**Table 12.** *TE* and *EE* dust contamination levels, $D_{\ell=500}$.

| Spectrum | Contamination level [$\mu$K$^2$] | | |
|---|---|---|---|
| | 100 GHz (G70) | 143 GHz (G50) | 217 GHz (G41) |
| $D^{TE}_{\ell=500}$ | | | |
| 100 GHz (G70) . . . . . . . | $0.14 \pm 0.042$ | $0.12 \pm 0.036$ | $0.3 \pm 0.09$ |
| 143 GHz (G50) . . . . . . . | | $0.24 \pm 0.072$ | $0.6 \pm 0.018$ |
| 217 GHz (G41) . . . . . . . | | | $1.8 \pm 0.54$ |
| $D^{EE}_{\ell=500}$ | | | |
| 100 GHz (G70) . . . . . . . | $0.06 \pm 0.012$ | $0.05 \pm 0.015$ | $0.11 \pm 0.033$ |
| 143 GHz (G50) . . . . . . . | | $0.1 \pm 0.02$ | $0.24 \pm 0.048$ |
| 217 GHz (G41) . . . . . . . | | | $0.72 \pm 0.14$ |

**Notes.** Values reported in the table correspond to the evaluation of the contamination level in each frequency by fitting the 353 GHz cross half-mission spectra against the CMB-corrected $353 \times 100$, $353 \times 143$ and $353 \times 217$ spectra over a range of multipoles. The CMB correction is obtained using the 100 GHz cross half-mission spectra (we have similar results at 143 GHz). Level reported here correspond to the amplitude of the contamination $\mathcal{D}_\ell$ at $\ell = 500$ in $\mu$K$^2$.

*The cosmic infrared background.* The CIB model has a number of differences from that used in Like13. First of all, it is now entirely parameterized by a single amplitude $\mathcal{D}^{\rm CIB}_{217}$ and a template $C^{\rm CIB}_\ell$:

$$\left(C^{\rm CIB}_{\nu\times\nu'}\right)_\ell = a^{\rm CIB}_\nu a^{\rm CIB}_{\nu'} C^{\rm CIB}_\ell \times \mathcal{D}^{\rm CIB}_{217}, \tag{26}$$

where the spectral coefficients $a^{\rm CIB}_\nu$ represent the CIB emission law normalized at $\nu = 217$ GHz.

In 2013, the template was an effective power-law model with a variable index with expected value $n = -1.37$ (when including the "highL" data from ACT and SPT). We did not assume any emission law and fitted the 143 GHz and 217 GHz amplitude, along with their correlation coefficient. The Planck Collaboration has studied the CIB in detail in Planck Collaboration XXX (2014) and now proposes a one-plus-two-halo model, which provides an accurate description of the *Planck* and IRAS CIB spectra from 3000 GHz down to 217 GHz. We extrapolate this model here, assuming it remains appropriate in describing the 143 GHz and 100 GHz data. The CIB emission law and template are computed following Planck Collaboration XXX (2014). The template power spectrum provided by this work has a very small frequency dependence that we ignore.

At small scales, $\ell > 2500$, the slope of the template is similar to the power law used in Like13. At larger scales, however, the slope is much shallower. This is in line with the variation we observed in 2013 on the power-law index of our simple CIB model when changing the maximum multipole. The current template is shown as the green line in the *TT* foreground component plots in Fig. 17.

In 2013, the correlation between the 143 GHz and 217 GHz CIB spectra was fitted, favouring a high correlation, greater than 90% (when including the "highL" data). The present model yields a fully correlated CIB between 143 GHz and 217 GHz.

We now include the the CIB contribution at 100 GHz, which was ignored in 2013. Another difference with the 2013 model is that the parameter controlling the amplitude at 217 GHz now directly gives the amplitude in the actual 217 GHz *Planck* band at $\ell = 3000$, i.e., it includes the colour correction. The ratio between the two is 1.33. The 2013 amplitude of the CIB contribution at $\ell = 3000$ (including the highL data) was $66 \pm 6.7$ $\mu$K$^2$, while our best estimate for the present analysis is $63.9 \pm 6.6$ $\mu$K$^2$ (PlanckTT+lowP).

*Point sources.* At the likelihood level, we cannot differentiate between the radio- and IR-point sources. We thus describe their combined contribution by their total emissivity per frequency pair,

$$\left(C^{\rm PS}_{\nu\times\nu'}\right)_\ell = \mathcal{D}^{\rm PS}_{\nu\times\nu'} / \mathcal{A}_{3000}, \tag{27}$$

where $\mathcal{D}_{\nu\times\nu'}$ is the amplitude of the point-source contribution in $\mathcal{D}_\ell$ at $\ell = 3000$. Contrary to 2013, we do not use a correlation parameter to represent the $143 \times 217$ point-source contribution; instead we use a free amplitude parameter. This has the disadvantage of not preventing a possible unphysical solution. However, it simplifies the parameter optimization, and it is easier to understand in terms of contamination amplitude.

*Kinetic SZ (kSZ).* We use the same model as in 2013. The kSZ emission is parameterized with a single amplitude and a fixed template from Trac et al. (2011),

$$\left(C^{\rm kSZ}_{\nu\times\nu'}\right)_\ell = C^{\rm kSZ}_\ell \times \mathcal{D}^{\rm kSZ}, \tag{28}$$

where $\mathcal{D}^{\rm kSZ}$ is the kSZ contribution at $\ell = 3000$.

*Thermal SZ (tSZ).* Here again, we use the same model as in 2013. The tSZ emission is also parameterized by a single amplitude and a fixed template using the $\epsilon = 0.5$ model from Efstathiou & Migliaccio (2012),

$$\left(C^{\rm tSZ}_{\nu\times\nu'}\right)_\ell = a^{\rm tSZ}_\nu a^{\rm tSZ}_{\nu'} C^{\rm tSZ}_\ell \times \mathcal{D}^{\rm tSZ}_{143}, \tag{29}$$

where $a^{\rm tSZ}_\nu$ is the thermal Sunyaev-Zeldovich spectrum, normalized to $\nu_0 = 143$ GHz and corrected for the *Planck* bandpass colour corrections. Ignoring the bandpass correction, we recall that the tSZ spectrum is given by

$$a^{\rm tSZ}_\nu = \frac{f(\nu)}{f(\nu_0)}, \; f(\nu) = \left(x \coth\left(\frac{x}{2}\right) - 4\right), \; x = \frac{h\nu}{k_{\rm B} T_{\rm cmb}}. \tag{30}$$

*Thermal SZ × CIB correlation.* Following Like13 the cross-correlation between the thermal SZ and the CIB, tSZ× CIB, is parameterized by a single correlation parameter, $\xi$, and a fixed template from Addison et al. (2012),

$$\left(C^{\rm tSZ\times CIB}_{\nu\times\nu'}\right)_\ell = \xi\,\sqrt{\mathcal{D}^{\rm tSZ}_{143}\,\mathcal{D}^{\rm CIB}_{217}} \times\left(a^{\rm tSZ}_\nu\,a^{\rm CIB}_{\nu'} + a^{\rm tSZ}_{\nu'}\,a^{\rm CIB}_\nu\right) \times C^{\rm tSZ\times CIB}_\ell, \tag{31}$$

where $a^{\rm tSZ}_\nu$ is the thermal Sunyaev-Zeldovich spectrum, corrected for the *Planck* bandpass colour corrections and $a^{\rm CIB}_\nu$ is the CIB spectrum, rescaled at $\nu$ = 217 GHz as in the previous paragraphs.

*SZ prior.* The kinetic SZ, the thermal SZ, and its correlation with the CIB are not constrained accurately by the *Planck* data alone. Besides, the tSZ×CIB level is highly correlated with the amplitude of the tSZ. In 2013, we reduced the degeneracy between those parameters and improved their determination by adding the ACT and SPT data. In 2015, we instead impose a Gaussian prior on the tSZ and kSZ amplitudes, inspired by the constraints set by these experiments. From a joint analysis of the *Planck* 2013 data with those from ACT and SPT, we obtain

$$\mathcal{D}^{\rm kSZ} + 1.6\mathcal{D}^{\rm tSZ} = (9.5 \pm 3)\,\mu{\rm K}^2, \tag{32}$$

in excellent agreement with the estimates from Reichardt et al. (2012), once they are rescaled to the *Planck* frequencies (see Planck Collaboration XIII 2016, for a detailed discussion).

As can be seen in Fig. 17, the kSZ, tSZ, and tSZ×CIB correlations are always dominated by the dust, CIB, and point-source contributions.

### 3.4. Instrumental modelling

The following sections describe the instrument modelling elements of the model vector, addressing the issues of calibration and beam uncertainties in Sects. 3.4.1–3.4.3, and describing the noise properties in Sect. 3.4.4. For convenience, Table 10 defines the symbol used for the calibration parameters and the priors later used for exploring them.

#### 3.4.1. Power spectra calibration uncertainties

As in 2013, we allow for a small recalibration of the different frequency power spectra, in order to account for residual uncertainties in the map calibration process. The mixing matrix in the model vector from Eq. (14) can be rewritten as

$$\begin{aligned}\left(M^{XY}_{ZW,\nu\times\nu'}\right)_\ell(\theta_{\rm inst}) &= G^{XY}_{\nu\times\nu'}(\theta_{\rm calib})\left(M^{XY,\rm other}_{ZW,\nu\times\nu'}\right)_\ell(\theta_{\rm other}),\\ G^{XY}_{\nu\times\nu'}(\theta_{\rm calib}) &= \frac{1}{y_{\rm P}^2}\left(\frac{1}{2\sqrt{c^{XX}_\nu c^{YY}_{\nu'}}} + \frac{1}{2\sqrt{c^{XX}_{\nu'} c^{YY}_\nu}}\right),\end{aligned} \tag{33}$$

where $c^{XX}_\nu$ is the calibration parameter for the $XX$ power spectrum at frequency $\nu$, $X$ being either $T$ or $E$, and $y_{\rm P}$ is the overall *Planck* calibration. We ignore the $\ell$-dependency of the weighting function between the $TE$ and $ET$ spectra at different frequencies that are added to form an effective cross-frequency $TE$ cross-spectrum. As in 2013, we use the $TT$ at 143 GHz as our inter-calibration reference, so that $c^{TT}_{143} = 1$.

We further allow for an overall *Planck* calibration uncertainty, whose variation is constrained by a tight Gaussian prior,

$$y_{\rm P} = 1 \pm 0.0025. \tag{34}$$

This prior corresponds to the estimated overall uncertainty, which is discussed in depth in Planck Collaboration I (2016).

The calibration parameters can be degenerate with the foreground parameters, in particular the point sources at high $\ell$ (for $TT$) and the Galaxy for 217 GHz at low $\ell$. We thus proceed as in 2013, and measure the calibration refinement parameters on the large scales and on small sky fractions near the Galactic poles. We perform the same estimates on a range of Galactic masks (G20, G30, and G41) restricted to different maximum multipoles (up to $\ell$ = 1500). The fits are performed either by minimizing the scatter between the different frequency spectra, or by using the `SMICA` algorithm (see Planck Collaboration VI 2014, Sect. 7.3) with a freely varying CMB and generic foreground contribution. For the $TT$ spectra, we obtained in both cases very similar recalibration estimates, from which we extracted the conservative Gaussian priors on recalibration factors,

$$c^{TT}_{100} = 0.999 \pm 0.001, \tag{35}$$

$$c^{TT}_{217} = 0.995 \pm 0.002. \tag{36}$$

These are compatible with estimates made at the map level, but on the whole sky; see Planck Collaboration VIII (2016).

#### 3.4.2. Polarization efficiency and angular uncertainty

We now turn to the polarization recalibration case. The signal measured by an imperfect PSB is given by

$$d = G(1+\gamma)\left[I + \rho(1+\eta)\left(Q\cos 2(\phi+\omega) + U\sin 2(\phi+\omega)\right)\right] + n, \tag{37}$$

where $I$, $Q$, and $U$ are the Stokes parameters; $n$ is the instrumental noise; $G$, $\rho$, and $\phi$ are the nominal photometric calibration factor, polar efficiency, and direction of polarization of the PSB; and $\gamma$, $\eta$, and $\omega$ are the (small) errors made on each of them (see, e.g., Jones et al. 2007). Due to these errors, the measured cross-power spectra of maps $a$ and $b$ are then contaminated by a spurious signal given by

$$\Delta C^{TT}_\ell = (\gamma_a + \gamma_b)\,C^{TT}_\ell, \tag{38a}$$

$$\Delta C^{TE}_\ell = \left(\gamma_a + \gamma_b + \eta_b - 2\omega_b^2\right)C^{TE}_\ell, \tag{38b}$$

$$\Delta C^{EE}_\ell = \left(\gamma_a + \gamma_b + \eta_a + \eta_b - 2\omega_a^2 - 2\omega_b^2\right)C^{EE}_\ell + 2\left(\omega_a^2 + \omega_b^2\right)C^{BB}_\ell, \tag{38c}$$

where $\gamma_x$, $\eta_x$, and $\omega_x$, for $x = a, b$, are the effective instrumental errors for each of the two frequency-averaged maps. Pre-flight measurements of the HFI polarization efficiencies, $\rho$, had uncertainties $|\eta_x| \approx 0.3\%$, while the polarization angle of each PSB is known to $|\omega_x| \approx 1°$ (Rosset et al. 2010). Analysis of the 2015 maps shows the relative photometric calibration of each detector at 100 to 217 GHz to be known to about $|\gamma_x| = 0.16\%$ at worst, with an absolute orbital dipole calibration of about 0.2%, while analysis of the Crab Nebula observations showed the polarization uncertainties to be consistent with the pre-flight measurements (Planck Collaboration VIII 2016).

Assuming $C^{BB}_\ell$ to be negligible, and ignoring $\omega^2 \ll |\eta|$ in Eq. (38), the Gaussian priors on $\gamma$ and $\eta$ for each frequency-averaged polarized map would have rms of $\sigma_\gamma = 2 \times 10^{-3}$ and

$\sigma_\eta = 3 \times 10^{-3}$. Adding those uncertainties in quadrature, the auto-power spectrum recalibration $c_\nu^{EE}$ introduced in Eq. (33) would be given, for an equal-weight combination of $n_\mathrm{d} = 8$ polarized detectors, by

$$c_\nu^{EE} = 1 \pm 2\sqrt{\frac{\sigma_\gamma^2 + \sigma_\eta^2}{n_\mathrm{d}}} = 1 \pm 0.0025. \tag{39}$$

The most accurate recalibration factors for $TE$ and $EE$ could therefore be somewhat different from $TT$. We found, though, that setting the $EE$ recalibration parameter to unity or implementing those priors makes no difference with respect to cosmology; i.e., we recover the same cosmological parameters, with the same uncertainties. Thus, for the baseline explorations, we fixed the $EE$ recalibration parameter to unity,

$$c_\nu^{EE} = 1, \tag{40}$$

and the uncertainty on $TE$ comes only from the $TT$ calibration parameter through Eq. (33).

