstructed (*top row*) and input (*middle row*) NPIPE simulations for the Stokes Q and U CMB maps. The *bottom row* shows the difference between the output and input sky maps. All maps are smoothed to a common angular resolution of 2° FWHM.

than the NPIPE directional uncertainties, and the HFI 2015 and 2018 longitude estimates disagree at the $8\,\sigma$ level. Part of this is explained by the fact that the HFI directional uncertainties do not include full estimates of systematic errors. In addition, as stated above, one of the main algorithmic differences between the HFI DPC and NPIPE analysis is that in the HFI analysis, *Planck* maps of the CMB were subtracted from the sky maps prior to estimation of the CMB dipole. This introduces an uncertainty that is perfectly correlated across the entire sky, and trades a large amount of statistical uncertainty from higher-order CMB fluctuations against a systematic error that is difficult to quantify in terms of the ability of the component separation algorithms to remove foregrounds in the central Galactic plane. To illustrate this point, Fig. 77 shows the NPIPE dipole amplitude as a function of sky coverage, where the dipole has now been estimated with a similar methodology as in Planck Collaboration III (2020), i.e., by first subtracting a *Planck* CMB temperature map prior to estimating the dipole parameters, and not imposing a constrained realization in the masked region. (The coloured curves are direct reproductions from Fig. 23 in Planck Collaboration III 2020.) Here we see that the NPIPE dipole is in fact nominally stable to a precision smaller than $0.5\,\mu$K when adopting this approach, which is comparable to the 100 GHz HFI DPC result. This demonstrates that the variations seen in the top panel of Fig. 76 are not due to, say, residual foreground fluctuations in the NPIPE approach, but rather due to the fact that the CMB dipole is estimated independently over each considered sky fraction. Correspondingly, the apparent stability seen in Fig. 77 (and Fig. 23 in Planck Collaboration III 2020) is due to the fact that the CMB fluctuations have been fixed based on 95% of the sky for all cases, and not only the nominal $f_{\rm sky}$ value indicated in the plot. With these important algorithmic points in mind, we conclude that the NPIPE and HFI DPC dipole amplitudes listed in Table 11 agree well within statistical uncertainties. The NPIPE uncertainties are somewhat more conservative than those presented in Planck Collaboration III (2020).

9. Optical depth to re-ionization

In this section we use determination of the optical depth to reionization, τ, as a validation of the quality of the large-scale

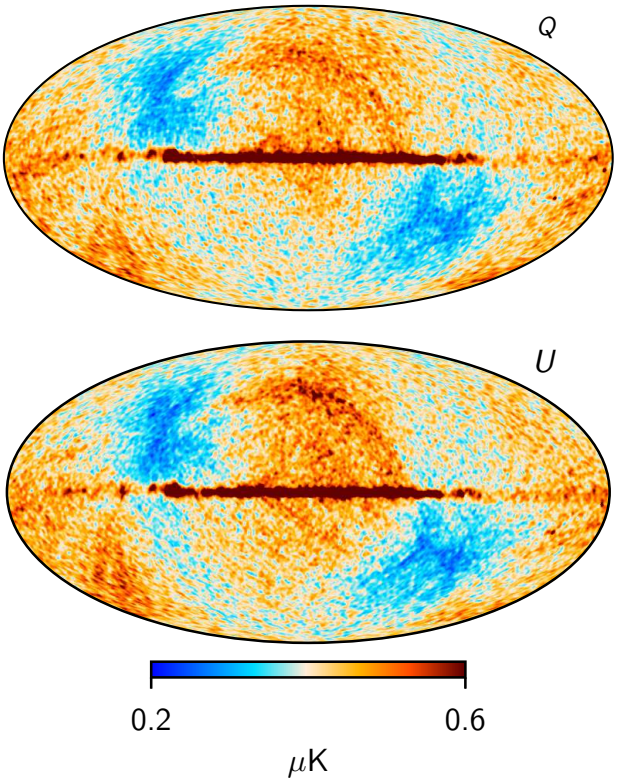


Fig. 72. Standard deviation evaluated from 100 end-to-end NPIPE full-mission simulations of CMB Q and U maps, as derived with Commander. Both maps are smoothed to 2° FWHM.

polarization in the NPIPE maps. Considering quadratic maximum-likelihood (QML) approaches, we use both auto-QML and cross-QML methods to extract the large-scale EE power spectrum, and then derive posterior distributions for τ. Simulations are used to infer the statistical and systematic uncertainties of our methods, and to demonstrate that the large-scale polarization transfer function (Sect. 4.3) is accounted for properly. As discussed, e.g., in Planck Collaboration V (2020) and Planck Collaboration VI (2020), estimated τ values depend on specific details of the analysis setup, including which part of the data is considered (e.g., EE-only or $TT-TE-EE$), whether additional cosmological parameters (e.g., A_s, r) are fixed to a reference value or marginalized over, and whether we model reionization as a sharp or an extended process. A full assessment of these effects for NPIPE maps is beyond the scope of this paper. Rather, here we focus on the consistency of estimated quantities (τ, but also foreground template amplitudes, power spectra, cleaned map goodness-of-fit) between different data cuts (e.g., frequencies, sky masks) as a measure of the cleanliness of the large-scale polarization of NPIPE maps.

9.1. Pixel-based analysis

The approach used for the pixel-based analysis of NPIPE maps is similar to the one used for the analysis of low-resolution LFI polarization maps described in the 2015 and 2018 releases (see Planck Collaboration V 2020 for methodological details). We summarize here the main differences between the two analyses:

Temperature map. We use the low-resolution 2015 Commander map, rather than the 2018 Commander map, taking advantadge of the larger sky coverage provided by the former.

Map apodization. We use cosine apodization to produce a low-resolution data set suitable for pixel-space analysis, as did the 2018 release, but the multipole range is different. As discussed in Sect. 4.2, NPIPE uses $\ell_1 = 1$, while the 2018 release uses $\ell_1 = 16$ ($\ell_2 = 48$ in both data sets).

Frequency channels. We perform the analysis on 44, 70, and 100 GHz channels.

Foreground cleaning. We use the 30 and 353 GHz maps as templates for synchrotron and thermal dust, to produce a cleaned frequency map at the target frequency. However, no templates modelling instrumental systematics are included in the fit, as was the case for the LFI analysis in the 2018 release.

Galaxy masks. We adopt the same masks as the LFI 2018 analysis. For an in depth discussion of the derivation and properties of these masks, we refer the reader to Sect. 2.3.1 of Planck Collaboration V (2020); here we just recall that masks are labelled according to the threshold R above which high signal pixels are excluded, so that higher values of R correspond to larger portions of the sky retained in the analysis. For optimum results, a self-consistent sets of masks based on NPIPE maps should be created following the procedure outlined in Planck Collaboration V (2020). However, to facilitate the comparison between NPIPE and PR3 results, and between different frequencies, we adopt the same masks as the LFI 2018 analysis.

Reference foreground scalings. In the LFI 2018 analysis, the foreground scaling coefficients were fixed to those estimated on the R2.2x polarization mask (retaining a sky fraction $f_{\rm sky} = 0.666$), while the cosmological analysis was performed on less aggressive masks. Here instead we use the same mask for both the template cleaning and the cosmological analysis. In particular, this implies that parameters estimated on different masks have been computed on maps corresponding to different template cleaning solutions.

As discussed in Sect. 4.3, the NPIPE calibration scheme using a polarized sky model results in suppression of the lowest multipoles, modelled as a spherically-symmetric transfer function (Figs. 20–22). The small dependence of the transfer function on the sky mask or on the cross-power spectral estimation approach (QML versus pseudo-C_ℓ), suggests that this modelling is an effective description of a more complicated structure. It is therefore important that the simulations used to calculate the transfer function be processed using the same tools and approaches that are applied to data. Pixel-based methods are the equivalent of harmonic auto-spectra methods, and we define the corresponding effective transfer function as

$$C_\ell^{\rm output} = k_\ell^2 C_\ell^{\rm CMB}, \tag{25}$$

where both $C_\ell^{\rm output}$ and $C_\ell^{\rm CMB}$ are QML estimates on $N_{\rm side} = 16$ maps, after applying the R1.8x ($f_{\rm sky} = 0.624$) polarization mask. Figure 78 shows the resulting E-mode transfer functions. Auto-spectra are more impacted by noise bias than cross-spectra, and with the available number of simulations we are not able to reliably measure the transfer function for $\ell > 7$ at 44 and 70 GHz, and $\ell > 11$ for 100 GHz. In addition, at modes $\ell > 4$ for LFI and $\ell > 7$ for 100 GHz, the measured transfer function is compatible with unity within 2 standard deviation (with the exception of $\ell = 7$ at 70 GHz). Those multipoles also roughly mark the scales above which the E-mode S/N falls below unity for values of τ compatible with current estimates, and have little impact on determination of the optical depth. Therefore, we conservatively enforce the transfer function to be unity at $\ell \geq 4$ for 44

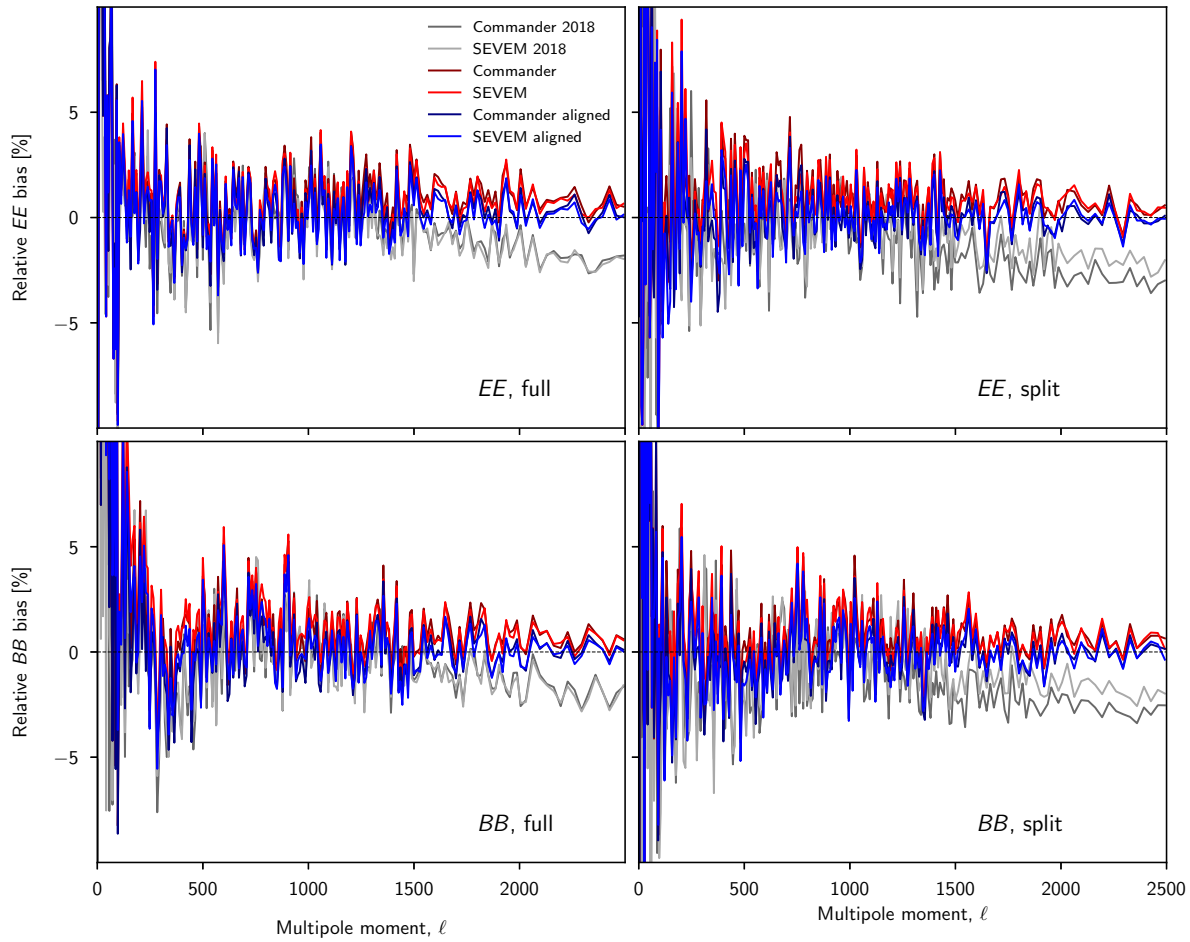


Fig. 73. Power spectrum consistency between the foreground-cleaned `Commander` (dark curves) and `SEVEM` (light curves) CMB polarization map and corresponding end-to-end-simulations. Each panel shows the fractional difference between the angular power spectrum computed from the observed data and the mean of the simulations. Blue and red curves show results derived for `NPIPE` data using simulations *with* and *without* noise alignment, respectively, while grey curves show similar results derived from *Planck* 2018 data using simulations *with* noise alignment. Rows show results for *EE* (*top*) and *BB* (*bottom*) spectra, while columns show results for full-mission (*left*) and split (*right*) data. In the latter case, A-B split results are shown for `NPIPE`, while half-mission splits are shown for *Planck* 2018.

and 70 GHz, and at $\ell \geq 7$ for 100 GHz. We nonetheless checked that using the full measured transfer function has only a minimal impact on the estimated values of τ (assuming a sharp transition to reionization). On the other hand, constraining models with extended reionization history also leverages features in the multipole range $\ell \simeq 10$–20, and any such study would need to assess whether pixel-based methods are suitable, given the above limitations on the transfer function. In the following, we do not account for the uncertainty in transfer-function estimates.

Figure 79 shows the scaling coefficients for synchrotron, α, and the thermal dust, β, and the excess χ^2 (defined as $\Delta\chi^2 = (\chi^2 - N_{\rm dof})/\sqrt{2N_{\rm dof}}$, where $N_{\rm dof}$ is the number of unmasked Q, U pixels as a function of the Galactic mask). In PR3, the analysis was restricted to masks for which $|\Delta\chi^2| < 3$, and the R1.8x mask was selected for the final likelihood. Both 44 and 70 GHz `NPIPE` maps meet the $|\Delta\chi^2| < 3$ criterion over the range of masks considered, while at 100 GHz this criterion is met (marginally) only for the R0.9 ($f_{\rm sky} = 0.379$) or less aggressive masks. At 70 GHz the level of stability over different masks is in line with Planck Collaboration V (2020) results, with an overall lower χ^2 excess. However, `NPIPE`maps seem to prefer higher scaling coefficients than what shown there. For the reference R1.8x mask, we find $\alpha = 0.062 \pm 0.004$ and $\beta = 0.0095 \pm 0.0003$, are about

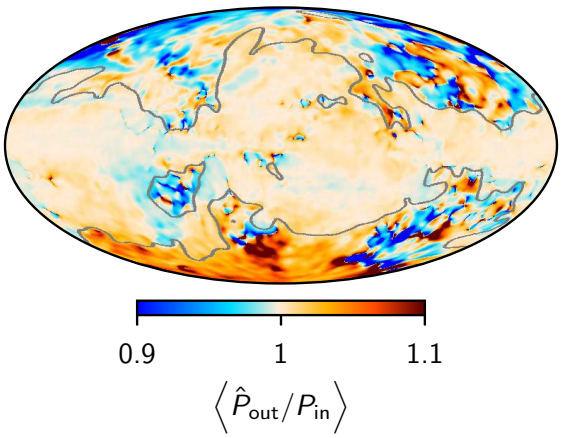


Fig. 74. Ratio between simulation output and input thermal dust polarization amplitude at 353 GHz, smoothed to an angular resolution of 3° FWHM. Gray lines indicate where the polarization amplitude is $3\,\mu{\rm K_{RJ}}$, corresponding roughly to a signal-to-noise ratio of unity.