We also explored the case of much looser priors, and found that best-fit calibration parameters deviate very significantly, and reach values of several percent (between 3% and 12% depending on the frequencies and on whether we fit the $EE$ or $TE$ case). This cannot be due to the instrumental uncertainties embodied in the prior. In the absence of an informative prior, this degree of freedom is used to minimize the differences between frequencies that stem from other effects, not included in the baseline modelling.

The next section introduces one such effect, the temperature-to-polarization leakage, which is due to combining detectors with different beams without accounting for it at the map-making stage (see Sect. 3.4.3). But anticipating the results of the analysis described in Appendix C.3.5, we note that when the calibration and leakage parameters are explored simultaneously without priors, they remain in clear tension with the priors (even if the level of recalibration decreases slightly, by typically 2%, showing the partial degeneracy between the two). In other words, when calibration and leakage parameters are both explored with their respective priors, there is evidence of residual unmodelled systematic effects in polarization – to which we will return.

### 3.4.3. Beam and transfer function uncertainties

The power spectra from map pairs are corrected by the corresponding effective beam window functions before being confronted with the data model. However, these window functions are not perfectly known, and we now discuss various related sources of errors and uncertainties, the impact of which on the reconstructed $C_\ell$s is shown in Fig. 21.

*Sub-pixel effects.* The first source of error, the so-called "sub-pixel" effect, discussed in detail in Like13, is a result of the *Planck* scanning strategy and map-making procedure. Scanning along rings with very low nutation levels can result in the centroid of the samples being slightly shifted from the pixel centres; however, the map-making algorithm assigns the mean value of samples in the pixel to the centre of the pixel. This effect, similar to the gravitational lensing of the CMB, has a non-diagonal influence on the power spectra, but the correction can be computed given the estimated power spectra for a given data selection, and recast into an additive, fixed component. We showed in Like13 that including this effect had little impact on the cosmological parameters measured by *Planck*.

*Masking effect.* A second source of error is the variation, from one sky pixel to another, of the effective beam width, which is averaged over all samples falling in that pixel. While all the `HEALPix` pixels have the same surface area, their shape – and therefore their moment of inertia (which drives the pixel window function) – depends on location, as shown in Fig. 22, and therefore makes the effective beam window function depend on the pixel mask considered. Of course the actual sampling of the pixels by *Planck* leads to individual moments of inertia slightly different from the intrinsic values shown here, but spot-check comparisons of this semi-analytical approach used by `QuickBeam` with numerical simulations of the actual scanning by `FEBeCoP` showed agreement at the $10^{-3}$ level for $\ell < 2500$ on the resulting pixel window functions for sky coverage varying from 40 to 100%.

In the various Galactic masks used here (Figs. 12–13) the contribution of the unmasked pixels to the total effective window function departs from the full-sky average (which is not included in the effective beam window functions), and we therefore expect a different effective transfer function for each mask. We ignored this dependence and mitigated its effect by using transfer functions computed with the Galactic mask G60 which retains an effective sky fraction (including the mask apodization) of $f_\mathrm{sky} = 60\%$, not too different from the sky fractions $f_\mathrm{sky}$ between 41 and 70% (see Sect. 3.2.2) used for computing the power spectra.

Figure 21 compares the impact of these two sources of uncertainty on the stated *Planck* statistical error bars for $\Delta\ell = 30$. It shows that, for $\ell < 1800$ where most of the information on $\Lambda$CDM lies, the error on the $TT$ power spectra introduced by the sub-pixel effect and by the sky-coverage dependence are less than about 0.1%, and well below the statistical error bars of the binned $C_\ell$. In the range $1800 \le \ell \le 2500$, which helps constrain one-parameter extensions to base $\Lambda$CDM (such as $N_\mathrm{eff}$), the relative error can reach 0.4% (note as a comparison that the high-$\ell$ ACT experiment states a statistical error of about 3% on the bin $2340 \le \ell \le 2540$, Das et al. 2014). The bottom panel shows the Monte Carlo error model of the beam window functions, which provides negligible ($\ell$-coupled) uncertainties. Even if this model is somewhat optimistic, since it does not include the effect of the ADC non-linearities and the colour-correction effect of beam measurements on planets (Planck Collaboration VII 2016), we note that even expanding them by a factor of 10 keeps them within the statistical uncertainty of the power spectra.

*Modelling the uncertainties.* As in the 2013 analysis, the beam uncertainty eigenmodes were determined from 100 (improved) Monte Carlo (MC) simulations of each planet observation used to measure the scanning beams, then processed through the same `QuickBeam` pipeline as the nominal beam to determine their effective angular transfer function $B(\ell)$. Thanks to the use of Saturn and Jupiter transits instead of the dimmer Mars used in 2013, the resulting uncertainties are now significantly smaller (Planck Collaboration VII 2016).

For each pair of frequency maps (and frequency-averaged beams) used in the present analysis, a singular-value decomposition (SVD) of the correlation matrix of 100 Monte Carlo based $B(\ell)$ realizations was performed over the ranges $[0, \ell_\mathrm{max}]$ with $\ell_\mathrm{max} = (2000, 3000, 3000)$ at (100, 143, 217 GHz), and the five leading modes were kept, as well as their covariance matrix (since the error modes do exhibit Gaussian statistics). We therefore have, for each pair of beams, five $\ell$-dependent templates,

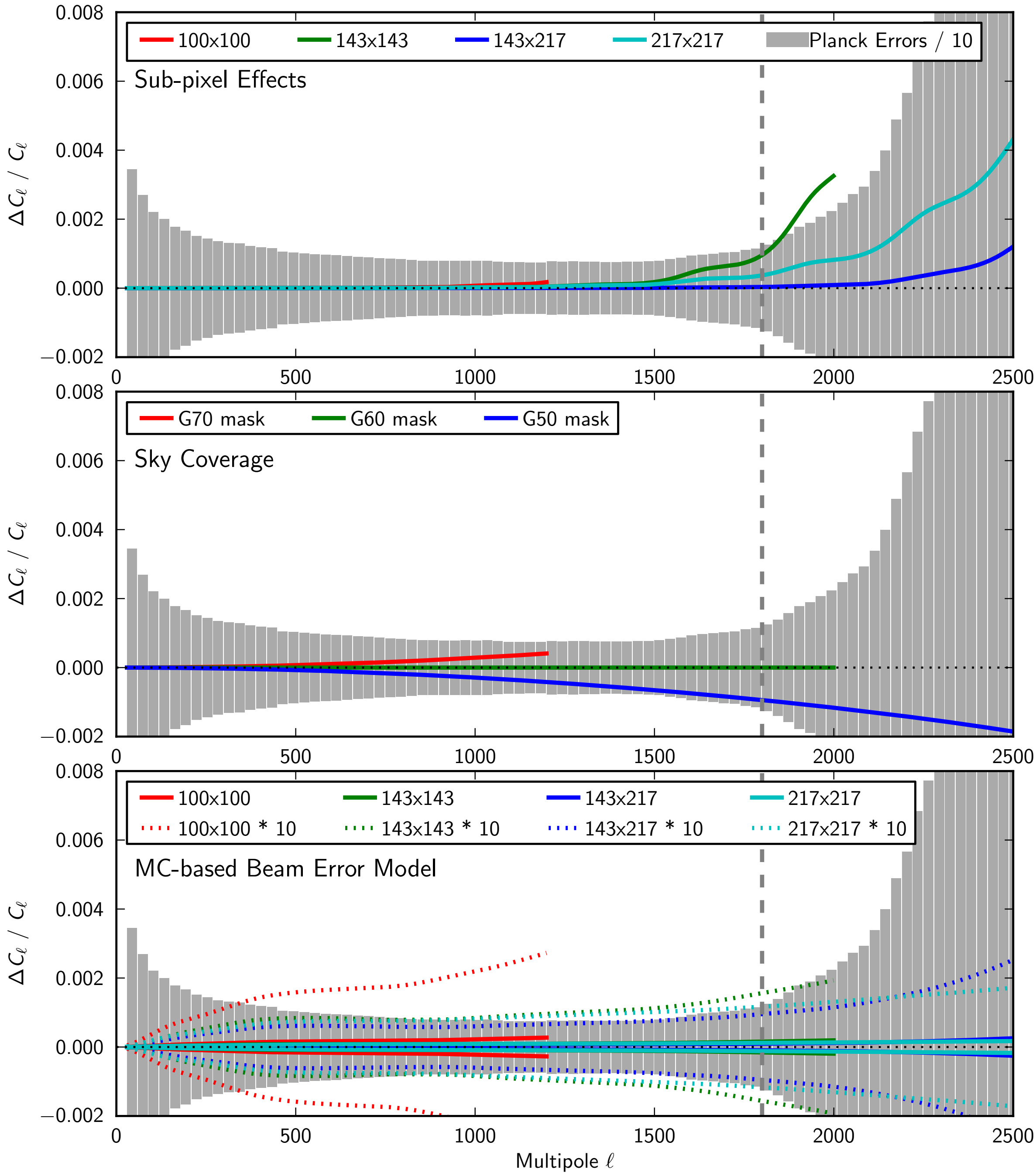


**Fig. 21.** Contribution of various beam-window-function-related errors and uncertainties to the $C_\ell$ relative error. In each panel, the grey histogram shows the relative statistical error on the *Planck* CMB $TT$ binned power spectrum (for a bin width $\Delta\ell = 30$) *divided by 10*, while the vertical grey dashes delineate the range $\ell < 1800$ that is most informative for base $\Lambda$CDM. *Top*: estimation of the error made by ignoring the sub-pixel effects for a fiducial $C_\ell$ including the CMB and CIB contributions. *Middle*: error due to the sky mask, for the Galactic masks used in the $TT$ analysis. *Bottom*: current beam window function error model, shown at $1\sigma$ (solid lines) and $10\sigma$ (dotted lines).

each associated with a Gaussian amplitude centred on 0, and a covariance matrix coupling all of them.

Including the beam uncertainties in the mixing matrix of Eq. (14) gives

$$\left(M^{XY}_{ZW,\nu\times\nu'}\right)_\ell(\theta_{\rm inst}) = \left(M^{XY,{\rm other}}_{ZW,\nu\times\nu'}\right)_\ell(\theta_{\rm other})\,\left(\Delta W^{ZW}_{\nu\times\nu'}\right)_\ell(\theta_{\rm beam}),$$
$$\left(\Delta W^{ZW}_{\nu\times\nu'}\right)_\ell(\theta_{\rm beam}) = \exp\sum_{i=1}^{5} 2\,\theta^{ZW,i}_{\nu\times\nu'}\left(E^{ZW,i}_{\nu\times\nu'}\right)_\ell, \qquad (41)$$

where $\left(\Delta W^{ZW}_{\nu\times\nu'}\right)_\ell(\theta_{\rm beam})$ stands for the beam error built from the eigenmodes $\left(E^{ZW,i}_{\nu\times\nu'}\right)_\ell$. The quadratic sum of the beam eigenmodes is shown in Fig. 21. This is much smaller (less than a percent) than the combined $TT$ spectrum error bars. This contrasts with the 2013 case where the beam uncertainties were greater; for instance, for the 100, 143, and 217 GHz channel maps, the rms of the $W(\ell) = B(\ell)^2$ uncertainties at $\ell = 1000$ dropped from $(61, 23, 20) \times 10^{-4}$ to $(2.2, 0.84, 0.81) \times 10^{-4}$, respectively. The fact that beam uncertainties are sub-dominant in the total error

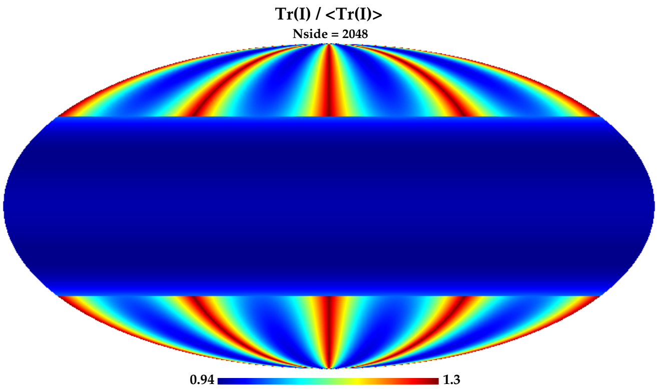


**Fig. 22.** Map of the relative variations of the trace of the `HEALPix` pixel moment of inertia tensor at $N_{\rm side} = 2048$ in Galactic coordinates.

budget is even more pronounced in polarization, where noise is higher. We use the beam modes computed from temperature data, combined with appropriate weights when used as parameters affecting the $TE$ and $EE$ spectra.

As in 2013, instead of including the beam error in the vector model, we include its contribution to the covariance matrix, linearizing the vector model so that

$$\left(C^{XY}_{\nu\times\nu'}\right)_\ell(\theta) = \left(C^{XY}_{\nu\times\nu'}\right)_\ell(\theta, \theta_{\rm beam}=0) + \left(\Delta W^{ZW}_{\nu\times\nu'}\right)_\ell(\theta_{\rm beam})\left(C^{XY}_{\nu\times\nu'}\right)^*_\ell, \tag{42}$$

where $\left(C^{XY}_{\nu\times\nu'}\right)^*_\ell$ is the fiducial spectrum $XY$ for the pair of frequencies $\nu\times\nu'$ obtained using the best cosmological and foreground model. We can then marginalize over the beam uncertainty, enlarging the covariance matrix to obtain

$$\mathsf{C}_{\rm beam\ marg.} = \mathsf{C} + \boldsymbol{C}^*\left\langle\Delta\boldsymbol{W}\Delta\boldsymbol{W}^{\sf T}\right\rangle\boldsymbol{C}^{*\sf T}, \tag{43}$$

where $\left\langle\Delta\boldsymbol{W}\Delta\boldsymbol{W}^{\sf T}\right\rangle$ is the Monte Carlo based covariance matrix, restricted to its first five eigenmodes.

In 2013, beam errors were marginalized for all the modes except the two greatest of the $100\times100$ spectrum. In the present release we instead marginalize over all modes in $TT$, $TE$, and $EE$. We also performed a test in which we estimated the amplitudes for all of the first five beam eigenmodes in $TT$, $TE$, and $EE$, and found no indication of any beam error contribution (see Sect. 4.1.3 and Fig. 35).

*Temperature-to-polarization leakage.* Polarization measurements are differential by nature. Therefore any unaccounted discrepancy in combining polarized detectors can create some leakage from temperature to polarization (Hu et al. 2003). Sources of such discrepancies in the current HFI processing include, but are not limited to: differences in the scanning beams that are ignored during the map-making; differences in the noise level, because of the individual inverse noise weighting used in HFI; and differences in the number of valid samples.

For this release, we did not attempt to model and remove a priori the form and amplitude of this coupling between the measured $TT$, $TE$, and $EE$ spectra; we rather estimate the residual effect by fitting *a posteriori* in the likelihood some flexible template of this coupling, parameterized by some new nuisance parameters that we now describe.

The temperature-to-polarization leakage due to beam mismatch is *assumed* to affect the spherical harmonic coefficients via

$$a^T_{\ell m} \longrightarrow a^T_{\ell m}, \tag{44a}$$

$$a^E_{\ell m} \longrightarrow a^E_{\ell m} + \varepsilon(\ell)a^T_{\ell m}, \tag{44b}$$

and, for each map, the spurious polarization power spectrum $C^{XY}_\ell \equiv \sum_m a^X_{\ell m}a^{Y*}_{\ell m}/(2\ell+1)$ is modelled as

$$\Delta C^{TE}_\ell = \varepsilon(\ell)\,C^{TT}_\ell, \tag{45a}$$

$$\Delta C^{EE}_\ell = \varepsilon^2(\ell)\,C^{TT}_\ell + 2\varepsilon(\ell)\,C^{TE}_\ell. \tag{45b}$$

Here $\varepsilon_\ell$ is a polynomial in multipole $\ell$ determined by the effective beam of the detector-assembly measuring the polarized signal. Considering an effective beam map $b(\hat{\boldsymbol{n}})$ (rotated so that it is centred on the north pole), its spherical harmonic coefficients are defined as $b_{\ell m} \equiv \int {\rm d}\hat{\boldsymbol{n}}\, b(\hat{\boldsymbol{n}})\, Y^*_{\ell m}(\hat{\boldsymbol{n}})$. As a consequence of the *Planck* scanning strategy, pixels are visited approximately every six months, with a rotation of the focal plane by 180°, and we expect $b_{\ell m}$ to be dominated by even values of $m$, and especially the modes $m = 2$ and 4, which describe the beam ellipticity. As noted by, e.g., Souradeep & Ratra (2001) for elliptical Gaussian beams, the *Planck*-HFI beams for a detector $d$ obey

$$b^{(d)}_{\ell m} \simeq \beta^{(d)}_m \ell^m b^{(d)}_{\ell 0}. \tag{46}$$

We therefore fit the spectra using a fourth-order polynomial

$$\varepsilon(\ell) = \varepsilon_0 + \varepsilon_2\ell^2 + \varepsilon_4\ell^4, \tag{47}$$

treating the coefficients $\varepsilon_0$, $\varepsilon_2$, and $\varepsilon_4$ as nuisance parameters in the MCMC analysis. Tests performed on detailed simulations of *Planck* observations with known mismatched beams have shown that Eqs. (45) and (47) describe the power leakage due to beam mismatch with an accuracy of about 20% in the $\ell$ range 100–2000.

The equations above suggest that the same polynomial $\varepsilon$ can describe the contamination of the $TE$ and $EE$ spectra for a given pair of detector sets. But in the current `Plik` analysis, the $TE$ cross-spectrum of two different maps $a$ and $b$ is the inverse-variance-weighted average of the cross-spectra $T_aE_b$ and $T_bE_a$, while $EE$ is simply $E_aE_b$. In addition, the temperature maps include the signal from SWBs, which is obviously not the case for the $E$ maps. We therefore describe the $TE$ and $EE$ corrections by different $\varepsilon$ parameters. Similarly, we treated the parameters for the $EE$ cross-frequency spectra as being uncorrelated with the parameters for the auto-frequency ones.

The leakage is driven by the discrepancy between the individual effective beams $b^{(d)}_{\ell m}$ making up a detector assembly, coupled with the details of the scanning strategy and relative weight of each detector. If we assumed a perfect knowledge of the beams, precise – but not necessarily accurate – numerical predictions of the leakage would be possible. However, we preferred to adopt a more conservative approach in which the leakage was free to vary over a range wide enough to enclose the true value. On the other hand, in order to limit the unphysical range of variations permitted by so many nuisance parameters, we need priors on the $\varepsilon_m$ terms used in the Monte Carlo explorations. We assume Gaussian distributions of zero mean with a standard deviation $\sigma_m$ representative of the dispersion found in simulations of the effect with realistic instrumental parameters. We found $\sigma_0 = 1\times10^{-5}$, $\sigma_2 = 1.25\times10^{-8}$, and $\sigma_4 = 2.7\times10^{-15}$. This procedure ignores correlations between terms of different $m$, and is therefore likely substantially too permissive.

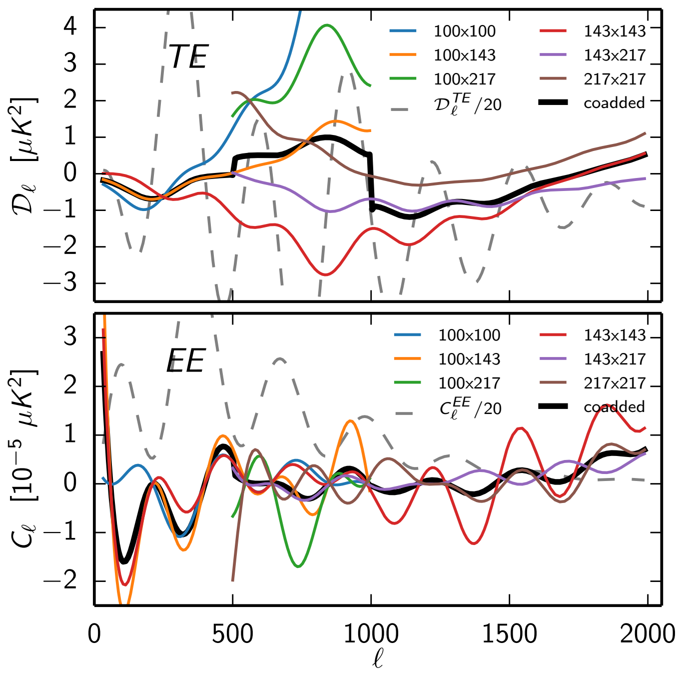


**Fig. 23.** Best fit of the power spectrum leakage due to the beam mismatch for $TE$ (Eq. (45a), *upper panel*) and $EE$ (Eq. (45b), *lower panel*). In each case, we show the correction for individual cross-spectra (coloured thin lines) and the co-added correction (black line). The individual cross-spectra corrections are only shown in the range of multipoles where the data from each particular pair is used. The individual correction can be much higher than the co-added correction. The co-added correction is dominated by the best S/N pair for each multipole. For example, up to $\ell = 500$, the $TE$ co-added correction is dominated by the $100 \times 143$ contribution. The grey dashed lines show the $TE$ and $EE$ best-fit spectra rescaled by a factor of 20, to give an idea of the location of the model peaks.

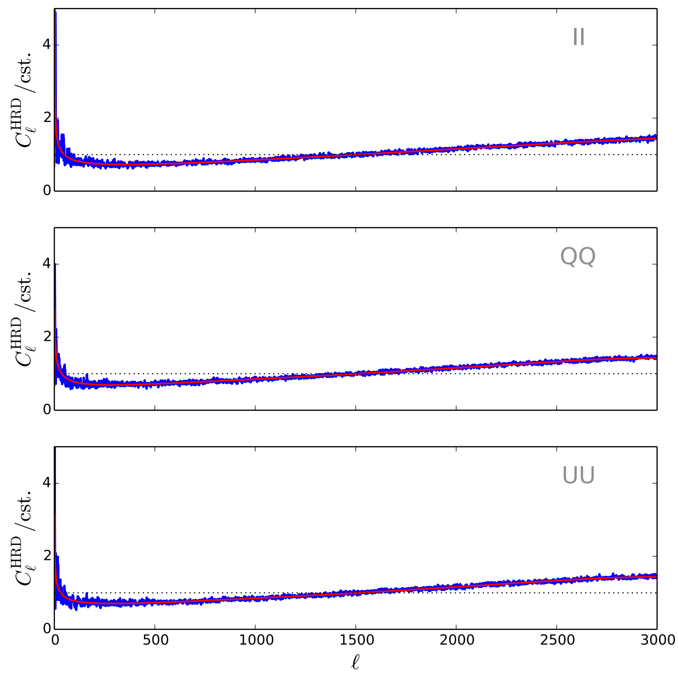


**Fig. 24.** Deviations from a white noise power spectrum induced by noise correlations. We show half-ring difference power spectra for 100 GHz half-mission 1 maps (blue lines) of Stokes parameters $I$ (*top panel*), $Q$ (*middle panel*), and $U$ (*bottom panel*). The best-fitting analytical model of the form Eq. (48) is over-plotted in red.

Another way of deriving the beam leakage would be to use a cosmological prior, i.e., by finding the best fit when holding the cosmological parameters fixed at their best-fit values for base ΛCDM. Figure 23 shows the result of this procedure for the cross-frequency pairs. The figure also shows the implied correction for the co-added spectra. This correction is dominated by the pair with the highest S/N at each multipole. The fact that different sets are used in different $\ell$-ranges leads to discontinuities in the correction template of the co-added spectrum. As can be seen in the figure, the co-added beam-leakage correction, of order $\mu K^2$, is much smaller than the individual corrections, which partially compensate each other on average (but improve the agreement between the individual polarized cross-frequency spectra).

It is shown in Appendix C.3.5 that neither procedure is fully satisfactory. The cosmological prior leads to nuisance parameters that vastly exceed the values allowed by the physical priors, and the physical priors are clearly overly permissive (leaving the cosmological parameters unchanged but with doubled error bars for some parameters). In any case, the agreement between the different cross-spectra remains much poorer in polarization than in temperature (see Sect. 4.4, Fig. 40, and Appendix C.3.5); they present oscillatory features similar to the ones produced by our beam leakage model, but the model is clearly not sufficient. For lack of a completely satisfactory global instrumental model, this correction is only illustrative and it is not used in the baseline likelihood.

### 3.4.4. Noise modelling

To predict the variance of the empirical power spectra, we need to model the noise properties of all maps used in the construction of the likelihood. As described in detail in Planck Collaboration VII (2016) and Planck Collaboration VIII (2016), the *Planck* HFI maps have complicated noise properties, with noise levels varying spatially and with correlations between neighbouring pixels along the scanning direction.

For each channel, full-resolution noise variance maps are constructed during the map-making process (Planck Collaboration VIII 2016). They provide an approximation to the diagonal elements of the true $n_{\rm pix} \times n_{\rm pix}$ noise covariance matrix for Stokes parameters $I$ (temperature only), or $I$, $Q$, and $U$ (temperature and polarization). While it is possible to capture the anisotropic nature of the noise variance with these objects, noise correlations between pixels remain unmodelled. To include deviations from a white-noise power spectrum, we therefore make use of half-ring difference maps. Choosing the 100 GHz map of the first half-mission as an example, we show the scalar (spin-0) power spectra of the three temperature and polarization maps in Fig. 24, rescaled by arbitrary constants. We find that the logarithm of the HFI noise power spectra as given by the half-ring difference maps can be accurately parameterized using a fourth-order polynomial with an additional logarithmic term,

$$\log(C_\ell^{\rm HRD}) = \sum_{i=0}^{4} \alpha_i\, \ell^i + \alpha_5 \log(\ell + \alpha_6). \tag{48}$$

Since low-frequency noise and processing steps like deglitching leave residual correlations between both half-ring maps, noise estimates derived from their difference are biased low, at the percent level at high-$\ell$ (where it was first detected and understood, see Planck Collaboration VI 2014). We correct for this

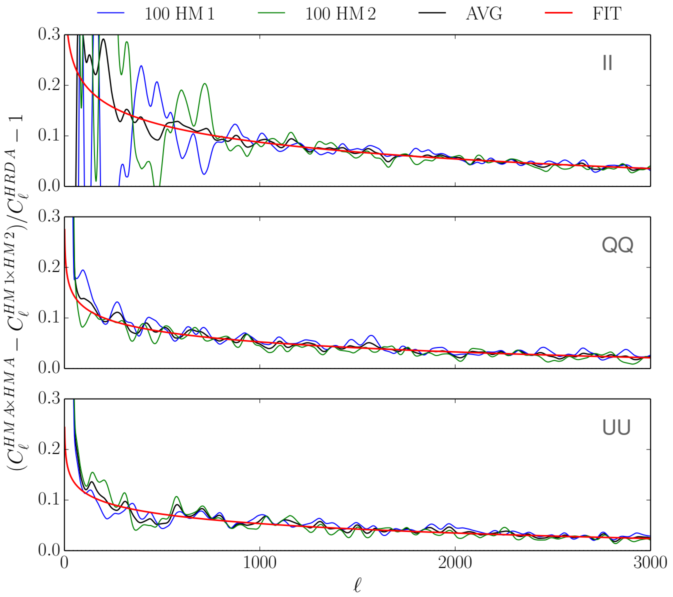


**Fig. 25.** Difference between auto and cross-spectra for the 100 GHz half-mission maps, divided by the noise estimate from half-ring difference maps (blue and green lines). Noise estimates derived from half-ring difference maps are biased low. We fit the average of both half-mission curves (black line) with a power law model (red line). The analysis procedure is applied to the Stokes parameter maps $I$, $Q$, and $U$ (*top to bottom*). All data power spectra are smoothed.