$1\,\sigma$ higher than Planck Collaboration V (2020) values. On the same R2.2x mask as Planck Collaboration V (2020), we observe

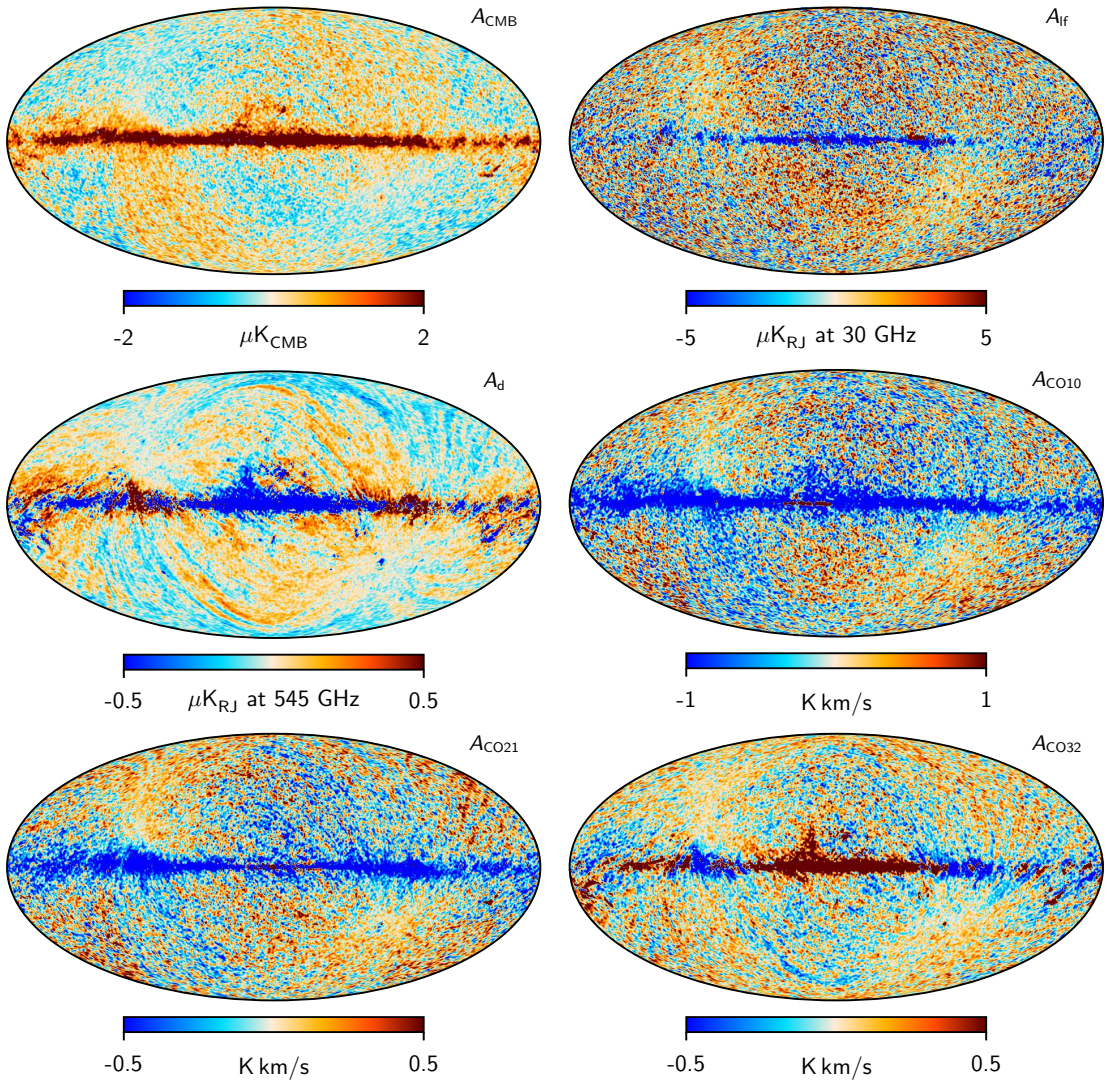


Fig. 75. Half-difference plots from the Commander high-resolution NPIPE splits. All maps are smoothed to a common resolution of 60′ FWHM beam.

Table 11. Comparison of Solar dipole measurements from COBE, WMAP, and *Planck*.

Experiment	Amplitude [μK_{CMB}]	Galactic coordinates l [deg]	Galactic coordinates b [deg]	Reference
COBE–DMR [a,b]	3358 ± 23	264.31 ± 0.16	48.05 ± 0.09	Lineweaver et al. (1996)
COBE–FIRAS	3343 ± 16	265.7 ± 0.5	48.3 ± 0.5	Fixsen et al. (1994)
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	48.265 ± 0.015	Planck Collaboration II (2019)
HFI 2018 [d]	3362.08 ± 0.99	264.021 ± 0.011	48.253 ± 0.005	Planck Collaboration III (2020)
NPIPE [a,c]	**3366.6 ± 2.7**	**263.986 ± 0.035**	**48.247 ± 0.023**	Section 8

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 CMB fluctuations are accounted for by subtracting a dipole-adjusted CMB map from frequency maps prior to dipole estimation.

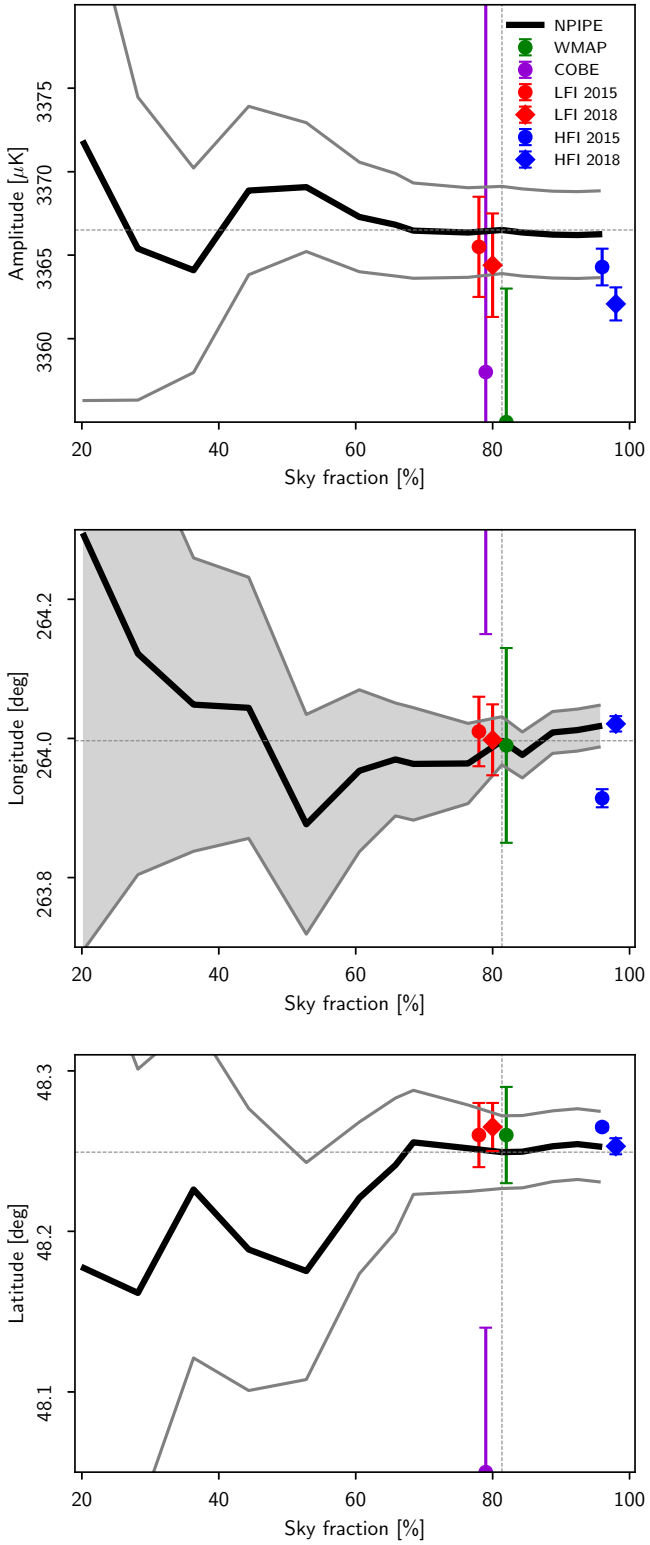


Fig. 76. CMB Solar dipole parameters as a function of sky fraction, as estimated from NPIPE data. The solid black lines show the posterior mean derived with a Wiener-filter estimator, and the grey bands show corresponding $\pm 1\,\sigma$ confidence regions including both statistical and systematic uncertainties. *From top to bottom*, the three panels show amplitude, longitude, and latitude parameters. For comparison, estimates from COBE, WMAP, and *Planck* LFI and HFI are shown as individual coloured data points. The dotted lines represent the NPIPE values that are adopted as final optimal estimates, and summarized in Table 11, defined with a sky fraction of $f_{\rm sky} = 0.81$.

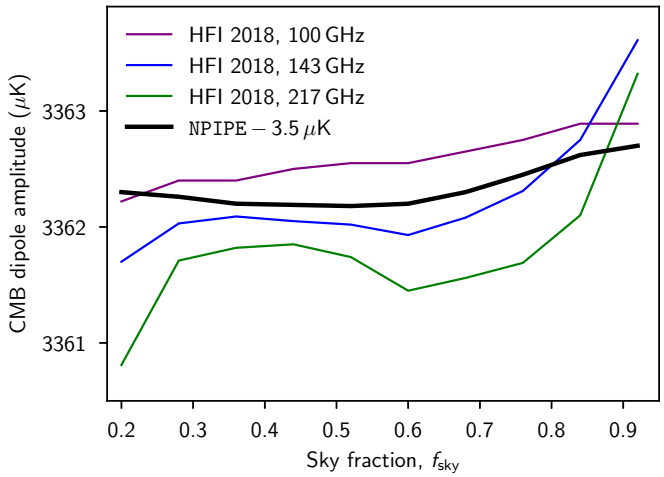


Fig. 77. NPIPE CMB dipole amplitude as a function of sky fraction when estimated using a similar methodology as reported in Planck Collaboration III (2020), i.e., by subtracting a foreground-cleaned *Planck* CMB temperature map prior to estimating the dipole parameters for each value of $f_{\rm sky}$. The coloured curves are direct reproductions from Fig. 23 in Planck Collaboration III (2020). Note that the NPIPE results have been offset by $3.5\,\mu$K for comparison purposes.

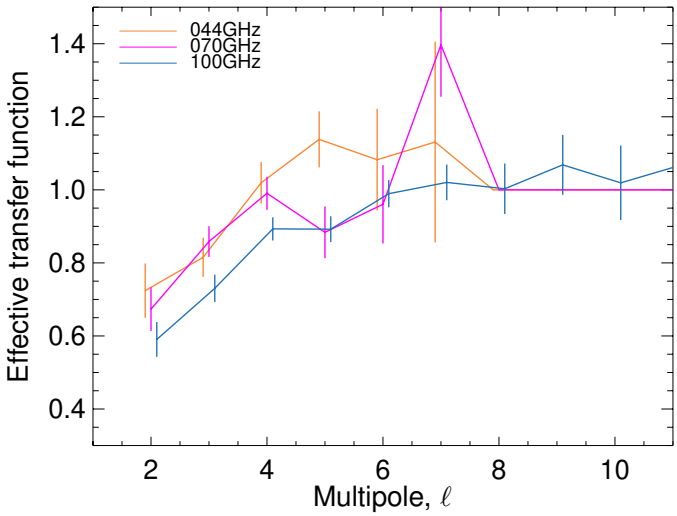


Fig. 78. Low-multipole effective transfer functions for the pixel-based analysis. Due to the higher impact of noise on auto-QML compared to cross-spectra, we are not able to measure the effective transfer function at $\ell > 7$ for 44 and 70 GHz, and at $\ell > 11$ for 100 GHz. In the analysis the transfer function is conservatively enforced to be unity at $\ell \geq 4$ for 44 and 70 GHz, and at $\ell \geq 7$ for 100 GHz. Note that the transfer functions here are, as expected, different from those shown in Figs. 20 and 21. As emphasized in the text, the simulations used to calculate the transfer function must be processed using the same tools and approaches that are applied to the data. The pixel-based analysis here and the spectrum-based analysis represented in Figs. 20 and 21 are different.

only minimal shifts in the estimates ($\alpha = 0.063 \pm 0.004$ and $\beta = 0.0094 \pm 0.0003$). The reason of this discrepancy is currently unclear, but we verified that adopting Planck Collaboration V (2020) scalings shifts τ estimates (discussed later in this section) by about $0.1\,\sigma$. Modelling the foreground spectral energy distributions as in Planck Collaboration IV (2020) (see also Sect. 7.1), we can convert the measured scalings into estimates of the polarized synchrotron spectral index, $\beta_{\rm s}$, and dust emissivity index $\beta_{\rm d}$. Fixing the thermal dust temperature $T_{\rm d} = 19.5$ K, we find

$\beta_s = 3.17 \pm 0.08$, $\beta_d = 1.61 \pm 0.02$, in good agreement with Planck Collaboration IV (2020) results. Results at the other frequencies considered are consistent with 70 GHz estimates, with $(\beta_s, \beta_d) = (3.23 \pm 0.05, 1.56 \pm 0.05)$ and $(\beta_s, \beta_d) = (3.11 \pm 0.08, 1.62 \pm 0.01)$ respectively at 44 and 100 GHz.

Figure 80 shows the power spectra of the corresponding cleaned maps. Some excess power is visible at 44 and 100 GHz, most noticeably for BB at $\ell = 2$ and 3 and for the EE at $\ell = 2$. A proper estimation of the BB transfer function would require a much larger number of simulations than are presently available, due to the much lower signal-to-noise ratio of B modes compared to E modes. Similar issues impact the determination of TB and EB transfer function. It is then not clear whether excesses are driven by residual contamination in the maps, a mismodelling of noise bias, an improper determination of the transfer functions, or a combination of all these effects.

Finally, Fig. 81 shows the τ likelihood distributions for the three NPIPE channels considered here, compared to the results for the PR3 70 GHz maps. The corresponding expectation values are shown in Table 12. Since the focus of this section is mainly on data consistency rather than full exploration of the parameter space, we fix $A_s e^{-2\tau}$ to 1.884, and all other parameters to their ΛCDM best-fit value from Planck Collaboration VI (2020). Because of the weak dependence of low-ℓ polarization on the other cosmological parameters, this assumption has only a minor impact on the recovered value and uncertainty of τ. However, in order to make a clearer comparison between different data sets, the value of τ quoted here for PR3 maps has been re-evaluated using the same setup as the NPIPE analysis.

Estimates from differing data sets are consistent overall, with 44 GHz preferring a roughly $1\,\sigma$ lower value than the other channels. In addition, while the pixel-based 100 GHz value is consistent with those from the other channels, it is about $2\,\sigma$ higher than the estimate based on the 100×143 GHz cross-spectrum (as discussed in the next section). Compared to cross-spectrum methods, a pixel-based approach is more sensitive to inaccuracies in estimation of the effective noise bias. It is possible that the excess χ^2 and the features seen in the 100 GHz power spectra discussed above are related to a failure of the adopted noise model at that frequency, which would also impact τ estimates. A comparison of the noise bias predicted by the 100 GHz NCVM with the power spectra of the corresponding half-ring half-difference (HRHD) map supports this idea, even though it is not clear whether this mismatch can completely account for the excess power at BB $\ell = 2$ and 3 (given the uncertainty on the B-mode transfer function discussed above).