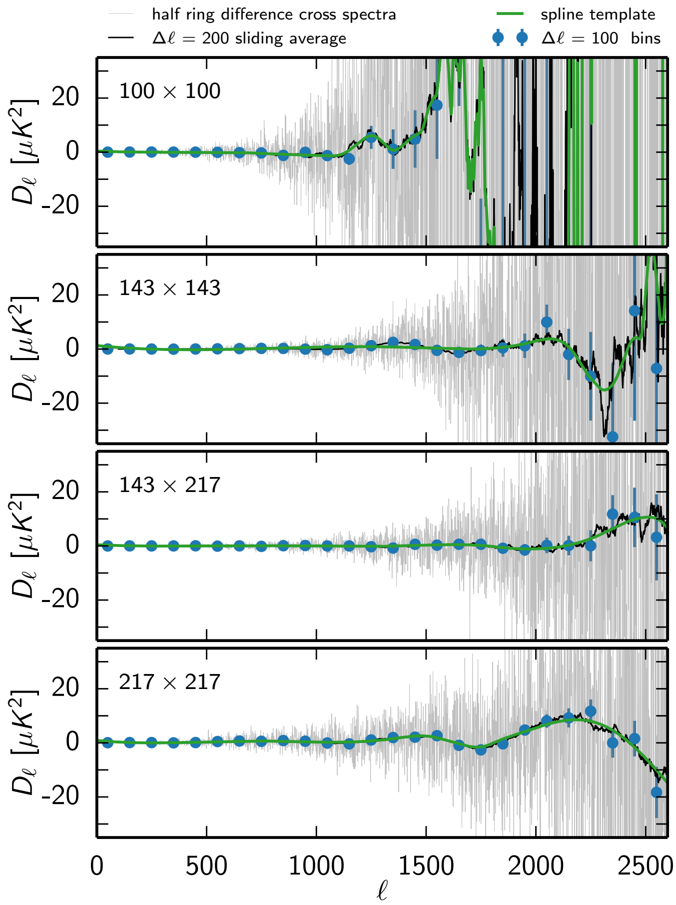


**Fig. 26.** Correlated noise model. In grey are shown the cross-detector $TT$ spectra of the half-ring difference maps. The black line show the same, smoothed by a $\Delta\ell = 200$ sliding average, while the blue data points are a $\Delta\ell = 100$ binned version of the grey line. Error bars simply reflect the scatter in each bin. The green line is the spline-smoothed version of the data that we use as our correlated noise template.

effect by comparing the difference of auto-power-spectra and cross-spectra (assumed to be free of noise bias) at a given frequency with the noise estimates obtained from half-ring difference maps. As shown in Fig. 25, we use a a power-law model with free spectral index to fit the average of the ratios of the first and second half-mission results to the half-ring difference spectrum, using the average to nullify chance correlations between signal and noise:

$$C_\ell^{\rm bias} = \alpha_0\,\ell^{\alpha_1} + \alpha_2. \quad (49)$$

At a multipole moment of $\ell = 1000$, we obtain correction factors for the temperature noise estimate obtained from half-ring difference maps of 9%, 10%, and 9% at 100, 143, and 217 GHz, respectively.

In summary, our HFI noise model is obtained as follows. For each map, we capture the anisotropic nature of the noise amplitude by using the diagonal elements of the pixel-space noise covariance matrix. The corresponding white-noise power spectrum is then modulated in harmonic space using the product of the two smooth fitting functions given in Eqs. (48) and (49).

*Correlated noise between detectors.* If there is some correlation between the noise in the different cuts in our data, the trick of only forming effective frequency-pair power spectra from cross-spectra to avoid the noise biases fails. In 2013, we evaluated the amplitude of such correlated noise between different detsets. The correlation, if any, was found to be small, and we estimated its effect on the cosmological parameter fits to be negligible. As stated in Sect. 3.2.1, the situation is different for the 2015 data. Indeed, we now detect a small but significant correlated noise contribution between the detsets. This is the reason we change our choice of data to estimate the cross-spectra, from detsets to half-mission maps. The correlated noise appears to be much less significant in the latter.

To estimate the amount of correlated noise in the data, we measured the cross-spectra between the half-ring difference maps of all the individual detsets. The cross-spectra are then summed using the same inverse-variance weighting that we used in 2013 to form the effective frequency-pair spectra. Figure 26 shows the spectra for each frequency pair. All of these deviate significantly from zero. We build an effective correlated noise template by fitting a smoothing spline on a $\Delta\ell = 200$ sliding average of the data. Given the noise level in polarization, we did not investigate the possible contribution of correlated noise in $EE$ and $TE$.

Section 4.1.1 shows that when these correlated noise templates are used, the results of the detsets likelihood are in excellent agreement with those based on the baseline, half-mission one.

### 3.5. Covariance matrix structure

The construction of a Gaussian approximation to the likelihood function requires building covariance matrices for the pseudo-power spectra. Mathematically exact expressions exist, but they are prohibitively expensive to calculate numerically at *Planck* resolution (Wandelt et al. 2001); we thus follow the approach taken in Like13 and make use of analytical approximations

(Hansen et al. 2002; Hinshaw et al. 2003; Efstathiou 2004; Challinor & Chon 2005).

For our baseline likelihood, we calculate covariance matrices for all 45 unique detector combinations that can be formed out of the six frequency-averaged half-mission maps at 100, 143, and 217 GHz. To do so, we assume a fiducial power spectrum that includes the data variance induced by the CMB and all foreground components described in Sect. 3.3; this variance is computed assuming these components are Gaussian-distributed. The effect of this approximation regarding Galactic foregrounds is tested by means of simulations in Sect. 3.6. The fiducial model is taken from the best-fit cosmological and foreground parameters; since they only become available after a full exploration of the likelihood, we iteratively refine our initial guess. As discussed in Sect. 3.1, the data vector used in the likelihood function of Eq. (13) is constructed from frequency-averaged power spectra. Following Like13, for each polarization combination, we therefore build averaged covariance matrices for the four frequencies $\nu_1, \nu_2, \nu_3, \nu_4$,

$$\mathrm{Var}(\hat{C}_\ell^{XY\ \nu_1,\nu_2}, \hat{C}_{\ell'}^{ZW\ \nu_3,\nu_4}) = \sum_{\substack{(i,j)\in(\nu_1,\nu_2)\\(p,q)\in(\nu_3,\nu_4)}} w_\ell^{XY\ i,j} w_{\ell'}^{ZW\ p,q} \times \mathrm{Var}\left(\hat{C}_\ell^{XY\ i,j}, \hat{C}_{\ell'}^{ZW\ p,q}\right), \tag{50}$$

where $X, Y, Z, W \in \{T, E\}$, and $w^{XY\ i,j}$ is the inverse-variance weight for the combination $(i, j)$, computed from

$$w_\ell^{XY\ i,j} \propto 1/\mathrm{Var}\left(\hat{C}_\ell^{XY\ i,j}, \hat{C}_\ell^{XY\ i,j}\right), \tag{51}$$

and normalized to unity. For the averaged $XY = TE$ covariance (and likewise for $ZW = TE$), the sum in Eq. (50) must be taken over the additional permutation $XY = ET$. That is, the two cases where the temperature map of channel $i$ is correlated with the polarization map of channel $j$ and vice versa are combined into a single frequency-averaged covariance matrix. These matrices are then combined to form the full covariance used in the likelihood,

$$\mathsf{C} = \begin{pmatrix} C^{TTTT} & C^{TTEE} & C^{TTTE} \\ C^{EETT} & C^{EEEE} & C^{EETE} \\ C^{TETT} & C^{TEEE} & C^{TETE} \end{pmatrix}, \tag{52}$$

where the individual polarization blocks are constructed from the frequency-averaged covariance matrices of Eq. (50) (Like13).

Appendix C.1.1 provides a summary of the equations used to compute temperature and polarization covariance matrices and presents a validation of the implementation through direct simulations. Let us note that, for the approximations used in the analytical computation of the covariance matrix to be precise, the mask power spectra have to decrease quickly with multipole moment $\ell$; this requirement gives rise to the apodization scheme discussed in Sect. 3.2.2. In the presence of a point-source mask, however, the condition may no longer be fulfilled, reducing the accuracy of the approximations assumed in the calculation of the covariance matrices. We discuss in Appendix C.1.4 the heuristic correction we developed to restore the accuracy, which is based on direct simulations of the effect.

### 3.6. FFP8 simulations

In order to validate the overall implementation and our approximations, we generated 300 simulated HFI half-mission map sets in the frequency range 100 to 217 GHz, which we analysed like the real data. For the CMB, we created realizations of the ΛCDM model with the best-fit parameters obtained in this paper. After convolving the CMB maps with beam and pixel window functions, we superimposed CIB, dust, and noise realizations from the FFP8 simulations (Planck Collaboration XII 2016) that capture both the correlation structure and anisotropy of foregrounds and noise. We then computed power spectra using the set of frequency-dependent masks described in Sect. 3.2.2 and created the corresponding `Plik` $TT$ likelihood. We modified the shape of the foreground spectra to fit the FFP8 simulations, but kept the parameterization used on the data. In the case of dust, we used priors similar to those used on data. Furthermore, in the following the dust amplitude parameter is named $\mathrm{gal}_{545}^{\nu\times\nu'}$. We then ran an MCMC sampler to derive the cosmological and foreground parameters posterior distributions for all dataset realizations.

**Table 13.** Shifts of parameters over 300 $TT$ simulations.

| Parameter | 300 sims | $r_\mathrm{A}^{30}$ | $r_\mathrm{A}^{65}$ | $r_\mathrm{A}^{100}$ |
|---|---|---|---|---|
| $\Omega_\mathrm{b}h^2$ . . . . . . . . | 0.27 | 0.62 | 0.50 | 0.52 |
| $\Omega_\mathrm{c}h^2$ . . . . . . . . | −0.71 | −0.65 | −0.44 | 0.00 |
| $\theta$ . . . . . . . . . . . | 1.48 | 1.67 | 1.67 | 1.29 |
| $\tau$ . . . . . . . . . . . | −0.57 | −0.38 | −0.56 | −0.38 |
| $\ln\left(10^{10}A_\mathrm{s}\right)$ . . . . | −0.70 | −0.52 | −0.65 | −0.35 |
| $n_\mathrm{s}$ . . . . . . . . . . | 1.86 | 1.87 | 1.46 | 0.78 |
| $A_\mathrm{CIB}^{217}$ . . . . . . . . | −0.99 | −1.09 | −1.44 | −1.34 |
| $\mathrm{gal}_{545}^{100}$ . . . . . . . | 0.31 | 0.13 | −0.09 | 0.04 |
| $\mathrm{gal}_{545}^{143}$ . . . . . . . | 0.40 | −0.21 | −0.23 | −0.19 |
| $\mathrm{gal}_{545}^{143-217}$ . . . . . | −0.22 | −0.35 | 0.36 | 0.22 |
| $\mathrm{gal}_{545}^{217}$ . . . . . . . | 1.61 | 1.48 | 2.19 | 2.04 |

**Notes.** Shifts are given in units of the posterior width rescaled by $1/\sqrt{300}$. If the parameters were uncorrelated, 68% of the shifts would be expected to lie within $\pm1\sigma$. The effect of varying the value of $\ell_\mathrm{min}$ is measured on the likelihood of the average spectra over 300 realizations, labelled $r_\mathrm{A}^{\ell_\mathrm{min}}$. A significant decrease of the bias on $n_\mathrm{s}$ is obtained by not including low-$\ell$ multipoles, at the cost, however, of a degradation in the determination of the foreground amplitudes $A_\mathrm{CIB}^{217}$, $\mathrm{gal}_{545}^{143-217}$, and $\mathrm{gal}_{545}^{217}$.

For each simulation, we computed the shift of the derived posterior mean parameters with respect to the input cosmology, normalized by their posterior widths $\sigma_\mathrm{post}$. When a Gaussian prior with standard deviation $\sigma_\mathrm{prior}$ is used, we rescale $\sigma_\mathrm{post}$ by $[1 - \sigma_\mathrm{post}^2/\sigma_\mathrm{prior}^2]^{1/2}$; this is the case for $\tau$ and for the Galactic dust amplitudes $\mathrm{gal}_{545}^{\nu}$ in the four cross-frequency channels used. In Fig. 27, we show histograms of the shifts we found for all 300 simulations for the six baseline cosmological parameters, as well as the FFP8 CIB and galactic dust amplitudes. As shown in the figure, we recover the input parameters with little bias and a scatter of the normalized parameter shifts around unity. The $p$-values of the Kolmogorov-Smirnov test that we ran are given in the legend and we do not detect significant departures from normality. The average reduced $\chi^2$ for the histograms of Fig. 27 is equal to 1.02.

Table 13 (second column) compiles the average shifts of Fig. 27, but in order to gauge whether they are as small as expected for this number of simulations (assuming no bias), the shifts are expressed in units of the posterior width rescaled by $1/\sqrt{300}$. We note that the shift of the average is above one (scaled)$\sigma$ in three cases out of a total of 11 parameters (68% of the Δs would be expected to lie within $1\sigma$ if the parameters

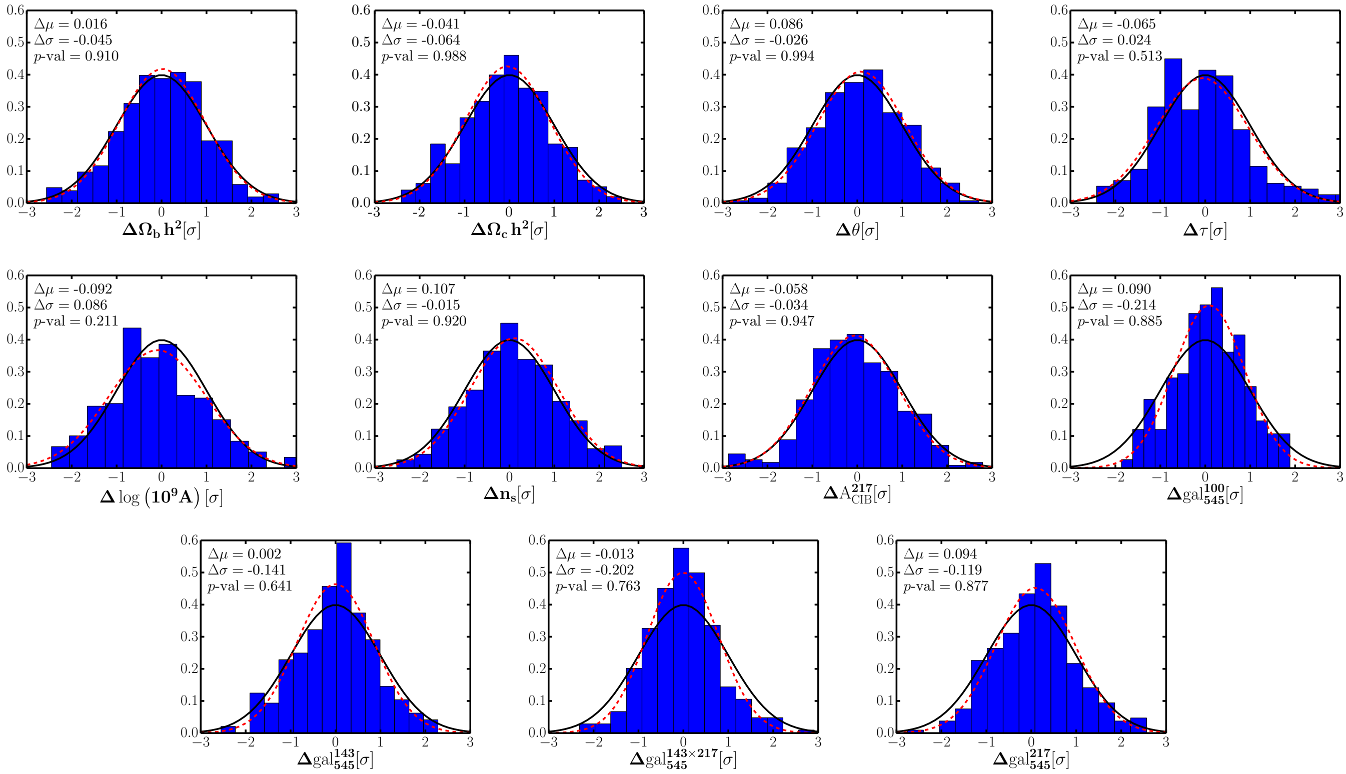


**Fig. 27.** `Plik` parameter results on 300 simulations for the six baseline cosmological parameters, as well as the FFP8 CIB and Galactic dust amplitudes. The simulations include quite realistic CMB, noise, and foregrounds (see text). The distributions of inferred posterior mean parameters are centred around their input values with the expected scatter. Indeed the dotted red lines show the best-fit Gaussian for each distribution, with a mean shift, $\Delta\mu$, and a departure $\Delta\sigma$ from unit standard deviation given in the legend; both are close to zero. These best fits are thus very close to Gaussian distributions with zero shift and unit variance, which are displayed for reference as black lines. The legend gives the numerical value of $\Delta\mu$ and $\Delta\sigma$, as well as the $p$-values of a Kolmogorov-Smirnov test of the histograms against a Gaussian distribution shifted from zero by $\Delta\mu$ and with standard deviation shifted from unity by $\Delta\sigma$. This confirms that the distributions are consistent with Gaussian distributions with zero mean and unit standard deviation, with a small offset of the mean.

were uncorrelated), with $\theta$, $n_{\rm s}$, and ${\rm gal}^{217}_{545}$ at the 1.7, 2.0, and 1.5 (scaled)$\sigma$ level, respectively.