In order to assess the impact of the low-ℓ features discussed above on 100 GHz τ estimates, we repeated the parameter estimation by projecting out EE $\ell = 2$, finding $\tau = 0.0598 \pm 0.0064$, decreasing to $\tau = 0.0564 \pm 0.0063$ if we further project out BB $\ell = 2$ and 3. Only a marginal shift to $\tau = 0.0560 \pm 0.0064$ is observed if we also exclude EE $\ell = 3$. This suggests that the current pixel-based analysis is not yet able to properly model all the sources of uncertainty affecting the largest scales at 100 GHz. Even after projecting out the low-ℓ modes, there remains a roughly $1\,\sigma$ discrepancy between the 100 GHz pixel-based results and the 100×143 cross-spectra results. Figure 80 shows a significant excess also for 100 GHz EE $\ell = 9$, and we performed a similar test to assess the impact of such a feature on our estimates. Projecting out this multipole leaves τ estimates virtually unchanged ($\tau = 0.0633 \pm 0.0065$), compared to the baseline estimate, suggesting that feature cannot account for the residual difference between 100 GHz pixel-based and 100×143 GHz results. On the other hand, pixel-based estimates

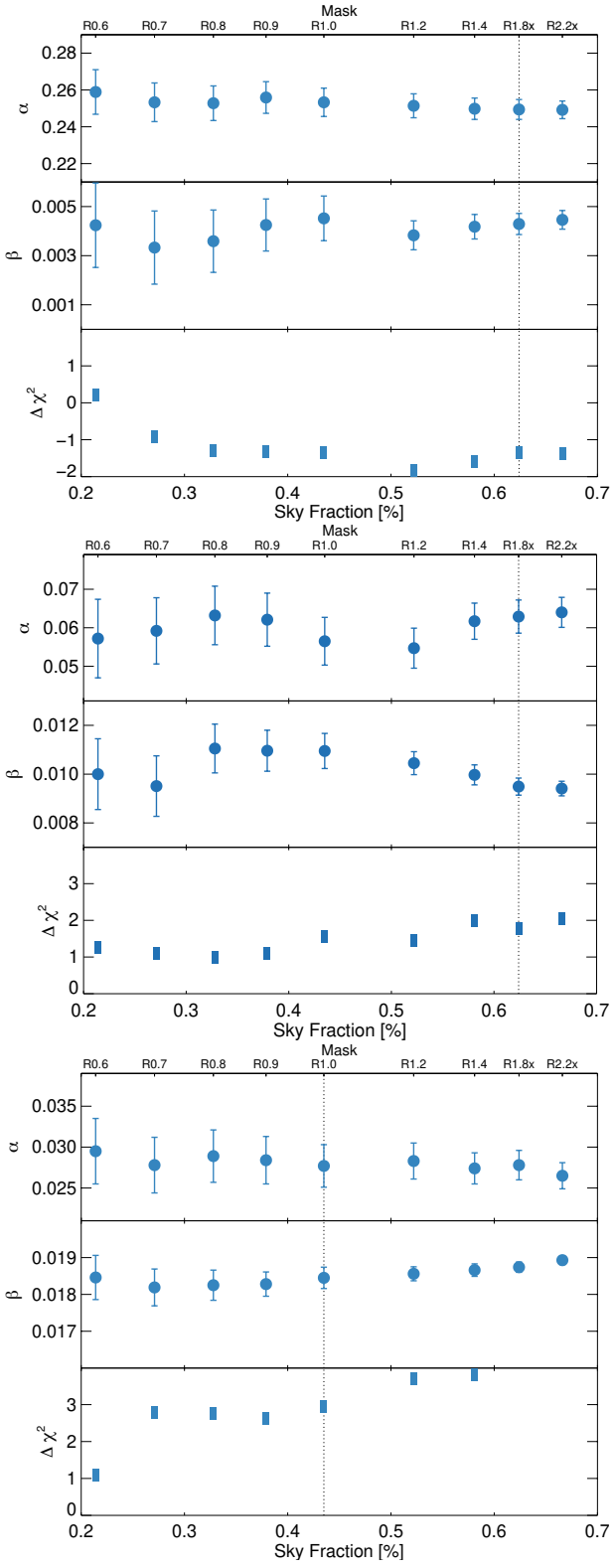


Fig. 79. Scaling coefficients for synchrotron spectral index α and thermal dust emissivity β, and excess $\Delta\chi^2 = (\chi^2 - N_{\rm dof})/\sqrt{2N_{\rm dof}}$, for 9 different masks. The panels *from top to bottom* show results at 44, 70, and 100 GHz, respectively. For 44 and 70 GHz data, the dotted vertical line shows the mask used for parameter estimation with the low-ℓ 2018 LFI likelihood, while for 100 GHz data it shows the largest mask for which $\Delta\chi^2 \leq 3$.

are based on (I, Q, U) maps, and therefore include all four power spectra (TT, EE, TE, and BB), while the cross-spectrum results

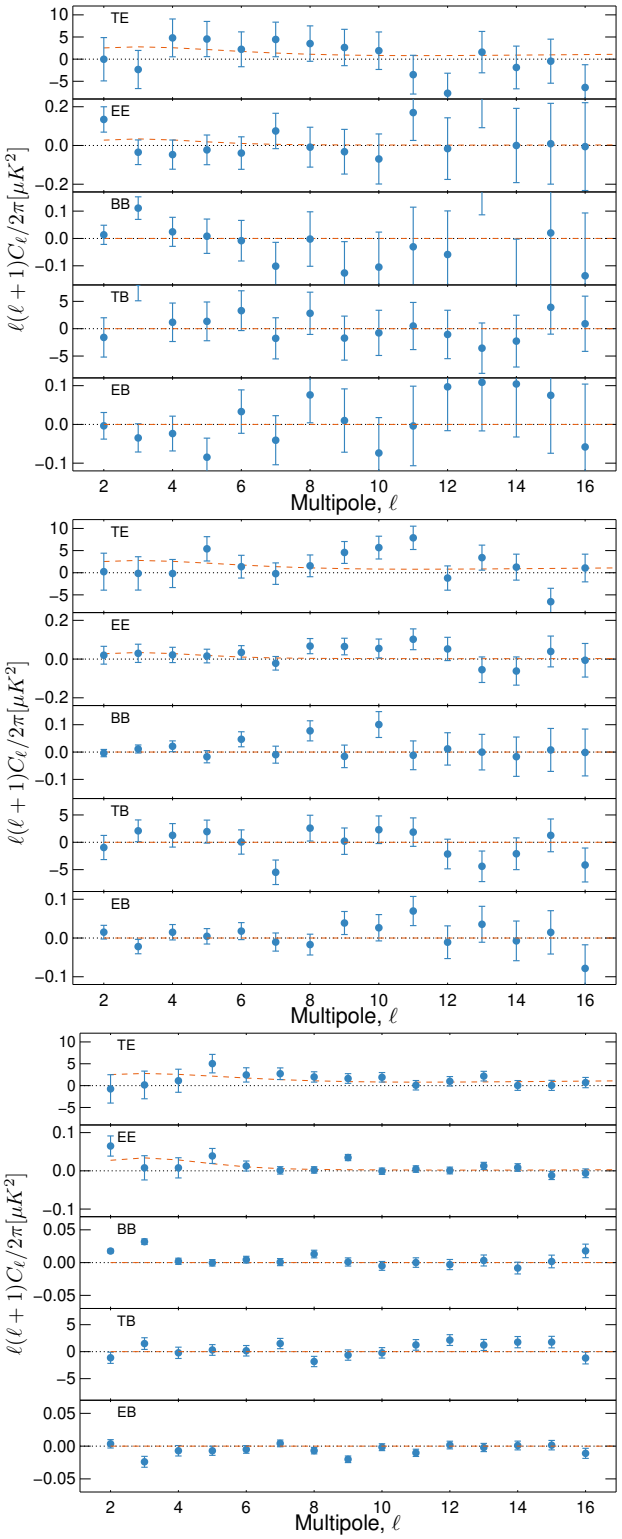


Fig. 80. Power spectra of template-cleaned, low-resolution maps. From top to bottom, the panels show results for 44, 70, and 100 GHz, respectively. The dashed red lines show a model with $\tau = 0.06$ (and not a fit to the data). Spectra for TT are not shown, since in all cases the `Commander` 2015 map is used for temperature. The polarization mask was R1.8x for 44 and 70 GHz, and R0.9 for 100 GHz. In all cases, we used the `Commander` 2015 mask for temperature.

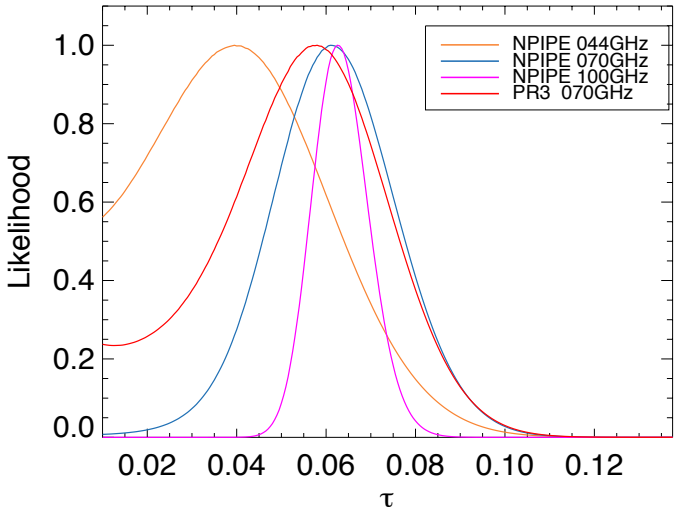


Fig. 81. Reionization optical depth τ likelihood for the `NPIPE` 44-, 70-, and 100 GHz maps, compared to that for the PR3 70 GHz map. Results were computed retaining a polarized sky fraction $f_{\rm sky} = 0.624$ (R1.8x mask) at 44 and 70 GHz, and $f_{\rm sky} = 0.379$ at 100 GHz (R0.9 mask). We fixed $A_{\rm s}e^{-2\tau}$ to 1.884, and the remaining parameters to their ΛCDM best-fit values.

Table 12. Reionization optical depth τ estimates from pixel-based analysis.

ν [GHz]	τ
44	0.0418 ± 0.0182
70	0.0617 ± 0.0142
100	0.0634 ± 0.0063
70 (2018)	0.0537 ± 0.0187

are based on EE spectra only. A direct comparison of the two sets of results is therefore not straightforward.

9.2. Cross-spectrum analysis

We use `lollipop`, the same low-ℓ, EE likelihood as was used to constrain the reionization history in Planck Collaboration Int. XLVII (2016). `Lollipop` is a spectrum-based likelihood following the approach proposed by Hamimeche & Lewis (2008) and extended to cross-spectra in Mangilli et al. (2015). Uncertainties are propagated using spectral covariance matrices estimated from the `NPIPE` end-to-end simulations, which include different realizations for the CMB signal, noise, and systematics (including first-order ADCNL).

We use the cross-correlations of the *Planck* polarized frequency maps from 70 to 217 GHz. Polarized foregrounds at those frequencies are dominated by Galactic dust and synchrotron emission. We use the 353 and 30 GHz *Planck* maps as templates to subtract dust and synchrotron emission, respectively, using a single coefficient for each component. The fit is performed over 52% of the sky, avoiding the inner Galactic plane, as well as the Galactic poles, where the S/N is too low (Fig. 82). Scaling coefficients (given in Table 13) are estimated using cross-correlation between detector-set maps in order to avoid the bias due to the noise in the template maps during the regression process. Once the coefficients are estimated, we clean each frequency map using detector-set maps at 353 and 30 GHz to avoid any noise bias when computing cross-frequency spectra.

Despite the use of these templates, foreground residuals in the cleaned maps are still dominant over the CMB polarized signal near the Galactic plane. For the power-spectrum estimation, we therefore apply a Galactic mask based on the amplitude of

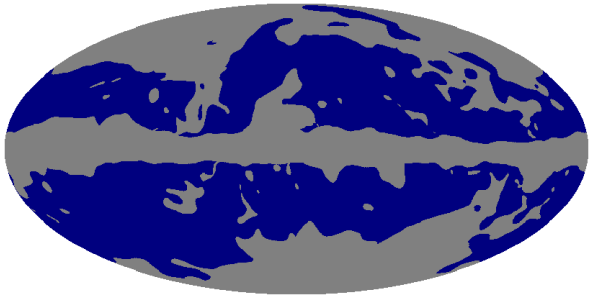


Fig. 82. Sky region ($f_{\rm sky}$ = 0.52) used for the regression of foreground templates on NPIPE frequency maps (blue pixels are kept for the fit).

Table 13. Template coefficients measured on data.

Channel [GHz]	c_{30}	c_{353}
70	0.0671 ± 0.0040	0.0085 ± 0.0005
100	0.0252 ± 0.0023	0.0189 ± 0.0003
143	0.0120 ± 0.0024	0.0399 ± 0.0003
217	0.0101 ± 0.0030	0.1292 ± 0.0004

Notes. Error bars including statistical noise and systematic uncertainties are estimated from NPIPE simulations. Note that the variation of the coefficient depending on the sky region used for the regression is larger than those error bars.

the polarized dust emission retaining either 41%, 52%, 63%, or 75% of the sky (similarly as in Planck Collaboration V 2020).

Cross-spectra are estimated using both a pseudo-C_ℓ estimator `Xpol` (a generalization to polarization of the algorithm presented in Tristram et al. 2005) and a so-called cross-QML estimator, which is adapted from `QML` (Tegmark & de Oliveira-Costa 2001) for cross-spectra (Vanneste et al. 2018), as already used in PR3 low-ℓ results. Both estimators lead to compatible results for computing E-mode power spectra on $N_{\rm side} = 16$ maps with slightly less variance for the latter. The final spectra are computed as the average of the cross-QML correlations between the frequency maps cleaned using detector-set A (or B) and the one cleaned using detector-set B (or A), thus avoiding any noise bias from the templates. Figure 83 shows the cross-frequency power spectra corrected for the `NPIPE` transfer function (principally affecting multipoles $\ell = 2$ and 3, as described in Sect. 4.3).

We use end-to-end simulations to propagate the uncertainties to the cross-power spectra, and all the way to the estimation of the reionization optical depth τ. The C_ℓ covariance matrix used for the likelihood is directly estimated from these Monte Carlos. Figure 83 shows the C_ℓ variance of the Monte Carlos including noise, systematics, and uncertainties related to foreground cleaning compared to cosmic variance.

We construct the likelihood based on the 100 × 143 EE cross-spectrum, and derive the posterior of the reionization optical depth τ in the ΛCDM model, fixing $A_{\rm s}e^{-2\tau}$ as well as other parameters to the *Planck* 2018 best-fit model (Planck Collaboration VI 2020). By default, the multipole range used is $\ell = 2$–20 containing all the statistical power of the reionization bump in EE. We show the results of the constraints coming from the 100 × 143 cross-spectrum, where we first vary the multipole range used in the likelihood (Fig. 84) and the sky fraction used for the computation of the power spectra (Fig. 85). We also compute the likelihood using the cross-spectrum from `Commander` detector-set maps (using the Monte Carlo simulations accordingly) and compare to the cross-frequency spectra in Fig. 86.

We observe consistent results for all frequencies from 70 to 217 GHz, over the range of multipoles between $\ell = 2$ and $\ell = 20$, as well as for sky fractions below 75%. Note that, unlike the results in Planck Collaboration V (2020), there is no Monte Carlo correction at the level of the likelihood to take into account residual biases, including foreground residuals depending on sky cuts.

In particular, the value for the reionization optical depth τ obtained for the same data set used in *Planck* 2018 (i.e., the 100× 143 EE cross-spectrum), using 63% of the sky, is

$$\tau = 0.051 \pm 0.006, \tag{26}$$

which is fully compatible with the value $\tau = 0.0506 \pm 0.0086$ (lowE) derived in Planck Collaboration VI (2020), but with a lower uncertainty due to the improvements described in this paper.

10. Conclusions

We have presented `NPIPE`, a data-processing pipeline for processing raw *Planck* timestream data into calibrated frequency maps. `NPIPE` runs outside the *Planck*-DPC architecture and can be deployed at almost any supercomputing centre. The software, input data, and configuration files are released (see Appendix A) to allow a motivated reader with sufficient computing resources to repeat our analysis and improve upon it.

Compared to PR3 results, we have demonstrated significant reduction in the overall noise and systematics across essentially all angular scales, most notably reducing HFI EE and BB noise and systematics variance at $\ell < 10$ by 50–90% and reducing degree-scale statistical noise variance by 10–30% across both instruments. The improvement in the large-scale polarization uncertainty has been shown to come from the application of a polarized sky model during calibration, and is associated with a measurable and correctable suppression of CMB E-mode power at large angular scales. We have also shown substantial improvements in the internal and external consistency of the frequency maps, providing for a demonstrably-better model of the microwave sky.

`NPIPE` processing modules differ, in some cases significantly, from the ones developed, tested, and applied in *Planck* processing over two decades. We have tested our results extensively against PR2 and PR3, and have documented the differences throughout this paper. Most of these differences are well-understood; they support the notion that `NPIPE` maps have lower noise and systematics.

We summarize here what we consider the benefits of the `NPIPE` processing and products, and also give several cautionary comments. Advantages of the `NPIPE` release:

- reduced levels of noise and systematics at all angular scales;
- improved consistency across frequencies, particularly in polarization;
- more Monte Carlo realizations of simulated data, and better agreement between the simulated maps and flight data;
- availability of single-detector temperature maps from 100 to 857 GHz, with resolution of $N_{\rm side} = 4096$ at 217 GHz and above;
- absence of certain HFI analysis artefacts, in particular "zebra" stripes and CO-template pixel boundaries;
- one publicly-released pipeline to process LFI and HFI data, accompanied with the release of the raw timestreams for future analysis and improvement.

The following are some cautionary comments about the `NPIPE` release.

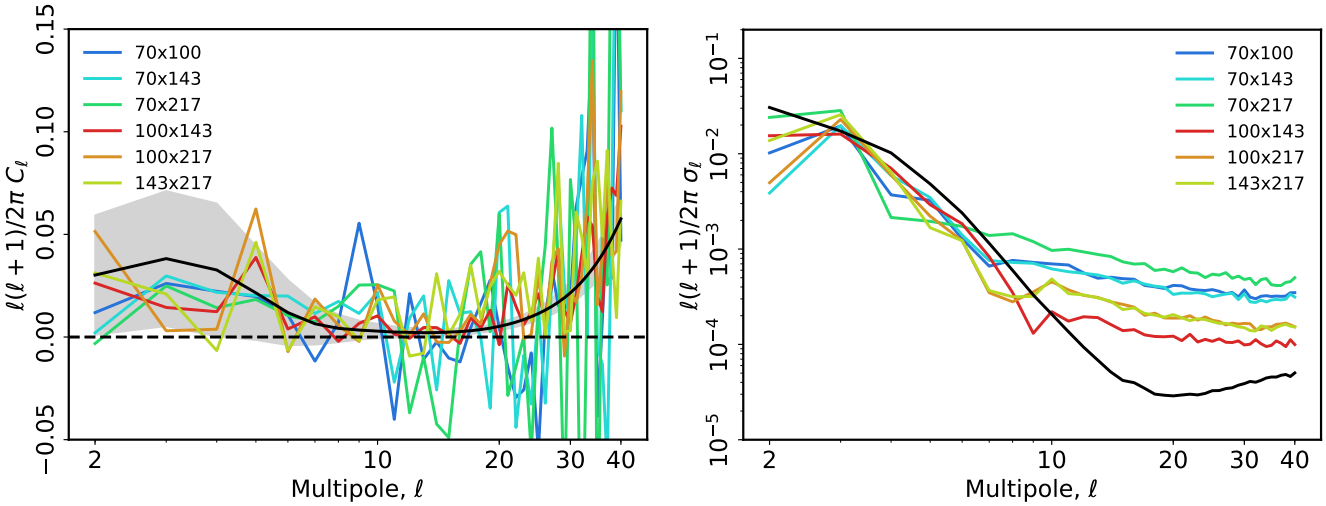


Fig. 83. *Left*: cross-power spectra for the four *Planck* frequency maps at 70, 100, 143, and 217 GHz compared to the *Planck*2018 best-fit model for *EE* (solid black line). *Right*: standard deviation of the sum of noise, systematics, and uncertainties related to foreground cleaning for the cross-frequency spectra computing using Monte Carlos as compared to the cosmic variance (solid black line).

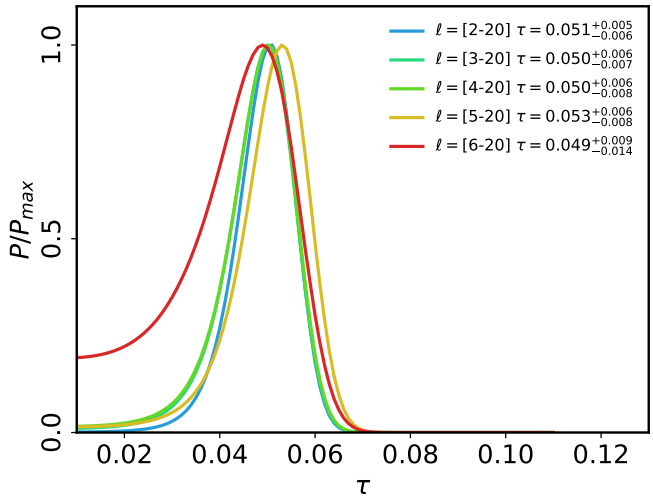


Fig. 84. Reionization optical depth τ posterior for the 100×143 *EE* cross-spectrum, varying the lowest multipole used in the likelihood.

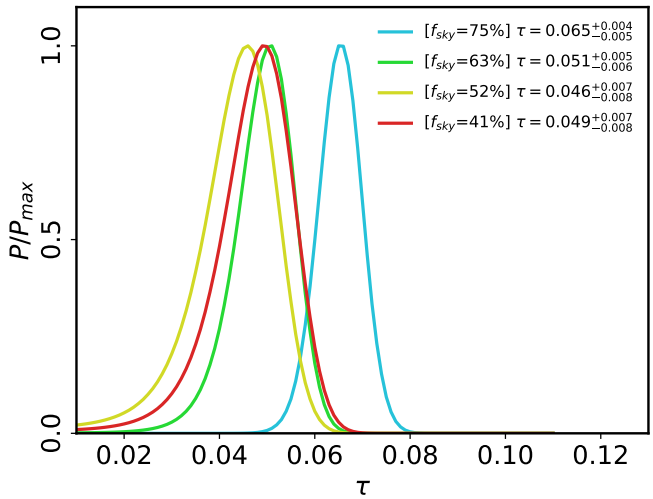


Fig. 85. Reionization optical depth τ posterior for the 100×143 *EE* cross-spectrum, varying the sky fraction used.

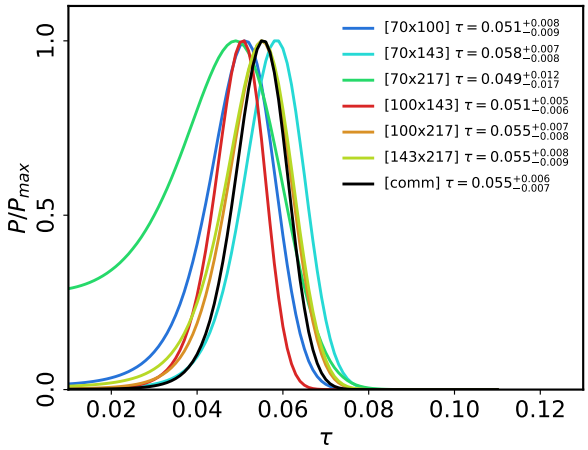


Fig. 86. Optical depth τ posterior depending on the cross-spectra used in the likelihood. We use frequency-cleaned maps from 70 to 217 GHz. We also show the posterior for the cross-power spectrum estimated using COMMANDER detector-set maps.

– Quantitative use of the NPIPE CMB polarization data on large angular scales ($\ell < 20$) requires accounting for the non-negligible transfer function. This will often necessitate processing the large (36 TB) body of NPIPE simulations to determine the degree of signal suppression in the cosmological observables of concern, at considerable computational expense.
– At the time of writing, the NPIPE data products have been extensively tested and validated through to the map level. With the exception of the demonstration cases of foreground separation and the determination of τ given in this paper, extraction of the full range of science results represented in the PR2 and PR3 releases remains for the future. While ongoing tests of cosmological parameter solutions indicate no disparity with the *Planck* 2018 results, presentation of work on this subject is deferred to a future publication. Although unlikely, it is possible that future analysis beyond what is presented here will uncover unidentified issues that impact the estimation of cosmological observables from NPIPE data.
– The component separation and sky model described in Sect. 7 use only *Planck* data, and do not include external data at

lower frequencies that help break degeneracies between synchrotron, free-free, and AME, as was done in PR2. We leave this for a future study. For now, the limitations of the low-frequency foreground model should be recognized.

NPIPE represents the first comprehensive effort to process all nine *Planck* frequencies using the same pipeline modules, and to leverage the wide spectral response of the two instruments to derive a consistent, multi-frequency data set. Our use of the 30, 217, and 353 GHz polarization maps as priors in calibrating the CMB frequencies can be considered a first attempt to incorporate component separation *into* the mapmaking pipeline. On-going and future efforts to expand the treatment into a full-fledged component-separation treatment may yet find significant gains over what is presented here.

Acknowledgements. The Planck Collaboration acknowledges the support of: ESA; CNES, and CNRS/INSU-IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MINECO, JA, and RES (Spain); Tekes, AoF, and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); ERC and PRACE (EU). A description of the Planck Collaboration and a list of its members, indicating which technical or scientific activities they have been involved in, can be found at http://www.cosmos.esa.int/web/planck/planck-collaboration. This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant numbers 776282, 772253 and 819478. This research would not have been possible without the resources of the National Energy Research Scientific Computing Center (NERSC), a US Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231.

References

Argüeso, F., Sanz, J. L., Herranz, D., López-Caniego, M., & González-Nuevo, J. 2009, MNRAS, 395, 649
Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJS, 208, 20
Chon, G., Challinor, A., Prunet, S., Hivon, E., & Szapudi, I. 2004, MNRAS, 350, 914
Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792
Delouis, J. M., Pagano, L., Mottet, S., Puget, J. L., & Vibert, L. 2019, A&A, 629, A38
Eriksen, H. K., O'Dwyer, I. J., Jewell, J. B., et al. 2004, ApJS, 155, 227
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
Fixsen, D., Cheng, E. S., Cottingham, D. A., et al. 1994, ApJ, 420, 445
Godard, B., Croon, M., Budnik, F., & Morley, T. 2009, Proceedings of the 21st International Symposium on Space Flight Dynamics (ISSFD) (Toulouse)
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Hamimeche, S., & Lewis, A. 2008, Phys. Rev. D, 77, 103013
Hinshaw, G., Weiland, J. L., Hill, R. S., et al. 2009, ApJS, 180, 225
Hivon, E., Mottet, S., & Ponthieu, N. 2017, A&A, 598, A25
Jewell, J., Levin, S., & Anderson, C. H. 2004, ApJ, 609, 1
Keihänen, E., Kurki-Suonio, H., Poutanen, T., Maino, D., & Burigana, C. 2004, A&A, 428, 287
Keihänen, E., Kurki-Suonio, H., & Poutanen, T. 2005, MNRAS, 360, 390
Keihänen, E., Keskitalo, R., Kurki-Suonio, H., Poutanen, T., & Sirviö, A. 2010, A&A, 510, A57
Kelsall, T., Weiland, J. L., Franz, B. A., et al. 1998, ApJ, 508, 44
Kurki-Suonio, H., Keihänen, E., Keskitalo, R., et al. 2009, A&A, 506, 1511
Leach, S. M., Cardoso, J., Baccigalupi, C., et al. 2008, A&A, 491, 597
Lenz, D., Hensley, B. S., & Doré, O. 2017, ApJ, 846, 38
Lineweaver, C. H., Tenorio, L., Smoot, G. F., et al. 1996, ApJ, 470, 38
López-Caniego, M., Herranz, D., González-Nuevo, J., et al. 2006, MNRAS, 370, 2047
Mangilli, A., Plaszczynski, S., & Tristram, M. 2015, MNRAS, 453, 3174
Mangilli, A., Aumont, J., Bernard, J. P., et al. 2019, A&A, 630, A74
Notari, A., & Quartin, M. 2015, JCAP, 1506, 047
Odegard, N., Weiland, J. L., Fixsen, D. J., et al. 2019, ApJ, 877, 40
Planck Collaboration I. 2014, A&A, 571, A1
Planck Collaboration II. 2014, A&A, 571, A2
Planck Collaboration V. 2014, A&A, 571, A5
Planck Collaboration VI. 2014, A&A, 571, A6
Planck Collaboration IX. 2014, A&A, 571, A9
Planck Collaboration X. 2014, A&A, 571, A10
Planck Collaboration XII. 2014, A&A, 571, A12
Planck Collaboration XIV. 2014, A&A, 571, A14
Planck Collaboration XXVII. 2014, A&A, 571, A27
Planck Collaboration I. 2016, A&A, 594, A1
Planck Collaboration II. 2016, A&A, 594, A2
Planck Collaboration V. 2016, A&A, 594, A5
Planck Collaboration VII. 2016, A&A, 594, A7
Planck Collaboration VIII. 2016, A&A, 594, A8
Planck Collaboration IX. 2016, A&A, 594, A9
Planck Collaboration X. 2016, A&A, 594, A10
Planck Collaboration XII. 2016, A&A, 594, A12
Planck Collaboration XXI. 2016, A&A, 594, A21
Planck Collaboration XXV. 2016, A&A, 594, A25
Planck Collaboration XXVI. 2016, A&A, 594, A26
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 VII. 2020, A&A, 641, A7
Planck Collaboration XI. 2020, A&A, 641, A11
Planck Collaboration Int. XLVI. 2016, A&A, 596, A107
Planck Collaboration Int. XLVII. 2016, A&A, 596, A108
Planck HFI Core Team 2011, A&A, 536, A6
Prezeau, G., & Reinecke, M. 2010, ApJS, 190, 267
Reinecke, M., & Seljebotn, D. S. 2013, A&A, 554, A112
Remazeilles, M., Delabrouille, J., & Cardoso, J.-F. 2011, MNRAS, 418, 467
Rosset, C., Tristram, M., Ponthieu, N., et al. 2010, A&A, 520, A13
Seljebotn, D. S., Bærland, T., Eriksen, H. K., Mardal, K. A., & Wehus, I. K. 2019, A&A, 627, A98
Szapudi, I., Prunet, S., & Colombi, S. 2001, ApJ, 561, L11
Tegmark, M., & de Oliveira-Costa, A. 2001, Phys. Rev. D, 64, 063001
Tristram, M., Macías-Pérez, J. F., Renault, C., & Santos, D. 2005, MNRAS, 358, 833
Tuttlebee, M. 2013, Herschel/Planck Star Tracker Performance Assessment and Calibration, Tech. Rep. PT-CMOC-OPS-RP-6435-HSO-GF, Herschel/Planck Flight Dynamics; SCISYS GmbH
Vanneste, S., Henrot-Versillé, S., Louis, T., & Tristram, M. 2018, Phys. Rev. D, 98, 103526
Villa, F., Terenzi, L., Sandri, M., et al. 2010, A&A, 520, A6
Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A. 2004, Phys. Rev. D, 70, 083511
Wehus, I. K., Fuskeland, U., Eriksen, H. K., et al. 2014, ArXiv e-prints [arXiv:1411.7616]
Zonca, A., Singer, L., Lenz, D., et al. 2019, J. Open Source Softw., 4, 1298