Before proceeding, let us note that an estimate (third column) of these shifts is obtained by simply computing the shift from a single likelihood using as input the average spectra of the 300 simulations. This effectively reduces cosmic variance and noise amplitude by a factor $\sqrt{300}$ and, more importantly, it decreases the cost and length of the overall computation, enabling additional tests. These shift estimates are noted $r_{\rm A}$. The table shows that significant improvement in the determination of $n_{\rm s}$ is obtained by removing low-$\ell$ multipoles. Indeed, Cols. 4 and 5 of Table 13 show the variation of the shift when the $\ell_{\rm min}$ of the high-$\ell$ likelihood is increased from 30 to 65 and 100. The shift in $n_{\rm s}$ is decreased by a factor two, while the decrease in the number of bins per cross-frequency spectrum is only reduced from 199 to 185 (having little impact on the size of the covariance matrix of cosmological parameters).

These changes with $\ell_{\rm min}$ therefore trace the small biases back to the lowest-$\ell$ bins. It suggests that the Gaussian approximation used in the high-$\ell$ likelihood starts to become mildly inaccurate at $\ell = 30$. Indeed, even if noticeable, this effect would contribute at most a $0.11\sigma$ bias on $n_{\rm s}$. This is further confirmed by the lack of a detectable effect found in Sect. 5.1 when varying the hybridization scale in $TT$ between `Commander` and `Plik`. However, the exclusion of low-$\ell$ information degrades our ability to accurately reconstruct the foreground amplitudes $A^{217}_{\rm CIB}$, ${\rm gal}^{143-217}_{545}$, and ${\rm gal}^{217}_{545}$. Indeed, the dust spectral amplitudes in the $143\times217$ and $217\times217$ channels are highest at low multipoles, and the CIB spectrum in the range $30 \le \ell \le 100$ also adds substantial information.

In spite of this low-$\ell$ trade-off between an accurate determination of $n_{\rm s}$ on the one hand and $A^{217}_{\rm CIB}$, ${\rm gal}^{143-217}_{545}$, and ${\rm gal}^{217}_{545}$ on the other, we can conclude that the `Plik` implementation is behaving as expected and can be used for actual data analysis.

Appendix C.2 extends this conclusion to the joint `PlikTT`, EE, TE likelihood case.

### 3.7. End-to-end simulations

While the previous section validated our methodology, our approximations, and the overall implementation, this does not yet give the sensitivity to residual systematic uncertainties undetected by data consistency checks. These are by their very nature very much more difficult to address realistically, since, when an effect is detected and sufficiently well understood, it can be modelled and is corrected for, in general at the TOI-processing stage; only the uncertainty of the correction needs to be addressed. Still, HFI has developed a complete model of the instrument which contains all identified systematic effects and enables realistic simulation of the instrumental response. We have therefore generated a number of full-mission time streams which we have then processed with the DPC TOI processing pipeline in order to create map datasets as close to instrumental reality as we can in order to assess the possible impact of low-level residual

**Table 14.** End-to-end parameter shifts for a single realization of CMB and foregrounds, along with five different noise realizations. Shifts are computed with respect to those obtained without noise and with instrumental effects turned off.

| Parameter | 1 | 2 | 3 | 4 | 5 | Mean | Median | $\sigma_{\rm FFP8}$ |
|---|---|---|---|---|---|---|---|---|
| $\Omega_{\rm b}h^2$ . . . . . . . . | 0.32 | 0.09 | −0.3 | −0.02 | −0.17 | −0.01 | −0.02 | 0.42 |
| $\Omega_{\rm c}h^2$ . . . . . . . . | −0.07 | −0.20 | −0.22 | −0.22 | 0.30 | −0.08 | −0.20 | 0.35 |
| $\theta$ . . . . . . . . . . . | 0.18 | 0.08 | 0.24 | −0.12 | 0.15 | 0.10 | 0.14 | 0.45 |
| $\tau$ . . . . . . . . . . . | 0.08 | 0.19 | −0.02 | 0.006 | −0.21 | 0.01 | 0.005 | 0.25 |
| $\ln\left(10^{10}A_{\rm s}\right)$ . . . . | 0.04 | 0.10 | −0.11 | −0.11 | −0.21 | −0.05 | −0.11 | 0.25 |
| $n_{\rm s}$ . . . . . . . . . . | 0.11 | 0.06 | 0.22 | −0.13 | −0.64 | −0.07 | 0.05 | 0.40 |

**Notes.** In columns labelled 1 to 5: shifts of the cosmological parameters of the five noise realizations of the end-to-end simulations with respect to those obtained for the simulations without noise and with instrumental effects turned off, normalized by the end-to-end simulations' posterior widths. In columns labelled "Mean" and "Median": the corresponding mean and median. In column labelled "$\sigma_{\rm FFP8}$": the standard deviation of the distribution obtained from the cosmological parameter shifts of 100 FFP8 simulations, varying the noise only with respect to the cosmological parameters of the CMB only.

**Table 15.** End-to-end parameter shifts for four different CMB realizations but comprising four pairs of realizations with the same noise realization with respect to those obtained without noise and with instrumental effects turned off.

| Parameter | 4 | 5 | 6 | 7 | 8 | $\Delta_{5-6}$ | $\Delta_{4-7}$ | $\Delta_{4-8}$ | $\Delta_{7-8}$ | $\sigma_{\rm FFP8}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $\Omega_{\rm b}h^2$ . . . . . . . . | −0.02 | −0.17 | −0.25 | −0.53 | 0.29 | 0.08 | 0.51 | 0.31 | 0.83 | 0.34 |
| $\Omega_{\rm c}h^2$ . . . . . . . . | −0.22 | 0.30 | −0.20 | 0.52 | −0.59 | 0.50 | 0.74 | 0.37 | 1.11 | 0.30 |
| $\theta$ . . . . . . . . . . . | −0.12 | 0.15 | −0.34 | −1.06 | 0.41 | 0.48 | 0.94 | 0.52 | 1.47 | 0.35 |
| $\tau$ . . . . . . . . . . . | 0.006 | −0.21 | −0.16 | −0.17 | −0.16 | 0.04 | 0.18 | 0.16 | 0.01 | 0.175 |
| $\ln\left(10^{10}A_{\rm s}\right)$ . . . . | −0.11 | −0.21 | −0.29 | −0.15 | −0.34 | 0.07 | 0.04 | 0.24 | 0.20 | 0.19 |
| $n_{\rm s}$ . . . . . . . . . . | −0.13 | −0.64 | −0.39 | −0.86 | 0.24 | 0.24 | 0.72 | 0.37 | 1.03 | 0.35 |

**Notes.** In columns labelled 4 to 8: shifts of the cosmological parameters of the end-to-end simulations with respect to those obtained for the simulations without noise and with instrumental effects turned off, normalized by the end-to-end simulations' posterior widths. We point out that realizations numbered 4 and 5 are common to the sets of Tables 14 and 15. In columns labelled "$\Delta_{5-6}$" to "$\Delta_{7-8}$", the absolute differences in the shifts within pairs of realizations having different CMB but the same foregrounds and noise realizations. In column labelled "$\sigma_{\rm FFP8}$": the standard deviation of the distribution obtained from the cosmological parameter shifts of 100 FFP8 simulations, varying the CMB only but keeping the same FFP8 noise realization, with respect to the cosmological parameters of the CMB only.

instrumental systematics, the effects of which might have remained undetected otherwise.

In this section, we report on the shifts in the values of the cosmological and foreground parameters induced by these specific residual systematic effects, comparing the results of a $TT$ likelihood analysis for two overlapping sets of five simulations:

1. five simulations of maps at 100 GHz, 143 GHz and 217 GHz, for a single realization of the CMB and of the foregrounds but for five different realizations of the noise,
2. five simulations of maps at 100 GHz, 143 GHz and 217 GHz, composed of four CMB realizations, two noise realizations, a single realization of the foregrounds, but forming four pairs of realizations having the same noise but different CMB.

These simulations sum to a total of eight distinct simulations and are numbered from 1 to 8 in Tables 14 and 15. To be more explicit, in the second set, among simulations numbered 4 to 8, simulations 4, 7 and 8 have different CMB, but the same noise as each other. Simulations 5 and 6 have different CMB, and the same noise as each other, but different from simulations 4, 7 and 8. Realizations 4 and 5, having the same CMB but different noise, are common to the two sets of five realizations.

Each of these have been performed twice, with the end-to-end (instrument plus TOI processing) pipeline and noise contribution switched either on or off. End-to-end simulations are computationally very costly, (typically a week for each simulated mission dataset) and hence only a few realizations were generated).

As explained in Sect. 5.4 of Planck Collaboration VII (2016), the end-to-end simulations are created by feeding the TOI processing pipeline with simulated data to evaluate and characterize the overall transfer function and the respective contribution of each individual effect on the determination of the cosmological parameters. Simulated TOIs are produced by applying the real mission scanning strategy to a realistic input sky specified by the *Planck* Sky Model (PSM; Delabrouille et al. 2013) containing a lensed CMB realization, galactic diffuse foregrounds, and the dipole components. To this sky-scanned TOI, we add a white-noise component, representing the phonon and photon noises. The very low-temporal-frequency thermal drift seen in the real data is also added to the TOI. The noisy sky TOI is then convolved with the appropriate bolometer transfer functions. Another white-noise component, representing Johnson noise and read-out noise, is also added. Simulated cosmic rays using the measured glitch rates, amplitudes, and shapes are added to the TOI. This TOI is interpolated to the electronic HFI fast-sampling frequency. It is then converted from analogue to digital using a simulated non-linear analogue-to-digital converter (ADC). Identified 4 K cooler spectral lines are added to the TOI. Both effects (ADC and 4 K lines) are derived from the measured in-flight behaviour. The TOI finally goes through the data compression/decompression algorithm used for communication between the *Planck* satellite and Earth. The simulated TOI is then processed in the same way as the real mission data for cleaning and systematic error removal, calibration, destriping, and map-making.

Some limitations of the current end-to-end approach follow. No pointing error is included, although previous (dedicated) simulations suggest that this has negligible effect. In addition, this effect was included in the dedicated simulations performed to assess the precision of the beam recovery procedure. The first step of the TOI processing is to correct the ADC non-linearity (ADC NL). For the flight data, the ADC NL was determined by using HFI's measured signal at the end of the HFI mission, with the instrument's cooling system switched off and an instrument temperature equal to 4 K. This determination relied on supposing the signal to be perfect white noise and therefore to correspond to the distortions brought in by the ADC. In the current implementation, we assume perfect knowledge of ADC NL and 4 K lines. This is of course not true for the real data and future end-to-simulations, accompanying *Planck*'s next data release, will improve our model of this effect. After ADC NL correction, the signal is converted to volts. Deglitching is then performed by flagging glitch heads and by using glitch tail time-lines. This enables the creation of the thermal baseline which is used for signal demodulation. The thermal baseline and glitch tails are subtracted, the signal is converted to watts, and the 4 K lines are removed. The resulting signal is then deconvolved by the bolometer transfer functions. We do not include uncertainties in the glitch tail shape used in the deglitching procedure, i.e., the templates are the same for the simulations and the processing; but here again, previous studies suggest any difference is a small effect.

The analysis of these sets of end-to-end simulations, and of their counterparts for which all instrumental effects are turned off, is performed similarly to that of the simulations described in Sect. 3.6. Angular power spectra for all cross-half-missions and for all frequency combinations are computed using the *Planck* masks described in Sect. 3.2.2 and with the appropriate beam functions. Noise levels are evaluated as described in Sect. 3.4.4. Templates for galactic foregrounds (CO, free-free, synchrotron, thermal and spinning dust), the kinetic and thermal SZ effects, the cosmic infrared background, and radio and IR point sources are constructed based on the PSM input foreground maps. The covariance matrix is computed with the method outlined in Appendix C.1 with the aforementioned input CMB power spectrum, input foreground spectra, noise levels, beam functions and masks.

All sets of power spectra and the inverse covariance matrix are then binned and used in the likelihood analysis performed using an MCMC sampler together with `Plik` and PICO in order to determine the best fit cosmological parameters. The shifts in cosmological parameter values induced by the imperfect correction of instrumental effets by the TOI processing pipeline are then computed for the end-to-end simulations with respect to those obtained for the simulations without noise and with instrumental effects turned off, normalized by the end-to-end simulations' posterior widths. Comparing shifts computed in this way cancels out cosmic variance and chance correlations between the CMB and the foregrounds and are thus fully attributable to the instrument and to the noise, which cannot be disentangled, as well as to CMB-noise chance correlations. That is, those shifts probe directly the scatter and possible biases induced by residual systematics effects.