[1] APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/lrfu, Observatoire de Paris, Sorbonne Paris Cité, 10 Rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France
[2] African Institute for Mathematical Sciences, 6-8 Melrose Road, Muizenberg, Cape Town, South Africa
[3] Astrophysics Group, Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, UK
[4] Astrophysics & Cosmology Research Unit, School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Private Bag X54001, Durban 4000, South Africa
[5] CITA, University of Toronto, 60 St. George St., Toronto, ON M5S 3H8, Canada
[6] CNRS, IRAP, 9 Av. Colonel Roche, BP 44346, 31028 Toulouse Cedex 4, France
[7] Cahill Center for Astronomy and Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA
[8] California Institute of Technology, Pasadena, CA, USA
[9] Computational Cosmology Center, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[10] Département de Physique Théorique, Université de Genève, 24 Quai E. Ansermet, 1211 Genève 4, Switzerland
[11] Département de Physique, École normale supérieure, PSL Research University, CNRS, 24 Rue Lhomond, 75005 Paris, France

[12] Departamento de Astrofísica, Universidad de La Laguna (ULL), 38206 La Laguna, Tenerife, Spain
[13] Departamento de Física, Universidad de Oviedo, C/ Federico García Lorca 18, Oviedo, Spain
[14] Department of Astrophysics/IMAPP, Radboud University, PO Box 9010, 6500 GL Nijmegen, The Netherlands
[15] Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC, Canada
[16] Department of Physics & Astronomy, University of the Western Cape, Cape Town 7535, South Africa
[17] Department of Physics, Gustaf Hällströmin katu 2a, University of Helsinki, Helsinki, Finland
[18] Department of Physics, Princeton University, Princeton, NJ, USA
[19] Department of Physics, University of California, Santa Barbara, CA, USA
[20] Dipartimento di Fisica e Astronomia G. Galilei, Università degli Studi di Padova, Via Marzolo 8, 35131 Padova, Italy
[21] Dipartimento di Fisica e Scienze della Terra, Università di Ferrara, Via Saragat 1, 44122 Ferrara, Italy
[22] Dipartimento di Fisica, Università La Sapienza, P. le A. Moro 2, Roma, Italy
[23] Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria, 16, Milano, Italy
[24] Dipartimento di Fisica, Università degli Studi di Trieste, Via A. Valerio 2, Trieste, Italy
[25] Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy
[26] European Space Agency, ESAC, Planck Science Office, Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, Madrid, Spain
[27] European Space Agency, ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
[28] Gran Sasso Science Institute, INFN, Viale F. Crispi 7, 67100 L'Aquila, Italy
[29] HEP Division, Argonne National Laboratory, Lemont, IL 60439, USA
[30] Haverford College Astronomy Department, 370 Lancaster Avenue, Haverford, PA, USA
[31] Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland
[32] IFPU – Institute for Fundamental Physics of the Universe, Via Beirut 2, 34014 Trieste, Italy
[33] INAF – OAS Bologna, Istituto Nazionale di Astrofisica – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Area della Ricerca del CNR, Via Gobetti 101, 40129 Bologna, Italy
[34] INAF – Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, Padova, Italy
[35] INAF – Osservatorio Astronomico di Trieste, Via G.B. Tiepolo 11, Trieste, Italy
[36] INAF – Osservatorio Astronomico di Trieste, Via G.B. Tiepolo 11, 34143 Trieste, Italy
[37] INAF, Istituto di Radioastronomia, Via Piero Gobetti 101, 40129 Bologna, Italy
[38] INAF/IASF Milano, Via E. Bassini 15, Milano, Italy
[39] INFN – CNAF, viale Berti Pichat 6/2, 40127 Bologna, Italy
[40] INFN, Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
[41] INFN, Sezione di Ferrara, Via Saragat 1, 44122 Ferrara, Italy
[42] INFN, Sezione di Milano, Via Celoria 16, Milano, Italy
[43] INFN, Sezione di Roma 2, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy
[44] IUCAA, Post Bag 4, Ganeshkhind, Pune University Campus, Pune 411 007, India
[45] Imperial College London, Astrophysics group, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK
[46] Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA
[47] Institut d'Astrophysique Spatiale, CNRS, Univ. Paris-Sud, Université Paris-Saclay, Bât. 121, 91405 Orsay Cedex, France
[48] Institut d'Astrophysique de Paris, CNRS (UMR7095), 98 bis Boulevard Arago, 75014 Paris, France
[49] Institute Lorentz, Leiden University, PO Box 9506, Leiden 2300 RA, The Netherlands
[50] Institute of Theoretical Astrophysics, University of Oslo, Blindern Oslo, Norway
[51] Instituto de Astrofísica de Canarias, C/Vía Láctea s/n, La Laguna, Tenerife, Spain
[52] Instituto de Física de Cantabria (CSIC-Universidad de Cantabria), Avda. de los Castros s/n, Santander, Spain
[53] Istituto Nazionale di Fisica Nucleare, Sezione di Padova, Via Marzolo 8, 35131 Padova, Italy
[54] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, USA
[55] Jodrell Bank Centre for Astrophysics, Alan Turing Building, School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester M13 9PL, UK
[56] Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[57] LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France
[58] Laboratoire de Physique Subatomique et Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, 53 Rue des Martyrs, 38026 Grenoble Cedex, France
[59] Laboratoire de Physique Théorique, Université Paris-Sud 11 & CNRS, Bâtiment 210, 91405 Orsay, France
[60] Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[61] Low Temperature Laboratory, Department of Applied Physics, Aalto University, Espoo 00076, AALTO, Finland
[62] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
[63] Mullard Space Science Laboratory, University College London, Surrey RH5 6NT, UK
[64] National Centre for Nuclear Research, ul. L. Pasteura 7, 02-093 Warsaw, Poland
[65] SISSA, Astrophysics Sector, Via Bonomea 265, 34136 Trieste, Italy
[66] San Diego Supercomputer Center, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
[67] School of Physical Sciences, National Institute of Science Education and Research, HBNI, Jatni 752050, Odissa, India
[68] School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff CF24 3AA, UK
[69] School of Physics and Astronomy, Sun Yat-sen University, 2 Daxue Rd, Tangjia, Zhuhai, PR China
[70] School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK
[71] School of Physics, Indian Institute of Science Education and Research Thiruvananthapuram, Maruthamala PO, Vithura, Thiruvananthapuram 695551, Kerala, India
[72] Simon Fraser University, Department of Physics, 8888 University Drive, Burnaby, BC, Canada
[73] Sorbonne Université, CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France
[74] Sorbonne Université, Observatoire de Paris, Université PSL, École normale supérieure, CNRS, LERMA, 75005 Paris, France
[75] Space Science Data Center – Agenzia Spaziale Italiana, Via del Politecnico snc, 00133 Roma, Italy
[76] Space Sciences Laboratory, University of California, Berkeley, CA, USA
[77] The Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, AlbaNova 106 91, Stockholm, Sweden
[78] Université de Toulouse, UPS-OMP, IRAP, 31028 Toulouse Cedex 4, France
[79] Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland

Appendix A: NPIPE software and data release

The NPIPE software is publicly accessible on Github[18] at https://github.com/hpc4cmb/toast-npipe. In addition to some general dependencies, NPIPE is built upon the TOAST[19] framework for time-ordered data processing. The final destriping is performed using libMadam[20] (Keihänen et al. 2005, 2010). NPIPE is a Python3 code with Cython and C extensions. libMadam is a Fortran 2003 code.

The pipeline is parallelized with mpi4py[21]. The beam-convolved timestreams are produced with libconviqt[22] (Prezeau & Reinecke 2010). Distributed spherical harmonic operations are performed using libsharp[23] (Reinecke & Seljebotn 2013). HEALPix map operations are from healpy[24] (Zonca et al. 2019). Other FITS files are manipulated using PyFITS[25] from Astropy[26]. The beam window functions are evaluated using the QuickPol (Hivon et al. 2017) code adapted to NPIPE files and included in the NPIPE repository. Power-spectrum estimation with mode decoupling is done using PolSpice[27] (Szapudi et al. 2001; Chon et al. 2004).

The frequency maps, sky models, simulations, beam window functions, time streams, and auxiliary files are all available at NERSC under /global/cfs/cdirs/cmb/data/planck2020. Interested parties are invited to apply for an account following the instructions at crd.lbl.gov/departments/computational-science/c3/c3-research/cosmic-microwave-background/cmb-data-at-nersc. Flight data products and limited release of the simulations will be made available on the *Planck* Legacy Archive (PLA) at https://www.cosmos.esa.int/web/planck/pla

Appendix B: Calibration

We consider two alternative approaches to correcting gain fluctuations in the detector data. The first is conceptually simpler and can be described as "total power calibration". The second approach is based on the destriping principle (Keihänen et al. 2004, 2005, 2010; Kurki-Suonio et al. 2009) and targets the gain fluctuations instead of the total signal strength.

B.1. Total-power calibration

A direct approach to calibration is to build a signal model, $\boldsymbol{s}$, and regress it against the detector data, $\boldsymbol{d}$, at some chosen gain steps. The fitting coefficients, g, are directly the detector gains during each gain step:

$$\boldsymbol{d} = g \cdot \boldsymbol{s} + b \cdot \boldsymbol{o} + \boldsymbol{n}, \tag{B.1}$$

where b is the noise offset, $\boldsymbol{o}$ is a constant offset template, and $\boldsymbol{n}$ is the instrumental noise without an offset. Maximum likelihood estimates of the template coefficients g and b follow from the familiar linear regression. First we collect the templates into columns of a template matrix,

$$\mathsf{F} = [\boldsymbol{s}, \boldsymbol{o}], \tag{B.2}$$

and then solve the template coefficients, $\boldsymbol{a} = [g, b]^\mathsf{T}$, from

$$\boldsymbol{d} = \mathsf{F}\,\boldsymbol{a} + \boldsymbol{n}. \tag{B.3}$$

Assuming a noise covariance matrix $\mathsf{N} = \langle \boldsymbol{n}\boldsymbol{n}^\mathsf{T} \rangle$ the maximum likelihood solution is

$$\boldsymbol{a} = \left(\mathsf{F}^\mathsf{T}\mathsf{N}^{-1}\mathsf{F}\right)^{-1} \mathsf{F}^\mathsf{T}\mathsf{N}^{-1}\boldsymbol{d}. \tag{B.4}$$

Aside from simplicity, the total power calibration has the benefit that it naturally produces gains that both trace gain fluctuations and calibrates the signal against known astrophysical sources such as the Solar dipole.

Multiple gain steps are easily accommodated by adding columns to the template matrix F. Each gain template is zeroed outside the intended gain step or one may even blend the templates to enforce continuity.

Total-power calibration suffers from incomplete knowledge of the sky signal. The sky estimate, $\boldsymbol{s}$, *has* to be incomplete. If we knew exactly what the sky was, the measurement would have been unnecessary in the first place. Any errors (noise and otherwise) in our sky estimate cause a bias towards zero due to a linear regression phenomenon know as "errors in variables". Ignoring the noise offset for a while, it can be shown that the gain is systematically underestimated:

$$\hat{g} = \frac{g}{1 + \sigma^2/\sigma_\mathrm{g}^2}, \tag{B.5}$$

where σ_g^2 is the variance of the "true" sky signal and σ^2 is the variance of errors in our estimate. The error is directly proportional to the overall gain:

$$\frac{g}{\hat{g}} = 1 + \sigma^2/\sigma_\mathrm{g}^2. \tag{B.6}$$

This form of the bias $g/\hat{g}$ is only accurate in the presence of a single template and becomes more complicated when several gain steps or additional templates are considered.

Lack of astrophysical gradients in the signal (e.g., scanning along the dipole equator) makes even this simple system degenerate and causes unchecked transfer of power between the gain and the noise offset.

B.2. Calibration via destriping

In this section we describe the calibration approach adopted in Planck Collaboration Int. XLVI (2016) for HFI analysis, and is a multi-detector polarized extension of the DaCapo algorithm used for LFI calibration (Planck Collaboration II 2016, 2019) with some critical differences.

Destriping assumes that the detector data are a combination of a sky-synchronous signal, a linear combination of noise templates (typically disjoint noise offsets), and uncorrelated white noise:

$$\boldsymbol{d} = \mathsf{P}\boldsymbol{m} + \mathsf{F}\,\boldsymbol{a} + \boldsymbol{n}. \tag{B.7}$$

This algorithm works to infer a maximum likelihood estimate of the sky, $\hat{\boldsymbol{m}}$, by first determining the noise template amplitudes in $\boldsymbol{a}$.

[18] github.com
[19] TOAST is the time-ordered astrophysics scalable toolset. It is publicly available at github.com/hpc4cmb/toast
[20] libMadam is a library version of the Madam generalized destriper. It can be downloaded from github.com/hpc4cmb/libmadam
[21] bitbucket.org/mpi4py/mpi4py
[22] github.com/hpc4cmb/libconviqt
[23] github.com/Libsharp/libsharp
[24] HealPy is a Python front-end (github.com/healpy/healpy) to the HEALPix library: healpix.sourceforge.io.
[25] https://pypi.org/project/pyfits/3.3/
[26] www.astropy.org
[27] www2.iap.fr/users/hivon/software/PolSpice/

It is possible to repurpose the destriping approach to fit for gain *fluctuations* rather than the total gain. We use the destriping projection operator:

$$\mathsf{Z} = \mathsf{I} - \mathsf{P}\left(\mathsf{P}^{\mathsf{T}}\mathsf{N}^{-1}\mathsf{P}\right)^{-1}\mathsf{P}^{\mathsf{T}}\mathsf{N}^{-1}, \tag{B.8}$$

where I is the identity matrix and P is the pointing matrix. The operator Z acts on time-domain objects such as $\boldsymbol{d}$ and $g \cdot \boldsymbol{s}$ by:

1. binning a map of the signal (object);
2. resampling a timeline from the map; and
3. subtracting the resampled signal from the original signal.

The result is a time-domain object with the sky-synchronous part removed. Such a signal will be void of all stationary sky signal, but the projection will also affect the non-stationary parts, such as noise, orbital dipole, and gain fluctuations.

If we consistently apply this projection operator to every time-domain object in the linear regression equation, Eq. (B.4), we find

$$\boldsymbol{a} = \left(\mathsf{F}^{\mathsf{T}}\mathsf{N}^{-1}\mathsf{Z}\mathsf{F}\right)^{-1}\mathsf{F}^{\mathsf{T}}\mathsf{N}^{-1}\mathsf{Z}\,\boldsymbol{d}. \tag{B.9}$$

We have simplified Eq. (B.9) using the projection matrix properties of Z to write $\mathsf{Z}^{\mathsf{T}}\mathsf{N}^{-1}\mathsf{Z} = \mathsf{N}^{-1}\mathsf{Z}$. We note that Eq. (B.9) is also the maximum likelihood solution of Eq. (B.7) in the absence of prior knowledge about the template amplitude covariance:

$$\mathsf{C}_{\mathrm{a}} = \langle \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}} \rangle \quad \Rightarrow \quad \mathsf{C}_{\mathrm{a}}^{-1} = 0 \tag{B.10}$$

In Eq. (B.9) we have translated the total power calibration (Eq. (B.4) into a subspace that lacks sky-synchronous degrees of freedom. Recovered gain-template amplitudes no longer reflect the overall gain, but rather deviations about a sky-synchronous average. The errors-in-variables bias now becomes

$$\widehat{\delta g} = \frac{\delta g}{1 + \sigma^2/\sigma_g^2}, \tag{B.11}$$

making the error proportional to the gain fluctuation, rather than the overall gain. The bias is now attenuating the magnitude of the fluctuation instead of systematically pulling down the overall gain. If we iterate over the calibration, the gains rapidly converge to a self-consistent pair of sky and gain estimates $(\hat{m}, \hat{g})$, even with errors in the gain template.