The mean and median shifts for the five simulations with a single CMB realization and a single foreground realization, but different noise realizations, are given in Table 14. In order to verify that these shifts are within expectations, we computed the shifts in cosmological parameters for 100 FFP8 simulations, each with identical CMB signal but different FFP8 noise, with respect to the cosmological parameters obtained for the CMB only, normalized by their posterior widths. The standard deviations of the resulting distributions are given in the column labelled "$\sigma_{\rm FFP8}$" of Table 14 and can be compared with the shifts obtained for the five end-to-end simulations. All shifts are within $1\sigma$ of the shifts expected from FFP8. In addition, there is no indication of any detectable bias. All shifts are thus compatible with scatter introduced by noise.

The shifts for the five realizations with four different CMB realizations, the same foregrounds, but comprising four pairs with the same noise realization, are given in Cols. 4 to 8 of Table 15. As mentioned at the beginning of this section, realizations numbered 4 and 5 are common to the sets of Tables 14 and 15. In the columns labelled "$\Delta_{5-6}$" to "$\Delta_{7-8}$", we computed the absolute differences in the shifts within pairs of realizations having different CMB but the same foreground and noise realizations. We compare these differences to the standard deviations of the distributions of cosmological parameter shifts of 100 FFP8 simulations, varying the CMB but keeping the same FFP8 noise realization, with respect to the cosmological parameters of the corresponding CMB but without noise (column labelled "$\sigma_{\rm FFP8}$"). These distributions quantify the impact of CMB-noise correlations on the determination of the cosmological parameters. The Table shows that among all $\Delta$'s, 11 are within $1\sigma_{\rm FFP8}$, 7 are within 1 to $2\sigma_{\rm FFP8}$, 3 are within 2 to $3\sigma_{\rm FFP8}$, 2 are within 3 to $4\sigma_{\rm FFP8}$. 50% of the differences are within $1\sigma$ and 78% within $2\sigma$. At the very worst, taking the example of $\Delta_{7-8}$ a cosmological parameter ($\theta$) moves a total of $4\sigma$, from $-3\sigma$ to $1\sigma$ in units of $\sigma_{\rm FFP8}$ when the CMB is changed but the noise is left the same. This is rare but can be expected in a few percent of simulations. As in the case of the shifts listed in Table 14, there is thus no detectable bias, with all shifts compatible with those expected from FFP8.

In summary, we have detected no sign as yet of systematic biases of the cosmological parameters due to known low-level instrumental effects as corrected by the current HFI TOI processing pipeline. An increase in the significance of these tests is left for further work once the simulation chain is further optimized for more massive numerical work.

### 3.8. High-multipole reference results

This section describes the results obtained using the baseline `Plik` likelihood, in combination with a prior on the optical depth to reionization, $\tau = 0.07 \pm 0.02$ (referred to, in $TT$, as `PlikTT`+tauprior). The robustness and validation of these results (presented in Sect. 4) can therefore be assessed independently of any potential low-$\ell$ anomaly, or hybridization issues. The full low-$\ell$ + high-$\ell$ likelihood will be discussed in Sect. 5.

Figure 28 shows the high-$\ell$ co-added CMB spectra in $TT$, $TE$, and $EE$, and their residuals with respect to the best-fit $\Lambda$CDM model in $TT$ (red line), both $\ell$-by-$\ell$ (grey points) and binned (blue circles). The blue error bars per bin are derived from the diagonal of the covariance matrix computed with the best-fit CMB as fiducial model. The bottom sub-panels with residuals also show (yellow lines) the diagonal of the $\ell$-by-$\ell$ covariance matrix, which may be compared to the dispersion of the individual $\ell$ determinations. Parenthetically, it provides graphical evidence that $TT$ is dominated by cosmic variance through $\ell \approx 1600$, while $TE$ is cosmic-variance dominated at $\ell \lesssim 160$ and $\ell \approx 260$–$460$. The jumps in the polarization diagonal-covariance

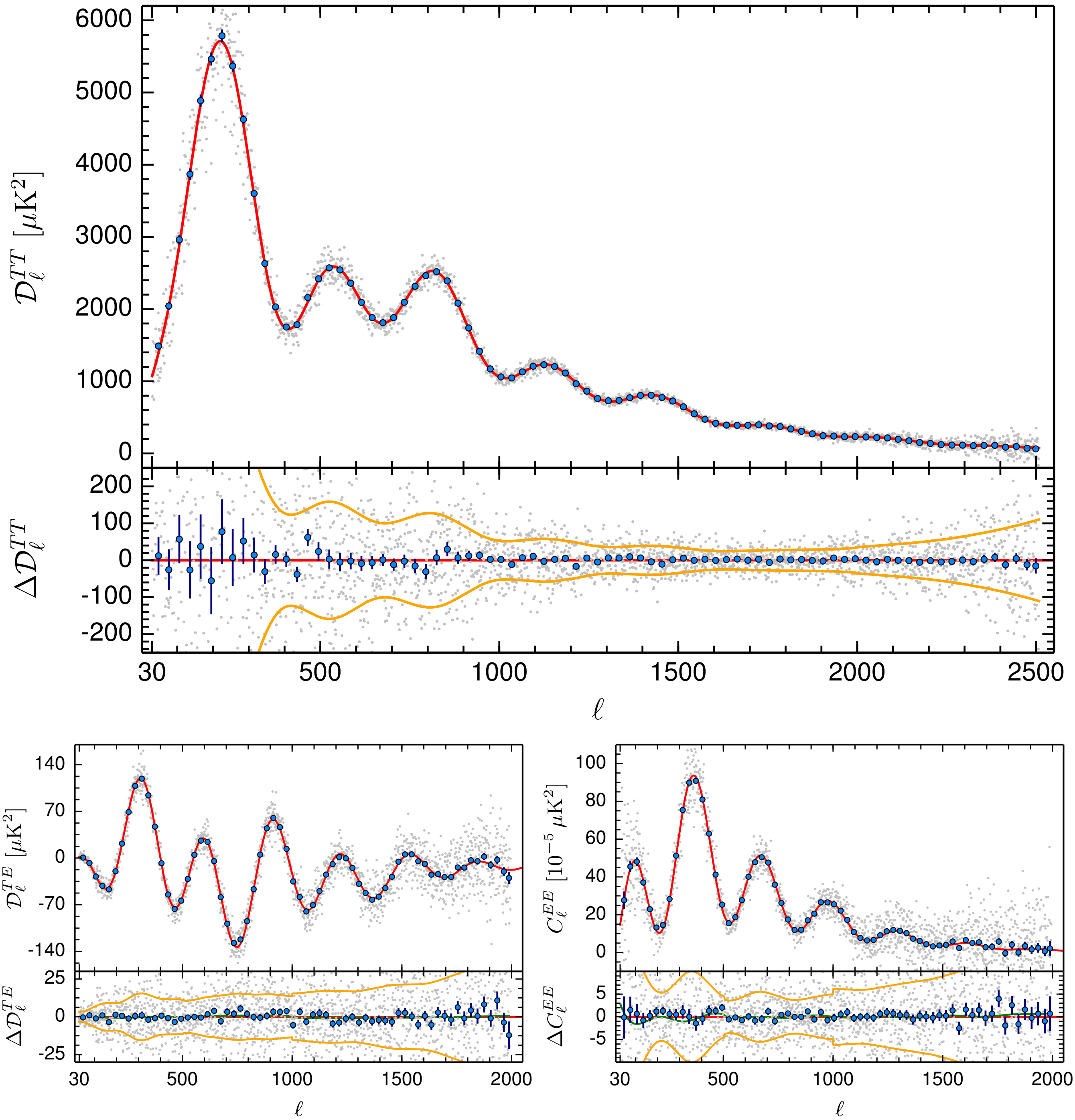


**Fig. 28.** `Plik` 2015 co-added $TT$, $TE$, and $EE$ spectra. The blue points are for bins of $\Delta\ell = 30$, while the grey points are unbinned. The *lower panels* show the residuals with respect to the best fit `PlikTT`+tauprior $\Lambda$CDM model. The yellow lines show the 68% unbinned error bars. For $TE$ and $EE$, we also show the best-fit beam-leakage correction (green line; see text and Fig. 23).

error-bars come from the variable $\ell$ ranges retained at different frequencies, which therefore vary the amount of data included discontinuously with $\ell$. Figure 29 zooms in to five adjacent $\ell$ ranges on the co-added spectra to allow close inspection of the data distribution around the model.

More quantitatively, Table 16 shows the $\chi^2$ values with respect to the $\Lambda$CDM best fit to the `PlikTT`+tauprior data combination for the unbinned CMB co-added power spectra (obtained as described in Appendix C.4). The $TT$ spectrum has a reduced $\chi^2$ of 1.03 for 2479 degrees of freedom, corresponding to a probability to exceed (PTE) of 17.2%; the base $\Lambda$CDM model is therefore in agreement with the co-added data. The best-fit $\Lambda$CDM model in $TT$ also provides an excellent description of the co-added polarized spectra, with a PTE of 12.8% in $TE$ and 34.6% in $EE$. This already suggests that extensions with, e.g., isocurvature modes can be severely constrained.

Despite this overall agreement, we note that the PTEs are not uniformly good for all cross-frequency spectra (see in particular the $100\times100$ and $100\times217$ in $TE$). This shows that the baseline instrumental model needs to include further effects to describe all of the data in detail, even if the averages over frequencies appear less affected. The green line in Fig. 28 (mostly visible in

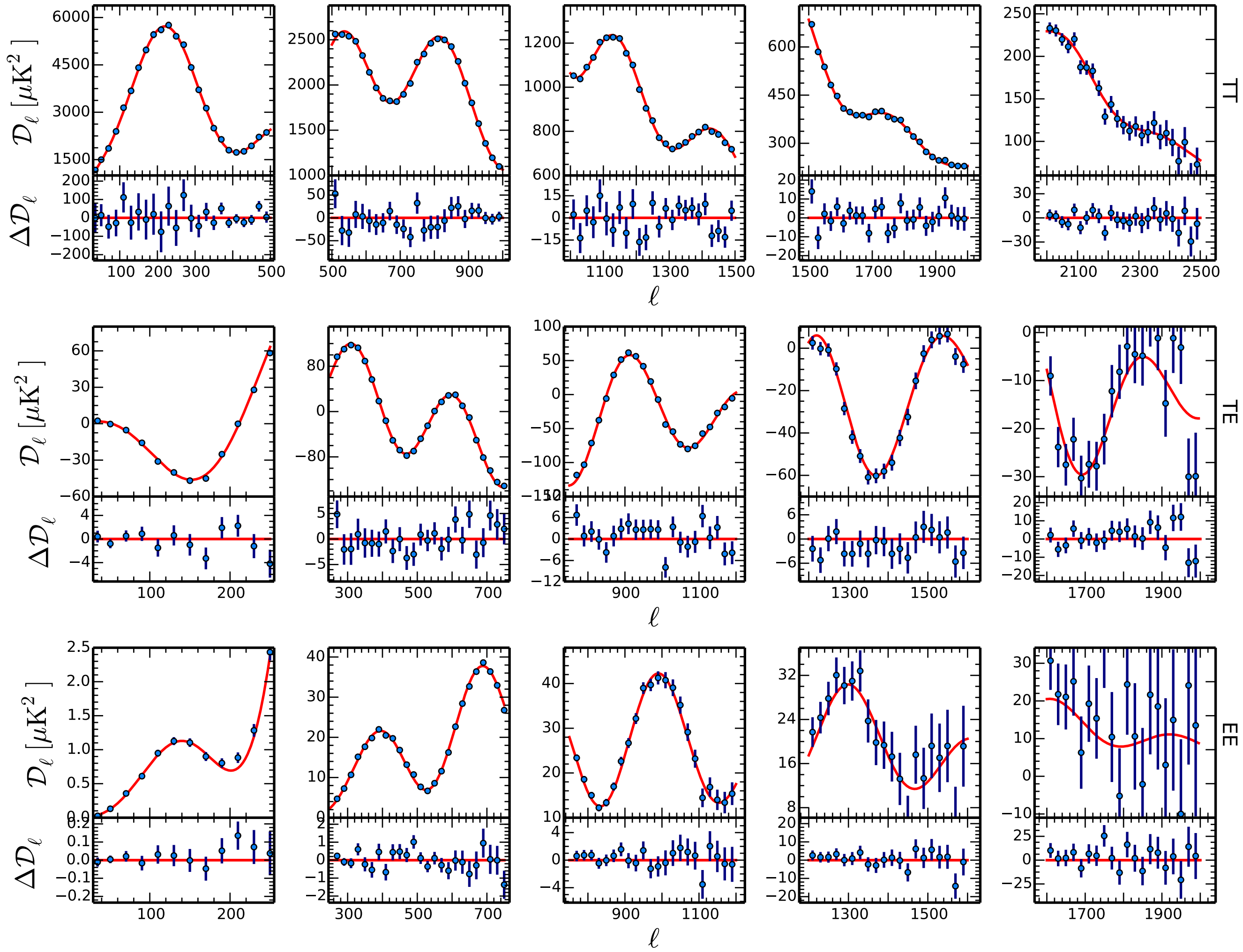

**Fig. 29.** Zoom in to various $\ell$ ranges of the HM co-added power spectra, together with the `Plik`TT+tauprior $\Lambda$CDM best-fit model (red line). We show the $TT$ (*top*), $TE$ (*centre*) and $EE$ (*bottom*) power spectra. The *lower panels* in each plot show the residuals with respect to the best-fit model.

the $\Delta C_\ell^{EE}$ plot) shows the best-fit leakage correction (shown on its own in Fig. 23), which is obtained when fixing the cosmology to the $TT$-based model. Let us recall, though, that this correction is for illustrative purposes only, and it is set to zero for all actual parameter searches. Indeed, we shall see that these leakage effects are not enough to bring all the data into full concordance with the model.

In more quantitative detail, Fig. 30 shows the binned ($\Delta\ell$ = 100) residuals for the co-added CMB spectra in units of the standard deviation of each data point, (data − model)/error. For $TT$, we find the greatest deviations at $\ell \approx 434$ ($-1.8\sigma$), 464 ($2.7\sigma$), 1214 ($-2.1\sigma$), and 1450 ($-1.8, \sigma$). At $\ell = 1754$, where we previously reported a deficit due to the imperfect removal of the $^4$He-JT cooler line (see Planck Collaboration XIII 2016, Sect. 3), there is now a less significant fluctuation, at the level of $-1.4\sigma$. The residuals in polarization show similar levels of discrepancy.