Appendix C: Far-sidelobe corrections to the dipole

`NPIPE` uses the same method as the 2018 LFI DPC processing to convolve the ideal dipole model with the `GRASP` estimates of the *Planck* sidelobes. The corrections are applied to the Solar dipole and associated quadrupole.

We begin by noting that the Doppler effect up to the quadrupole term can be written (omitting a monopole term) as (e.g., Notari & Quartin 2015):

$$D(\hat{\boldsymbol{n}}) = T_0\left[\boldsymbol{\beta}\cdot\hat{\boldsymbol{n}}\,(1 + q\,\boldsymbol{\beta}\cdot\hat{\boldsymbol{n}})\right], \tag{C.1}$$

where T_0 is the CMB monopole temperature, $\hat{\boldsymbol{n}}$ is the observing direction, $\boldsymbol{\beta} = \boldsymbol{v}/c$ is the total velocity divided by the speed of light and q is the frequency-dependent quadrupole factor:

Table C.1. Frequency-dependent, second-order quadrupole factors.

ν [GHz]	q
0	1.00
30	1.02
44	1.05
70	1.12
100	1.25
143	1.48
217	2.00
353	3.12
545	4.80
857	7.55

$$q = \frac{x}{2}\cdot\frac{e^x + 1}{e^x - 1}, \quad \text{with} \quad x = \frac{h\nu}{k_{\mathrm{B}}T_0}, \tag{C.2}$$

where ν is the observing frequency (see also Planck Collaboration XXVII 2014). The quadrupole factors for each *Planck* frequency are shown in Table C.1. LFI-DPC processing used unity in place of q.

A detector observing the sky at direction $\hat{\boldsymbol{n}}_0$ will see the dipole and the quadrupole in Eq. (C.1) convolved with the instrumental beam response, $B(\hat{\boldsymbol{n}})$:

$$\widetilde{D}(\hat{\boldsymbol{n}}_0) = \int \mathrm{d}\Omega\, B(\hat{\boldsymbol{n}}, \hat{\boldsymbol{n}}_0) D(\hat{\boldsymbol{n}}). \tag{C.3}$$

If we write $D(\hat{\boldsymbol{n}})$ from Eq. (C.1) in terms of its Cartesian components, $\boldsymbol{\beta} = (\beta_{\mathrm{x}}, \beta_{\mathrm{y}}, \beta_{\mathrm{z}})$ and $\hat{\boldsymbol{n}} = (x, y, z)$ we find

$$\begin{aligned}\widetilde{D}(\hat{\boldsymbol{n}}_0) = T_0 \int \mathrm{d}\Omega\, B(\hat{\boldsymbol{n}}, \hat{\boldsymbol{n}}_0)\Big[&\beta_{\mathrm{x}}x + \beta_{\mathrm{y}}y + \beta_{\mathrm{z}}z \\ &+ q\Big(\beta_{\mathrm{x}}^2x^2 + \beta_{\mathrm{x}}\beta_{\mathrm{y}}xy + \beta_{\mathrm{x}}\beta_{\mathrm{z}}xz \\ &+ \beta_{\mathrm{y}}\beta_{\mathrm{x}}yx + \beta_{\mathrm{y}}^2y^2 + \beta_{\mathrm{y}}\beta_{\mathrm{z}}yz \\ &+ \beta_{\mathrm{z}}\beta_{\mathrm{x}}zx + \beta_{\mathrm{z}}\beta_{\mathrm{y}}zy + \beta_{\mathrm{z}}^2z^2\Big)\Big].\end{aligned} \tag{C.4}$$

If we now assume that we always rotate the velocity, $\boldsymbol{\beta}$, into a constant frame where $\hat{\boldsymbol{n}}_0$ is along the z-axis (or any other frame where the beam, B is constant), then the integrals can be pre-evaluated and we have

$$\begin{aligned}\widetilde{D} = T_0\Big[&S_{\mathrm{x}}\beta_{\mathrm{x}} + S_{\mathrm{y}}\beta_{\mathrm{y}} + S_{\mathrm{z}}\beta_{\mathrm{z}} \\ &+ q\Big(S_{\mathrm{xx}}\beta_{\mathrm{x}}^2 + S_{\mathrm{xy}}\beta_{\mathrm{x}}\beta_{\mathrm{y}} + S_{\mathrm{xz}}\beta_{\mathrm{x}}\beta_{\mathrm{z}} \\ &+ S_{\mathrm{yx}}\beta_{\mathrm{y}}\beta_{\mathrm{x}} + S_{\mathrm{yy}}\beta_{\mathrm{y}}^2 + S_{\mathrm{yz}}\beta_{\mathrm{y}}\beta_{\mathrm{z}} \\ &+ S_{\mathrm{zx}}\beta_{\mathrm{z}}\beta_{\mathrm{x}} + S_{\mathrm{zy}}\beta_{\mathrm{z}}\beta_{\mathrm{y}} + S_{\mathrm{zz}}\beta_{\mathrm{z}}^2\Big)\Big],\end{aligned} \tag{C.5}$$

where (for example), $S_{\mathrm{xy}} \equiv \int \mathrm{d}\Omega\, B(\hat{\boldsymbol{n}})\, x\, y$.

We carry out the integrals over entire 4π beams and store the resulting S-parameters for every *Planck* detector, allowing us to convolve the Doppler field with the full beam response on-the-fly for every detector sample. We then only need to rotate the total velocity, $\boldsymbol{\beta}$, and carry out the sum in Eq. (C.5).

Appendix D: Anomalies around the Galactic centre

NPIPE fixes a known issue in PR3 polarized HFI maps around the Galactic centre (Mangilli et al. 2019, Fig. 9 lower panel). The CO templates used in 2018 bandpass-mismatch correction were downgraded to low resolution ($N_{\rm side} = 128$ or $27\farcm5$) and the outlines of these pixels can be detected around the Galactic centre in the polarization maps. We show images of the Galactic centre in Fig. D.1.

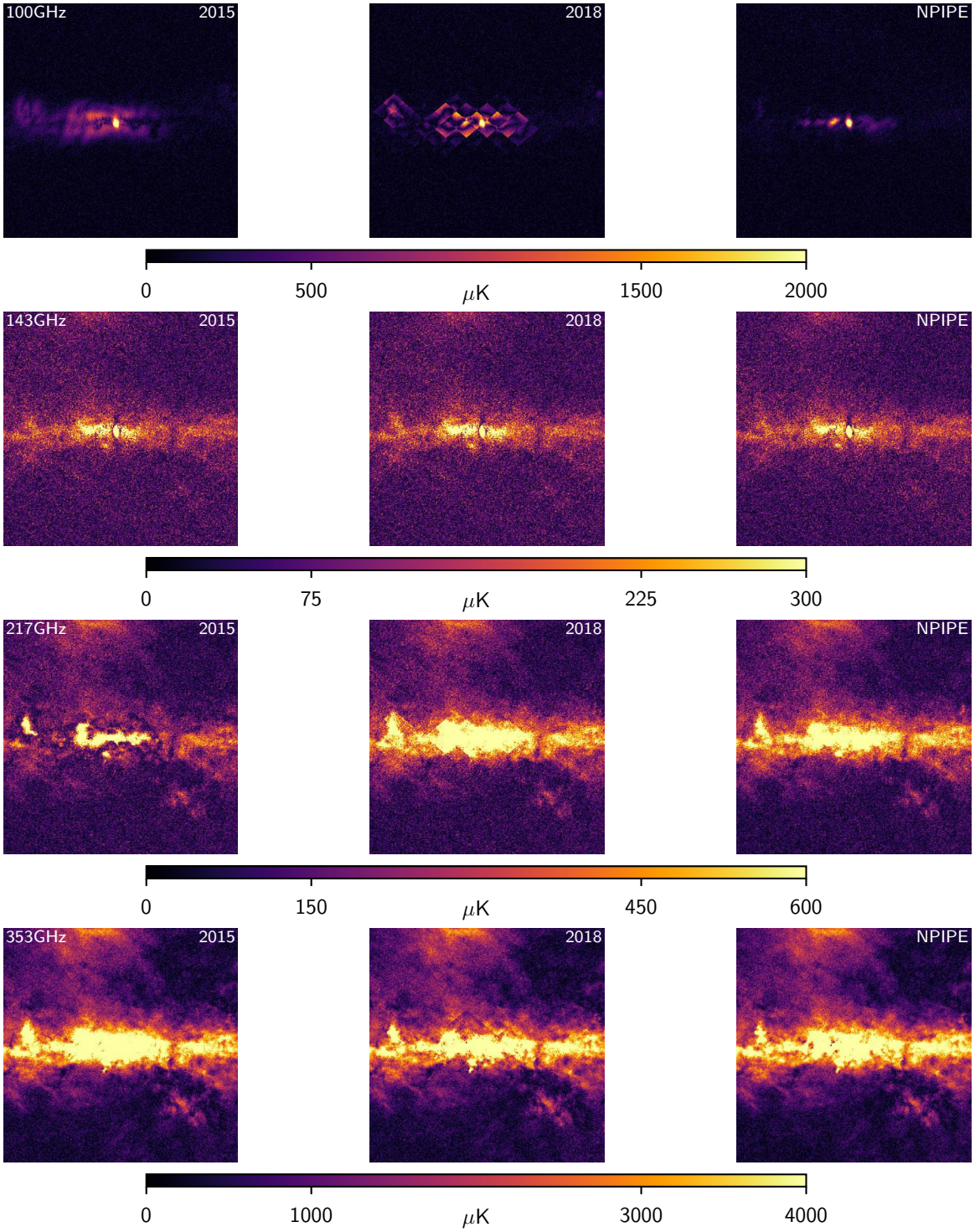


Fig. D.1. Polarization amplitude in an $8^\circ \times 8^\circ$ patch centred around the Galactic centre. The linear colour scale was chosen to demonstrate the low-resolution CO template residuals in the 2018 maps. The residuals are most pronounced at 100 GHz, where the CO corrections are largest, and absent at 143 GHz, where there is no CO correction needed. Since the 2015 maps are not corrected for bandpass mismatch, they do not display the same artefacts..

Appendix E: Visualizations of the destriping templates

NPIPE suppresses systematics by fitting and removing time-domain templates (see Sect. 2.4). The destriping templates are stored as columns of the template matrix, F, in Eq. (10). In this Appendix we visualize the templates, first by binning their full time-domain representations as a function of the spacecraft spin phase and pointing period (ring) index (Fig. E.1). In Fig. E.2 we bin the first survey (six months and approximately 5500 rings) onto more intuitive HEALPix maps. Most of the templates are time-dependent, so the full mission span of the templates cannot be binned into a single map without loosing the time-dependent features.

The gain and distortion templates are represented as single panels but, in reality, are split into a number of disjoint time steps and columns in the template matrix. Each of these steps is fitted as a separate template.

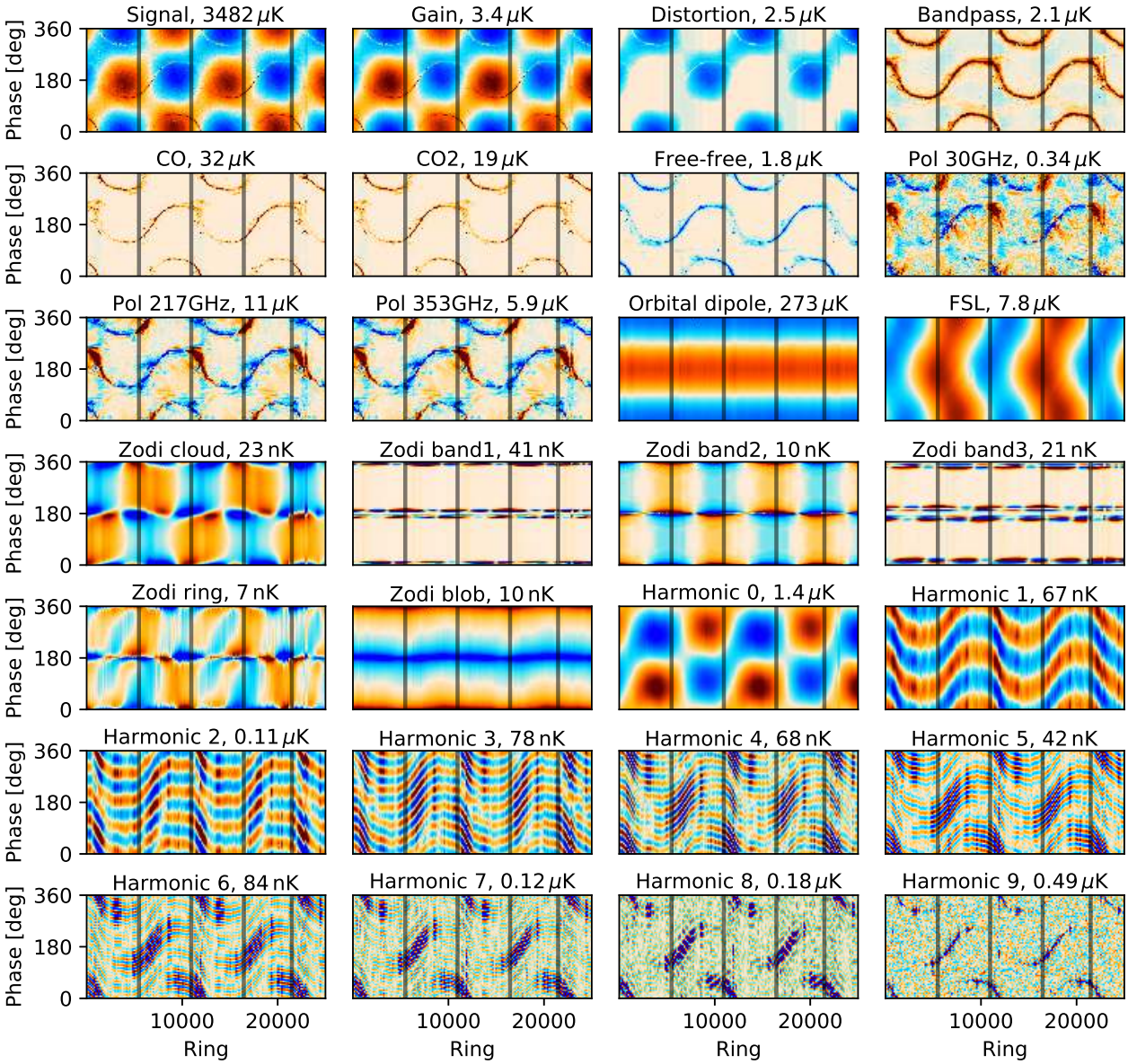


Fig. E.1. Signal and systematics templates for detector 100-1a, plotted as a function of pointing period (ring) and spacecraft spin phase. The gain and signal distortion templates are actually split into several disjoint steps that vary in length depending on the S/N. The templates for 100-1b are otherwise identical, but the 30, 217, and 353 GHz polarization templates are multiplied by −1. The far sidelobe (FSL) template is not fitted because of degeneracies, but it is estimated and subtracted. The polarization templates across all detectors share a single fitting amplitude. The zodiacal emission-template amplitudes are similarly shared. For 353 GHz and above, the harmonic templates are doubled to include frequency-dependent gain. At 100–217 GHz, only relative time-shift between frequency bins is modelled. The last harmonic template includes all frequencies not included in the other harmonic templates. The templates are scaled to match the rms amplitude of each systematic across the 100 GHz detectors, and the plotting ranges are chosen to match the $2\,\sigma$ range of each panel. To save space, the amplitude is reported in the title of each panel rather than as a colour bar. The grey vertical lines indicate the survey boundaries. Figure E.2 shows HEALPix maps of these templates that include only the first survey.