In order to assess whether these deviations are specific to one particular frequency channel or appear as a common signal in all the spectra, Fig. 31 shows foreground-cleaned $TT$ power spectra differences across all frequencies, in units of standard deviations (details on how this is derived can be found in Appendix C.3.2). The agreement between $TT$ spectra is clearly quite good. Figure 32 then shows the residuals per frequency for the $TT$ power spectra with respect to the $\Lambda$CDM `Plik`TT+tauprior best-fit model (see also the zoomed-in residual plots in Fig. C.5). The $\ell \approx 434$, 464, and 1214 deviations from the model appear to be common to all frequency channels, with differences between the frequencies smaller than $2\sigma$. However, the deviation at $\ell \approx 1450$ is higher at $217 \times 217$ than in the other channels. In particular, the inter-frequency differences (Fig. 31) between the $217 \times 217$ power spectrum and the $100 \times 100$, $143 \times 143$, and $143 \times 217$ ones show deviations at $\ell \approx 1450$ at the roughly 1.7, 2.6, and $3.4\sigma$ levels, respectively.

This inter-frequency difference is due to a deficit in the residuals of the $217 \times 217$ channel of about $-3.4\sigma$ in the bin centred at 1454 in Fig. 32. To better quantify this deviation, we also fit for a feature of the type $\cos^2((\pi/2)(\ell - \ell_p)/(\Delta\ell))$, with maximum amplitude centred at $\ell_p = 1460$, width $\Delta\ell = 25$ (we impose the feature to be zero at $|\ell - \ell_p| > \Delta\ell$) and with an independent amplitude in each frequency channel. At $217 \times 217$, we find an amplitude of $(-37.44 \pm 9.5)\mu\mathrm{K}^2$, while in the other channels we find $(-15.0 \pm 7.8)\mu\mathrm{K}^2$ at $143 \times 143$ and $(-19.7 \pm 7.9)\mu\mathrm{K}^2$ at $143 \times 217$. This outlier seems to be at least in part due to chance correlation between the CMB and dust. Indeed, the amplitude of the feature in the different spectra is in rough agreement with the dust emission law. Moreover, the feature can also be found when varying the retained sky fraction in the galactic mask, again with an amplitude scaling compatible with a dust

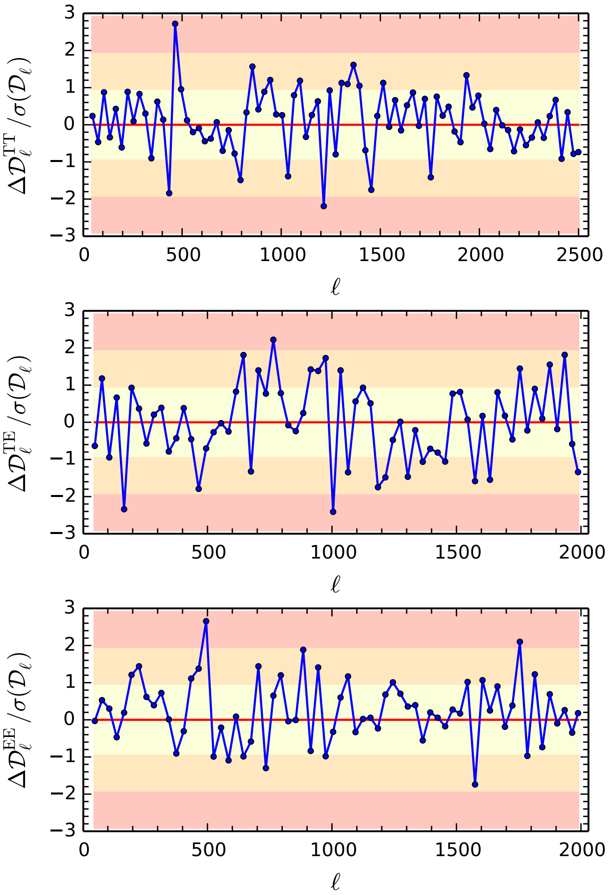


**Fig. 30.** Residuals of the co-added CMB $TT$ power spectra, with respect to the PlikTT+tauprior best-fit model, in units of standard deviation. The three coloured bands (from the centre, yellow, orange, and red) represent the $\pm1$, $\pm2$, and $\pm3\sigma$ regions.

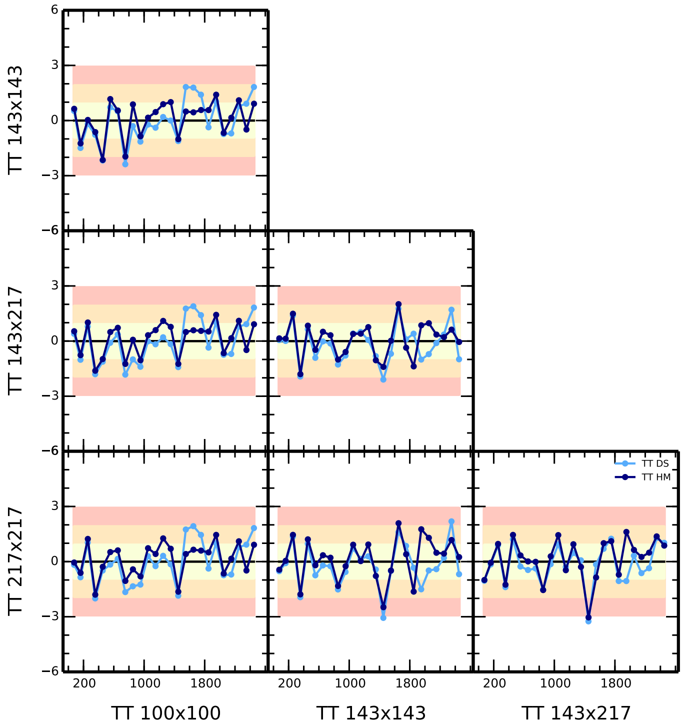


**Fig. 31.** Inter-frequency foreground-cleaned $TT$ power spectra differences, in $\mu\mathrm{K}^2$. Each of the sub-panels shows the difference, after foreground subtraction, between pairs of frequency power spectra (the spectrum named on the vertical axis minus the one named on the horizontal axis), in units of standard deviation. The coloured bands identify deviations that are smaller than one (yellow), two (orange), or three (red) standard deviations. We show the differences for both the HM power spectra (blue points) and the DS power spectra (light blue points) after correlated noise correction. Figure 41 displays the same quantities for the $TE$ and $EE$ spectra.

origin. We discuss below the impact on cosmological parameters, see the case "CUT $\ell$ = 1404–1504" in Fig. 35.

Finally we note that there is a deficit in the $\ell$ = 500–800 region (in particular between $\ell$ = 700 and 800) in the residuals of all the frequency spectra, roughly in correspondence with the position of the second and third peaks. Section 4.1 is dedicated to the study of these deviations and their impact on cosmological parameters. In spite of these marginally significant deviations from the model, the $\chi^2$ values shown in Table 16 indicate that the $\Lambda$CDM model is an acceptable fit to each of the unbinned individual frequency power spectra, with PTEs always $\mathcal{P} \gtrsim 10\%$ in $TT$. We therefore proceed to examine the parameters of the best-fit model.

The cosmological parameters of interest are summarized in Table 17. Let us note that the cosmological parameters inferred here are obtained using the same codes, priors, and assumptions as in Planck Collaboration XIII (2016), except for the fact that we use the much faster PICO (Fendt & Wandelt 2007a) code instead of CAMB when estimating cosmological parameters[12] from $TT, TE$ or $TT$, $TE$, $EE$ using high-$\ell$ *Planck* data. Appendix C.5 establishes that the results obtained with the two codes only differ by small fractions of a standard deviation (less than 15% for most parameters, with a few more extreme deviations). However, we still use the CAMB code for results from $EE$ alone, since in this case the parameter space explored is so wide that it includes regions outside the PICO interpolation region (see Appendix C.5 for further details).

Figure 33 shows the posterior distributions of each pair of parameters of the base $\Lambda$CDM model from PlikTT+tauprior. The upper-right triangle compares the $1\sigma$ and $2\sigma$ contours for the full likelihood with those derived from only the $\ell < 1000$ or the $\ell \geq 1000$ data. Section 4.1.6 addresses the question of whether the results from these different cases are consistent with what can be expected statistically. The lower-left triangle further shows that the results are not driven by the data from a specific channel, i.e., dropping any of the 100, 143, or 217 GHz map data from the analysis does not lead to much change. The next section provides a quantitative analysis of this and other jack-knife tests.

We now turn to polarization results. Inter-frequency comparisons and residuals for $TE$ and $EE$ spectra are analysed in detail in Sect. 4.4. Suffice it to say here that the results are less satisfactory than in $TT$, both in the consistency between frequency spectra and in the detailed $\chi^2$ results. This shows that the instrumental data model for polarization is less complete than for temperature, with residual effects at the $\mu\mathrm{K}^2$ level. The model thus needs to be further developed to take full advantage of the HFI data in polarization, given the level of noise achieved. We thus consider the high-$\ell$ polarized likelihood as a "beta" version. Despite these limitations, we include it in the product delivery, to allow external reproduction of the results, even though the tests that we show indicate that it should not be used when searching for weak deviations (at the $\mu\mathrm{K}^2$ level) from the baseline model.

[12] The definition of $A_{\rm L}$ differs in PICO and CAMB; see Appendix C.5.

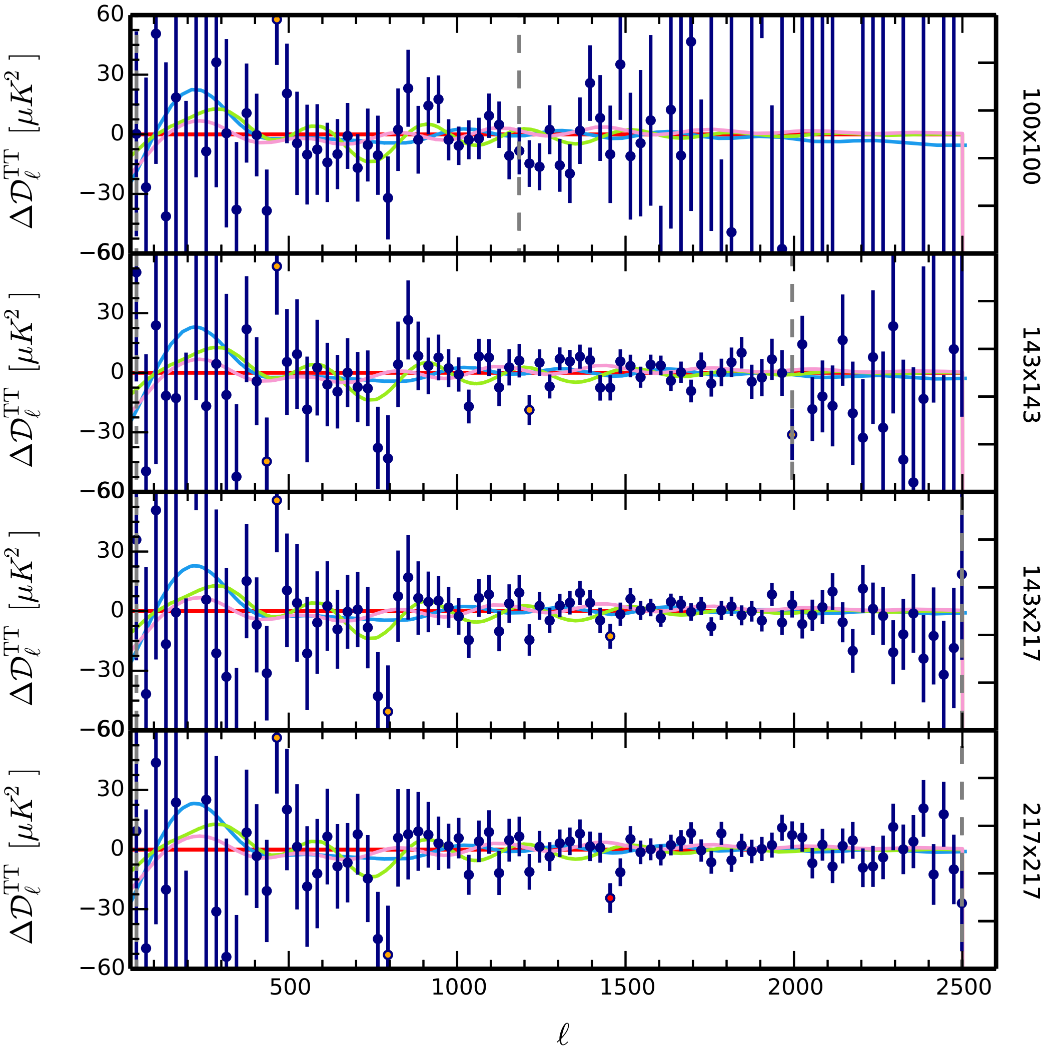


**Fig. 32.** Residuals in the half-mission $TT$ power spectra after subtracting the PlikTT+tauprior ΛCDM best-fit model (blue points, except for those which differ by at least 2 or $3\sigma$, which are coloured in orange or red, respectively). The light blue line shows the difference between the best-fit model obtained assuming a ΛCDM+$A_{\rm L}$ model and the ΛCDM best-fit baseline; the green line shows the difference of best-fit models using the $\ell_{\max} = 999$ likelihood (fixing the foregrounds to the baseline solution) minus the baseline best-fit (both in the ΛCDM framework); while the pink line is the same as the green one but for $\ell_{\max} = 1404$ instead of $\ell_{\max} = 999$; see text in Sect. 4.1. For the $TE$ and $EE$ spectra, see Fig. 40.

Nevertheless, we generally find agreement between the $TT$, $TE$, and $EE$ spectra. Figure 34 shows the $TE$, and $EE$ residual spectra conditioned on $TT$, which are close to zero. This is particularly the case for $TE$ below $\ell = 1000$, which gives some confidence in the polarization model. Most of the data points for $TE$ and $EE$ lie in the $\pm 2\sigma$ range. As for all $\chi^2$-based evaluations, the interpretation of this result depends crucially on the quality of the error estimates, i.e., on the quality of our noise model (see Sect. 3.4.4). We further note that the agreement is consistent with the finding that unmodelled instrumental effects in polarization are at the $\mu$K$^2$ level.

## 4. Assessment of the high-multipole likelihood

This section describes tests that we performed to assess the accuracy and robustness of the reference results of the high-$\ell$ likelihood that were presented above. First we establish the robustness of the $TT$ results using Plik alone in Sect. 4.1 and with other likelihoods in Sect. 4.2. We verify in Sect. 4.3 that the amplitudes of the compact-source contributions derived at various frequencies are consistent with our current knowledge of source counts. We then summarize in Sect. 4.4 the results of the detailed tests of the robustness of the polarization results, which are expanded upon in Appendix C.3.5. The paper Planck Collaboration XVI (2016) examines the dependence of the power spectrum on angular direction.