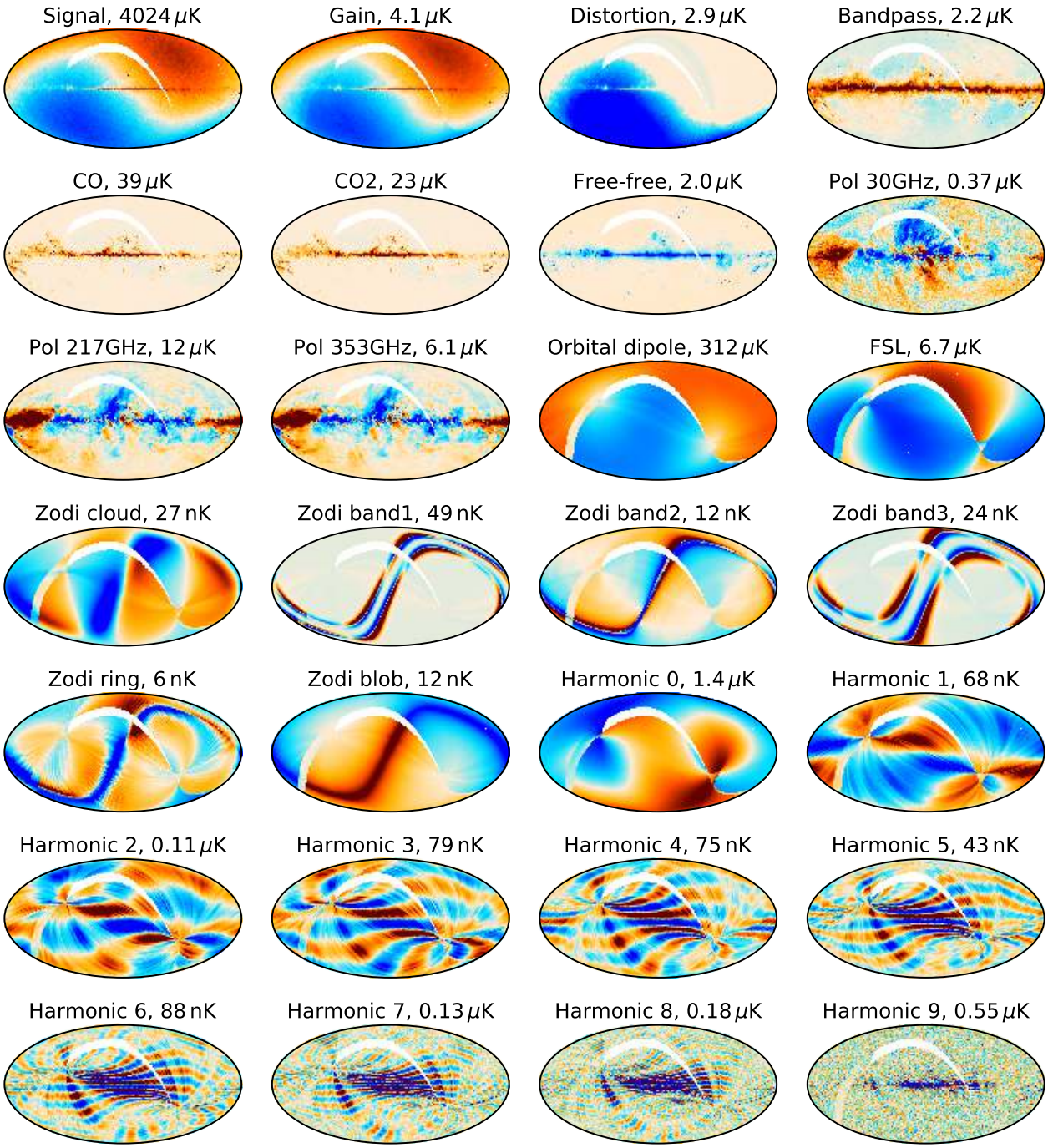


Fig. E.2. Binned maps of detector 100-1a signal and systematics templates for the first survey. For details, see the caption for Fig. E.1.

Appendix F: Validating the NPIPE ADCNL formalism

The NPIPE model for ADC nonlinearity adds another expansion order over the linear gain model used in PR3 (Planck Collaboration III 2020). That puts the complexity of the NPIPE model between the SRoll solution in PR3 and the SRoll2 solution (Delouis et al. 2019). It was discussed in Sect. 5 that the NPIPE simulation set only includes ADCNL compatible with the linear gain model. This is acceptable, if the applied ADCNL correction is powerful enough to remove the original ADCNL and replace it with the statistical and systematic template uncertainties.

To test the efficacy of the NPIPE ADCNL correction, we produced two sets of simulated 143 GHz TOD: one with a full model of ADCNL as was discussed in Planck Collaboration III (2020) and one without ADCNL. Other aspects of the simulated TOD were identical. Running NPIPE on these two simulations we may quantify the amount of ADCNL left in the NPIPE frequency maps by measuring the difference between the full and ADCNL-free maps. We show the ADCNL residual maps in Fig. F.1 and compare them to a noise estimate map derived from the half-ring, half-difference maps. In Fig. F.2 we show the power spectra of the residual maps. These figures demonstrate that the residual ADCNL is expected at a level lower than the pure instrumental noise. They also demonstrate that the linear gain correction alone is not enough to meet the instrumental noise threshold.

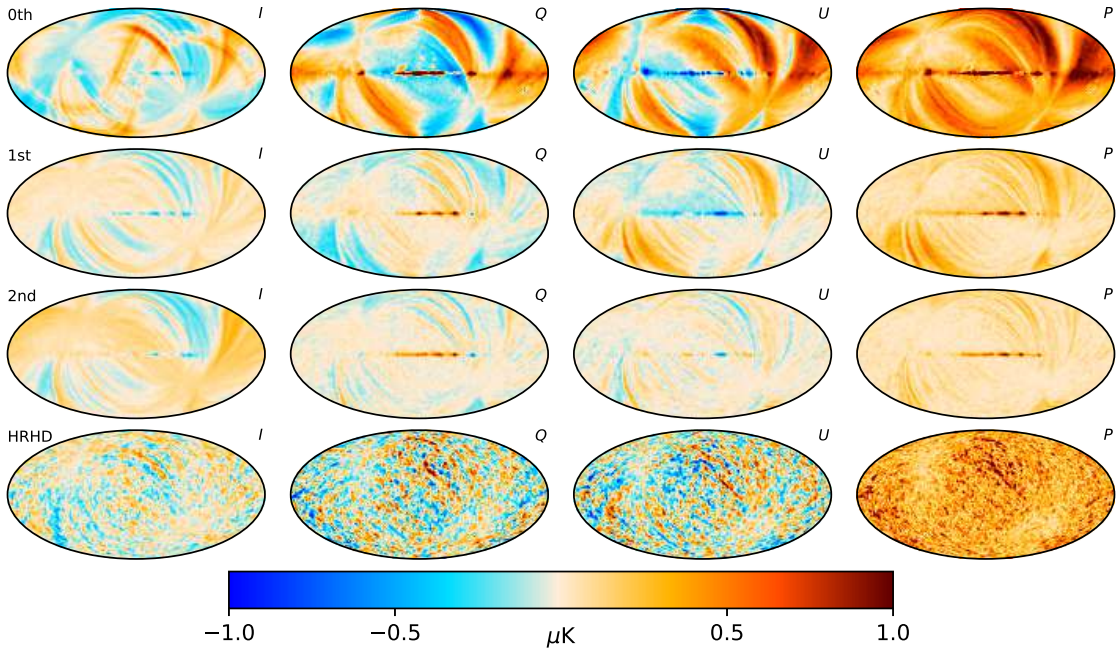


Fig. F.1. Simulated full-mission maps of HFI ADCNL at 143 GHz. *Top row*: full, unmitigated effect. *Second row*: residuals after fitting and correcting ADCNL using the linear gain model (first-order correction), as was done in PR3. *Third row*: residual after fitting for gain and distortion terms, as is done in NPIPE. The NPIPE transfer function (Sect. 4.3) applies to both the full and ADCNL-free simulations. *Fourth* and *last row*: half-ring, half-difference map from the same simulation to compare the magnitude of the effect to instrumental noise. All maps were smoothed with a 3° Gaussian beam. The residual power spectra are shown in Fig. F.2.

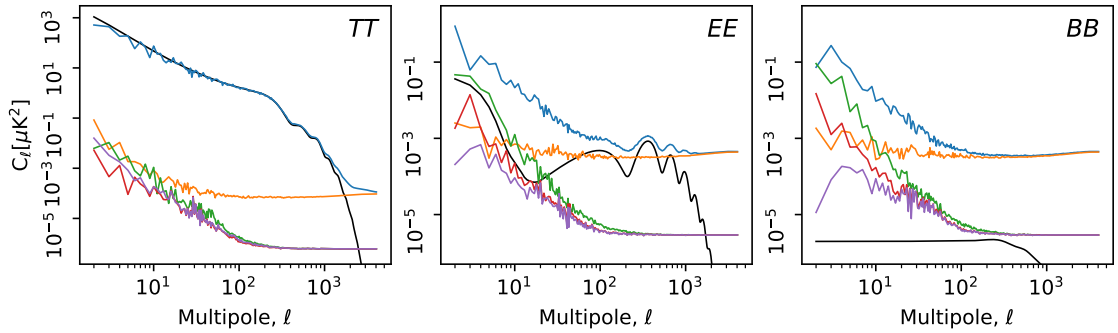


Fig. F.2. Power spectra of the simulated full-mission ADCNL effect over 50% of the sky, corrected for the sky fraction. The green curve shows the full, unmitigated effect. Red shows the residual after the linear (first order) gain correction, and magenta shows the residual after applying the gain and distortion templates (second order), as is done in NPIPE. For scale, the power spectrum of the full 143 GHz frequency map is shown in blue, and an instrumental noise estimate (from the half-ring, half-difference map) is shown in orange. The theoretical input power spectrum ($\tau = 0.06$) used in the simulation is in black. The residual maps are shown in Fig. F.1. The *EE* residuals are deconvolved from the NPIPE transfer function (Sect. 4.3).

Appendix G: Tabulated transfer function

In Table G.1 we show the values of the measured E transfer functions that correspond to Figs. 20 and 21.

Table G.1. NPIPE E transfer functions over 60% of the sky corresponding to Figs. 20 and 21.

Multipole	30 GHz	44 GHz	70 GHz	100 GHz	143 GHz	217 GHz	353 GHz
2	1.020 ± 0.087	0.761 ± 0.031	0.663 ± 0.020	0.584 ± 0.016	0.560 ± 0.015	0.562 ± 0.018	0.936 ± 0.072
3	0.968 ± 0.042	0.817 ± 0.018	0.811 ± 0.012	0.496 ± 0.013	0.506 ± 0.013	0.498 ± 0.014	1.073 ± 0.058
4	1.015 ± 0.035	0.942 ± 0.019	0.924 ± 0.016	0.876 ± 0.008	0.850 ± 0.008	0.858 ± 0.010	1.075 ± 0.048
5	1.034 ± 0.041	0.899 ± 0.025	0.870 ± 0.018	0.812 ± 0.010	0.813 ± 0.010	0.798 ± 0.012	0.980 ± 0.058
6	1.077 ± 0.046	0.860 ± 0.035	0.935 ± 0.030	0.868 ± 0.014	0.864 ± 0.013	0.876 ± 0.016	0.923 ± 0.061
7	1.005 ± 0.056	0.904 ± 0.054	0.918 ± 0.041	0.892 ± 0.018	0.903 ± 0.016	0.923 ± 0.021	0.853 ± 0.079
8	0.913 ± 0.064	1.008 ± 0.071	0.881 ± 0.058	0.890 ± 0.021	0.857 ± 0.020	0.916 ± 0.026	0.886 ± 0.097
9	0.911 ± 0.079	0.974 ± 0.092	0.866 ± 0.068	0.895 ± 0.027	0.941 ± 0.024	0.940 ± 0.030	0.995 ± 0.119
10	1.017 ± 0.088	1.139 ± 0.096	0.938 ± 0.073	0.914 ± 0.028	0.900 ± 0.025	0.846 ± 0.036	1.159 ± 0.133
11	0.831 ± 0.097	0.794 ± 0.117	0.888 ± 0.083	0.856 ± 0.032	0.906 ± 0.030	0.810 ± 0.039	1.274 ± 0.143
12	0.918 ± 0.102	0.798 ± 0.127	0.760 ± 0.093	0.899 ± 0.033	0.863 ± 0.030	0.939 ± 0.042	0.806 ± 0.157
13	1.155 ± 0.115	0.526 ± 0.147	1.186 ± 0.119	0.949 ± 0.040	0.965 ± 0.037	0.942 ± 0.050	0.716 ± 0.165
14	1.163 ± 0.110	0.967 ± 0.150	0.884 ± 0.122	0.917 ± 0.040	0.932 ± 0.035	0.890 ± 0.046	0.908 ± 0.159
15	0.980 ± 0.114	1.047 ± 0.148	0.812 ± 0.119	0.992 ± 0.038	0.903 ± 0.036	0.830 ± 0.051	1.035 ± 0.150
16	1.049 ± 0.098	1.143 ± 0.146	0.922 ± 0.110	0.989 ± 0.037	0.944 ± 0.033	0.962 ± 0.044	0.762 ± 0.157
17	1.012 ± 0.101	0.770 ± 0.140	0.941 ± 0.111	1.046 ± 0.036	0.951 ± 0.033	0.948 ± 0.047	0.919 ± 0.146
18	0.991 ± 0.095	0.762 ± 0.141	0.888 ± 0.105	1.031 ± 0.035	0.944 ± 0.033	0.915 ± 0.043	1.003 ± 0.143
19	0.930 ± 0.104	1.142 ± 0.128	0.828 ± 0.101	0.984 ± 0.034	1.021 ± 0.028	0.878 ± 0.043	0.882 ± 0.143
20	0.950 ± 0.097	1.161 ± 0.116	0.763 ± 0.093	0.907 ± 0.031	0.988 ± 0.027	0.993 ± 0.036	0.603 ± 0.136
21	0.941 ± 0.090	0.995 ± 0.110	1.112 ± 0.087	0.886 ± 0.026	0.932 ± 0.024	0.955 ± 0.035	0.877 ± 0.125
22	1.048 ± 0.084	0.929 ± 0.113	0.966 ± 0.079	1.005 ± 0.025	0.994 ± 0.022	0.955 ± 0.032	0.949 ± 0.118
23	0.958 ± 0.087	0.903 ± 0.097	0.985 ± 0.076	0.974 ± 0.023	0.978 ± 0.020	0.966 ± 0.029	1.142 ± 0.116
24	0.931 ± 0.080	0.967 ± 0.093	1.006 ± 0.066	1.033 ± 0.022	0.980 ± 0.019	0.981 ± 0.027	1.268 ± 0.120
25	1.176 ± 0.081	0.875 ± 0.086	0.959 ± 0.066	1.017 ± 0.021	0.961 ± 0.017	1.010 ± 0.026	1.077 ± 0.115
26	1.043 ± 0.078	1.111 ± 0.084	0.757 ± 0.066	0.994 ± 0.019	1.022 ± 0.018	1.008 ± 0.025	1.068 ± 0.114
27	1.025 ± 0.072	1.028 ± 0.078	0.873 ± 0.061	1.014 ± 0.018	1.000 ± 0.017	0.997 ± 0.022	1.193 ± 0.104
28	0.962 ± 0.078	0.976 ± 0.072	1.005 ± 0.058	0.986 ± 0.018	0.995 ± 0.015	0.994 ± 0.022	1.088 ± 0.108
29	0.999 ± 0.076	0.983 ± 0.073	0.991 ± 0.060	0.994 ± 0.016	1.009 ± 0.015	0.998 ± 0.020	0.855 ± 0.106
30	1.019 ± 0.081	0.995 ± 0.067	0.964 ± 0.053	1.009 ± 0.016	1.011 ± 0.015	0.999 ± 0.020	1.057 ± 0.108
31	0.928 ± 0.078	1.042 ± 0.061	0.925 ± 0.051	0.996 ± 0.015	0.969 ± 0.014	0.994 ± 0.019	1.166 ± 0.106
32	1.012 ± 0.078	0.867 ± 0.062	0.947 ± 0.048	0.989 ± 0.014	0.982 ± 0.013	0.971 ± 0.019	1.150 ± 0.110
33	1.051 ± 0.075	0.988 ± 0.061	0.875 ± 0.049	0.997 ± 0.013	1.000 ± 0.013	1.017 ± 0.017	1.112 ± 0.103
34	0.911 ± 0.073	0.992 ± 0.058	1.034 ± 0.047	0.987 ± 0.013	0.986 ± 0.012	1.002 ± 0.017	1.006 ± 0.103
35	0.934 ± 0.078	0.888 ± 0.054	1.009 ± 0.044	0.975 ± 0.012	0.999 ± 0.012	0.989 ± 0.017	0.968 ± 0.110
36	0.927 ± 0.082	1.014 ± 0.053	0.961 ± 0.044	0.993 ± 0.012	1.017 ± 0.011	0.994 ± 0.016	0.990 ± 0.107
37	1.083 ± 0.080	0.871 ± 0.052	0.907 ± 0.040	1.003 ± 0.012	1.013 ± 0.011	0.977 ± 0.015	0.881 ± 0.106
38	0.932 ± 0.083	0.974 ± 0.048	0.971 ± 0.039	1.000 ± 0.012	1.012 ± 0.010	0.997 ± 0.015	0.960 ± 0.109
39	1.273 ± 0.080	0.967 ± 0.049	0.914 ± 0.039	1.033 ± 0.011	0.992 ± 0.010	0.976 ± 0.013	0.959 ± 0.105
40	1.042 ± 0.080	0.976 ± 0.046	1.063 ± 0.037	0.983 ± 0.011	0.981 ± 0.010	0.986 ± 0.014	1.163 ± 0.110

Notes. The uncertainties represent the uncertainty of the transfer function fit, not the full dispersion of the simulations. These transfer function values apply to each element of the spherical harmonic expansion. Their effect in the power spectrum is squared.

Appendix H: Degeneracy in bandpass mismatch and polarization templates

Table H.1. NPIPE polarization-template amplitudes.

	Template frequency		
Fitting frequency	30 GHz	217 GHz	353 GHz
44 GHz	0.221	0.227	−0.0236
70 GHz	0.0482	0.274	−0.0251
100 GHz	0.0133	0.338	−0.0215
143 GHz	0.00979	0.382	−0.0185
217 GHz	0.0114	n/a	0.0493

During reprocessing (Sect. 2.4) of the CMB channels, NPIPE fits the polarized frequency maps from the foreground channels as time-domain templates (Sect. 2.4.13). This allows fitting for the other time-domain templates over a temperature-only sky, breaking some significant degeneracies that otherwise render the large-scale polarization in the maps very noisy. It is tempting to ask, if the polarization templates can be combined into a high S/N estimate of the polarized sky at each of the CMB frequencies. Unfortunately the answer is negative for two reasons:

1. the polarization templates are degenerate with the bandpass-mismatch templates and;
2. the polarization template description does not support spectral index variation across the sky.

The fitted polarization template amplitudes are shown in Table H.1. It is straightforward to demonstrate that a polarization map constructed with these amplitudes is missing some of the total polarization at each of the CMB frequencies. The remainder is captured in the bandpass-mismatch templates that, through temperature-to-polarization leakage, translate into Galactic polarization resembling the channel-map templates. We were able to demonstrate this effect with a simplified simulation of the 217 GHz channel processing. In this test we simulated noise-free TOD and fed them to the NPIPE reprocessing. We also replaced the frequency-map-based polarization templates with the actual polarization template used in the simulation. One would expect this setup to yield unit amplitude for the fitted polarization template but instead we recovered 0.61. Since the simulation did not include systematics, it was possible to disable templates in the template matrix one by one and repeat the simulation. We found that the low fitting amplitude persisted despite disabling orbital-dipole fitting, zodiacal emission, and the whole hierarchy of the HFI transfer function residual templates. Once we disabled the bandpass-mismatch correction, the fitted polarization template amplitude jumped to 1.01. There is no way to repeat the test with flight data as we cannot disable the true bandpass mismatch.

The fact that the polarization templates are degenerate with the bandpass-mismatch correction may seem risky, as it could compromise the vital bandpass-mismatch correction and bias the Galactic polarization in the NPIPE CMB frequency maps (polarization templates are not fitted at 30 or 353 GHz). However, the potential bias is avoided by the two-step approach used in reprocessing the CMB frequencies.

1. During the first $N - 1$ iterations, the time-dependent gain or ADCNL correction and other templates are fitted while marginalizing over the bandpass-mismatch and polarization-template amplitudes. Fitting is done over an intensity-only sky, approximating that the polarized signal is fitted by the TOD templates. As long as the combination of these templates is a reasonable description of the polarization modulation in the TOD, the other templates are not affected.
2. During the final iteration, the gain solution is held fixed, fitting is done over a polarized IQU sky, and the polarization templates are not involved.

As the TOD are being calibrated and ADCNL-corrected, it matters only that the *total* polarization model is complete enough not to bias the time-varying gain and ADCNL solution. The fact that the polarization model is built from degenerate templates does not bias the solution, although it may slow down the convergence of the solver. The bandpass-mismatch correction is not affected by the degeneracy either, because it is ultimately solved over a polarized sky *without* the polarization prior.

Appendix I: Simulated uncertainty and bias

Comparison of simulated and real null (noise) maps in Sect. 5.3 shows that the overall uncertainty in the simulations agrees with the flight data. With 600 Monte Carlo realizations, we may also ask if the processing residuals have zero mean or if there are detectable biases, even much below the overall uncertainty. Figure I.1 shows the total (noise and systematics) uncertainty of 1-degree smoothed IQU maps; Fig. I.2 shows the corresponding bias. The uncertainty is measured as the per-pixel rms of the smoothed residual (output−input) maps, and the bias is the average of those same maps.

We remind the reader that the NPIPE simulations treat as fixed systematics such as bandpass mismatch, gain fluctuations, ADCNL, bolometric transfer-function residuals, and beam asymmetry. They are not drawn from a distribution, but rather applied to each Monte Carlo simulation the same way. This means that the residuals associated with these systematics do not average out across the simulations. The averaged residual in Fig. I.2 demonstrates the total level of systematic residuals without the statistical noise but should not be taken as a measurement of the actual residuals in the flight data maps.

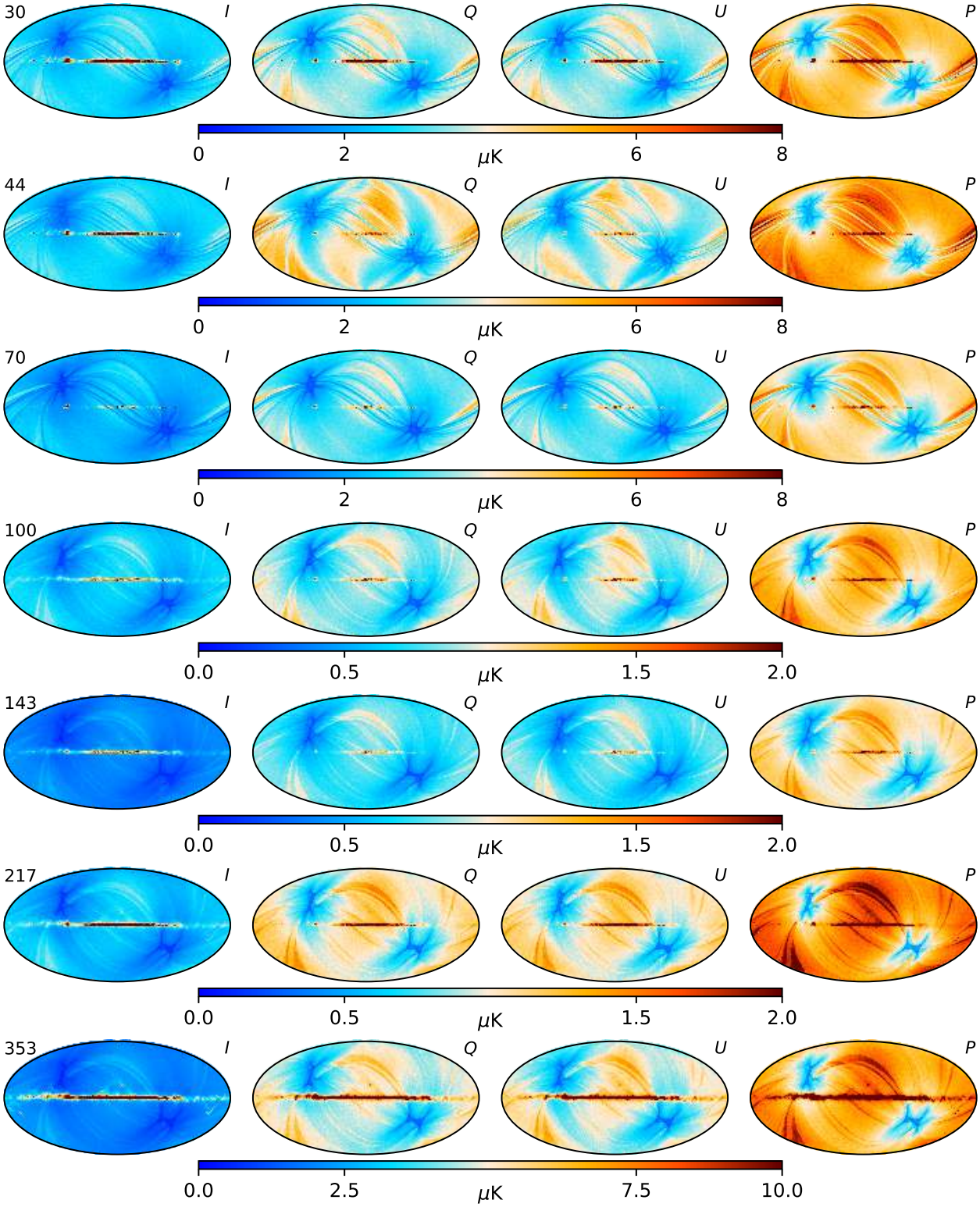


Fig. I.1. Simulation error. The rms of the residual maps smoothed to 1° is a measure of the total per-pixel uncertainty at degree scales and larger.

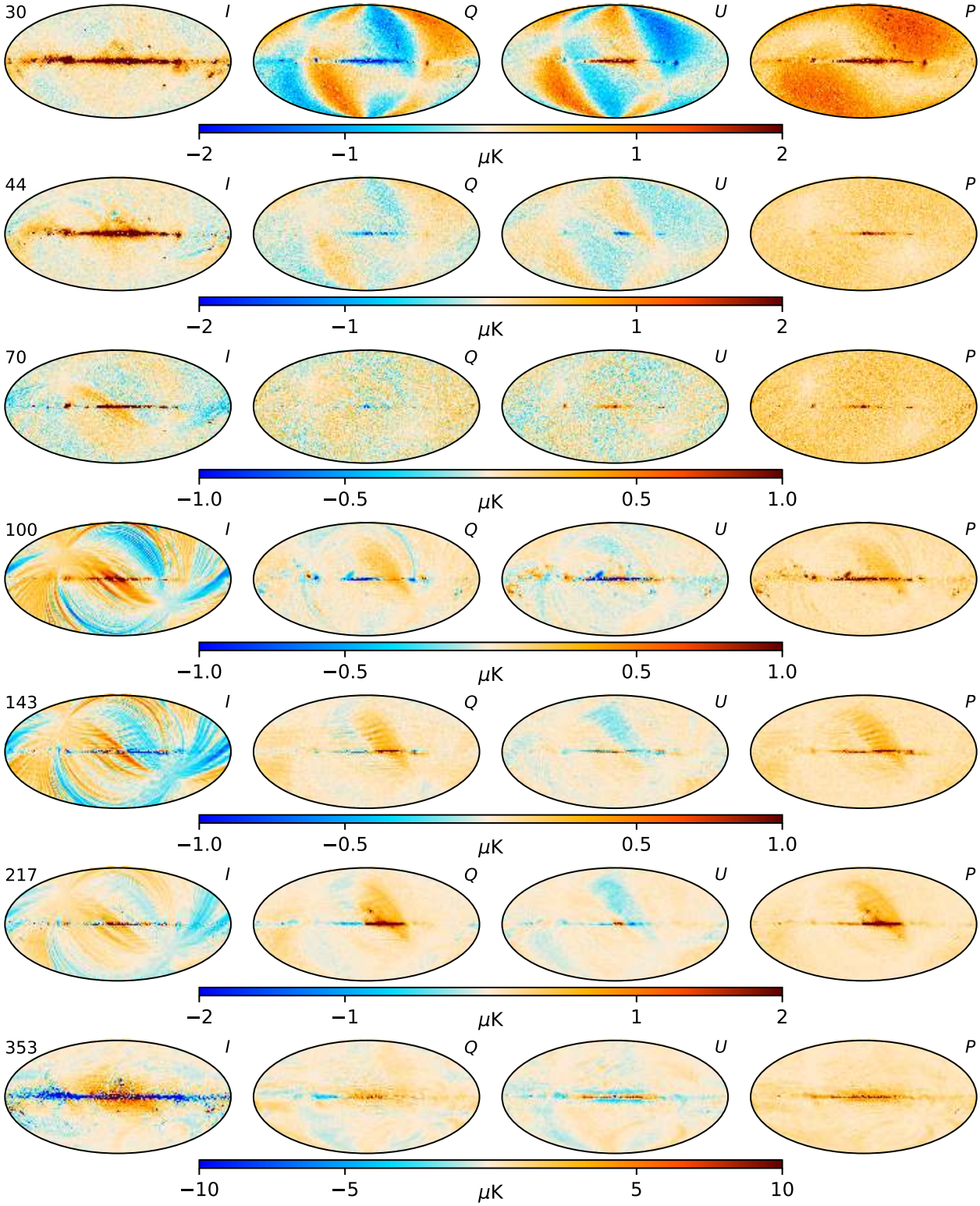


Fig. I.2. Simulation bias. The mean of the residual maps smoothed to 1° is a measure of persistent error for this particular realization of systematics. The magnitude of the bias can be compared to the total error shown in Fig. I.1. The dipole residual in the 30 GHz polarization maps is consistent with a small relative calibration error between the radiometers, and reflects the degeneracy between calibration and the large scale polarization due to the *Planck* scan strategy. The error translates to 44 GHz by means of the polarization prior. HFI maps exhibit a ringing structure coming from the transfer function residuals and a faint ADCNL error above and to the right of the Galactic centre.