### 4.1. TT robustness tests

Figure 35 shows the marginal mean and the 68% CL error bars for cosmological parameters calculated assuming different data choices, likelihoods, parameter combinations, and data combinations. The 31 cases shown assume a base-ΛCDM framework, except when otherwise specified. The reference case uses the PlikTT+tauprior data combination. Figure 36 adds the specific results for the lensing parameter $A_{\rm L}$ (left) in a ΛCDM+$A_{\rm L}$ framework and for the effective number of relativistic species $N_{\rm eff}$ (right) in a ΛCDM+$N_{\rm eff}$ extended framework.

In both figures, the grey bands show the standard deviation of the parameter shifts relative to the baseline likelihood expected when using a sub-sample of the data (e.g., excising $\ell$-ranges

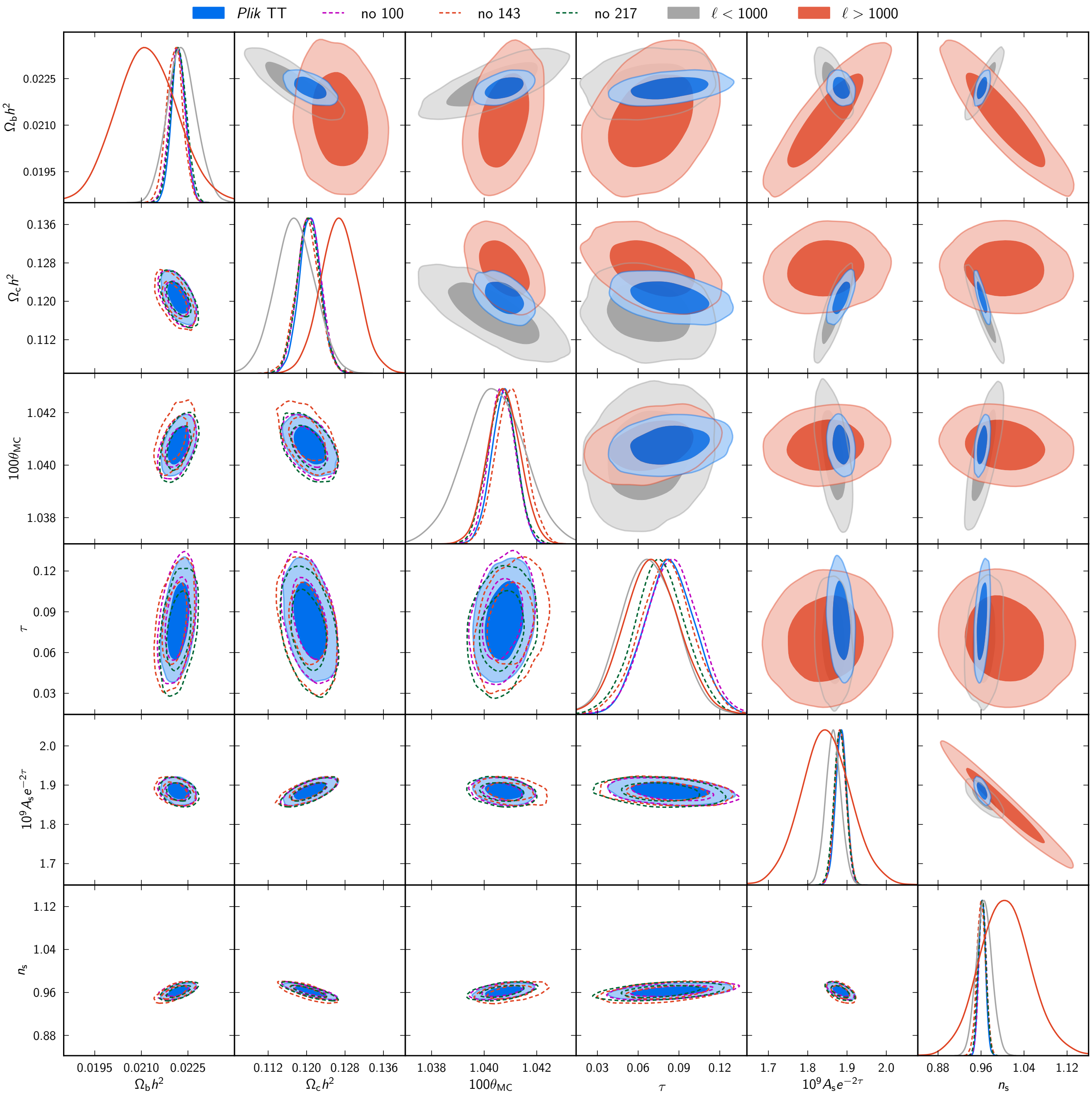


**Fig. 33.** ΛCDM parameters posterior distribution for `PlikTT+tauprior`. The lower left triangle of the matrix displays how the constraints are modified when the information from one of the frequency channels is dropped. The upper right triangle displays how the constraints are modified when the information from multipoles $\ell$ greater or less than 1000 is dropped. All the results shown in this figure were obtained using the `CAMB` code.

or frequencies). Because the data sets used to make inferences about a model are changed, one would naturally expect the inferences themselves to change, simply because of the effects of noise and cosmic variance. The inferences could also be influenced by inadequacies in the model, deficiencies in the likelihood estimate, and systematic effects in the data. Indeed, one may compare posterior distributions from different data subsets with each other and with those from the full data set, in order to assess the overall plausibility of the analysis.

To this end it is useful to have some idea about the typical variation in posteriors that one would expect to see even in the ideal case of an appropriate model being used to fit data sets with correct likelihoods and no systematic errors. It can be shown (Gratton & Challinor, in prep.) that if $Y$ is a subset of a data set $X$, and $\boldsymbol{P}_X$ and $\boldsymbol{P}_Y$ are vectors of the maximum-likelihood parameter values for the two data sets, then the sampling distribution of the differences of the parameter values is given by

$$\overline{(\boldsymbol{P}_Y - \boldsymbol{P}_X)(\boldsymbol{P}_Y - \boldsymbol{P}_X)^{\mathsf{T}}} = \mathrm{cov}(\boldsymbol{P}_Y) - \mathrm{cov}(\boldsymbol{P}_X), \quad (53)$$

i.e., the covariance of the differences is simply the difference of their covariances. Here the covariances are approximated by the

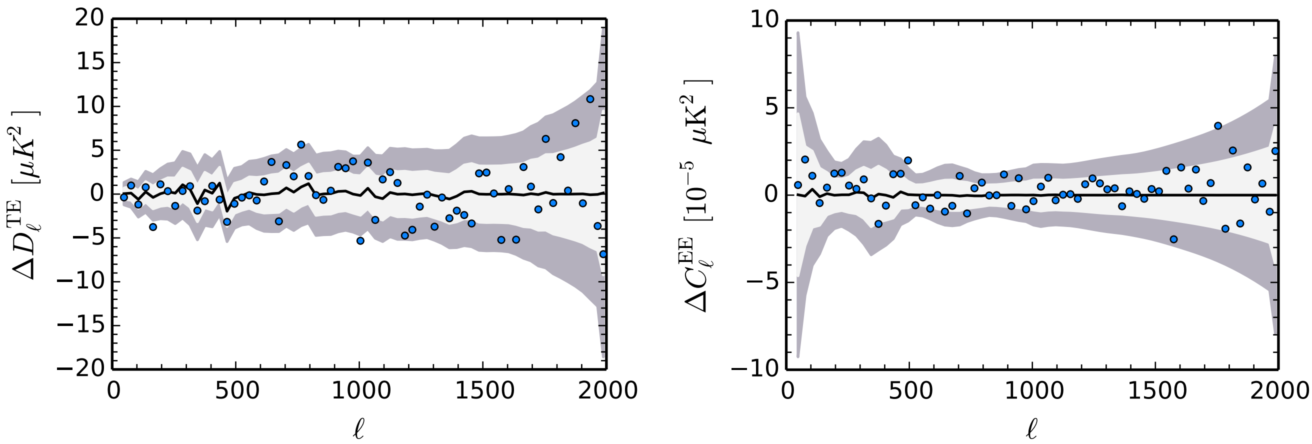


**Fig. 34.** *TE* (*left*) and *EE* (*right*) residuals conditioned on the *TT* spectrum (black line) with 1 and 2$\sigma$ error bands. The blue points are the actual *TE* and *EE* residuals. We do not include any beam-leakage correction here.

**Table 16.** Goodness-of-fit tests for the `Plik` temperature and polarization spectra at high $\ell$.

| Frequency [GHz] | $f_{sky}$[%][a] | Multipole range | $\chi^2$ | $\chi^2/N_\ell$ | $N_\ell$ | $\Delta\chi^2$ $\sqrt{2N_\ell}$[b] | PTE [%][c] | $\chi_{norm}$[d] | PTE$_\chi$[%][e] |
|---|---|---|---|---|---|---|---|---|---|
| ***TT*** | | | | | | | | | |
| 100 × 100 . . . . . . . . | 66 | 30–1197 | 1234.91 | 1.06 | 1168 | 1.38 | 8.50 | -0.30 | 76.44 |
| 143 × 143 . . . . . . . . | 57 | 30–1996 | 2034.59 | 1.03 | 1967 | 1.08 | 14.09 | -0.39 | 69.91 |
| 143 × 217 . . . . . . . . | 49 | 30–2508 | 2567.11 | 1.04 | 2479 | 1.25 | 10.63 | -1.07 | 28.25 |
| 217 × 217 . . . . . . . . | 47 | 30–2508 | 2549.40 | 1.03 | 2479 | 1.00 | 15.87 | -0.17 | 86.72 |
| Co-added . . . . . . . . . | | 30–2508 | 2545.50 | 1.03 | 2479 | 0.94 | 17.22 | -0.16 | 87.17 |
| ***TE*** | | | | | | | | | |
| 100 × 100 . . . . . . . . | 67 | 30–999 | 1089.75 | 1.12 | 970 | 2.72 | 0.43 | 3.70 | 0.02 |
| 100 × 143 . . . . . . . . | 50 | 30–999 | 1033.38 | 1.07 | 970 | 1.44 | 7.72 | 0.92 | 35.66 |
| 100 × 217 . . . . . . . . | 41 | 505–999 | 527.85 | 1.07 | 495 | 1.04 | 14.85 | 5.05 | 0.00 |
| 143 × 143 . . . . . . . . | 50 | 30–1996 | 2028.18 | 1.03 | 1967 | 0.98 | 16.45 | -2.21 | 2.69 |
| 143 × 217 . . . . . . . . | 41 | 505–1996 | 1606.06 | 1.08 | 1492 | 2.09 | 2.02 | -0.75 | 45.19 |
| 217 × 217 . . . . . . . . | 41 | 505–1996 | 1431.65 | 0.96 | 1492 | -1.10 | 86.60 | 1.33 | 18.20 |
| Co-added . . . . . . . . . | | 30–1996 | 2038.54 | 1.04 | 1967 | 1.14 | 12.76 | 0.09 | 93.09 |
| ***EE*** | | | | | | | | | |
| 100 × 100 . . . . . . . . | 70 | 30–999 | 1027.14 | 1.06 | 970 | 1.30 | 9.89 | 1.13 | 25.88 |
| 100 × 143 . . . . . . . . | 52 | 30–999 | 1048.77 | 1.08 | 970 | 1.79 | 3.94 | 1.77 | 7.72 |
| 100 × 217 . . . . . . . . | 43 | 505–999 | 479.49 | 0.97 | 495 | -0.49 | 68.33 | -3.01 | 0.26 |
| 143 × 143 . . . . . . . . | 50 | 30–1996 | 2001.48 | 1.02 | 1967 | 0.55 | 28.87 | 3.74 | 0.02 |
| 143 × 217 . . . . . . . . | 43 | 505–1996 | 1430.95 | 0.96 | 1492 | -1.12 | 86.89 | -0.71 | 47.70 |
| 217 × 217 . . . . . . . . | 41 | 505–1996 | 1409.48 | 0.94 | 1492 | -1.51 | 93.66 | -1.39 | 16.45 |
| Co-added . . . . . . . . . | | 30–1996 | 1991.37 | 1.01 | 1967 | 0.39 | 34.55 | 1.88 | 6.00 |

**Notes.** [(a)] Effective fraction of the sky retained in the analysis. For the *TE* cross-spectra between two different frequencies, we show the smaller $f_{sky}$ of the *TE* or *ET* combinations. [(b)] $\Delta\chi^2 = \chi^2 - N_\ell$ is the difference from the mean, assuming the best-fit *TT* base-$\Lambda$CDM model is correct, here expressed in units of the expected dispersion, $\sqrt{2N_\ell}$. [(c)] Probability to exceed the tabulated value of $\chi^2$. [(d)] Weighted linear sum of deviations, scaled by the standard deviation, as defined in Eq. (60). [(e)] Probability to exceed the absolute value $|\chi_{norm}|$.

inverses of the appropriate Fisher information matrices evaluated for the true model. One might thus expect the scatter in the modes of the posteriors to follow similarly, and to be able, if the parameters are well-constrained by the data, to use covariances of the appropriate posteriors on the right-hand side.

#### 4.1.1. Detset likelihood

We have verified (case "DS") that the results obtained using the half-mission cross-spectra likelihood are in agreement with those obtained using the detset (DS) cross-spectra likelihood. As explained in Sect. 3.4.4, the main difficulty in using the DS likelihood is that the results might depend on the accuracy of the correlated noise correction. Reassuringly, we find that the results from the HM and DS likelihoods agree within 0.2$\sigma$. This is an important cross-check, since we expect the two likelihoods to be sensitive to different kinds of temporal systematics. Direct differences of half-mission versus detset-based *TT* cross-frequency spectra are compared in Fig. 31 (Fig. 41 shows similar plots for the *TE* and *EE* spectra.).

When using the detsets, we fit the calibration coefficients of the various detector sets with respect to a reference. The

**Table 17.** Cosmological parameters used in this analysis.

| Parameter | Prior range | Baseline | Definition |
|---|---|---|---|
| $\omega_b \equiv \Omega_b h^2$ . . . . . . . | [0.005, 0.1] | . . . | Baryon density today |
| $\omega_c \equiv \Omega_c h^2$ . . . . . . . | [0.001, 0.99] | . . . | Cold dark matter density today |
| $\theta \equiv 100\theta_{MC}$ . . . . . | [0.5, 10.0] | . . . | $100 \times$ approximation to $r_*/D_A$ (used in CosmoMC) |
| $\tau$ . . . . . . . . . . . . . | [0.01, 0.8] | . . . | Thomson scattering optical depth due to reionization |
| $\tau$ . . . . . . . . . . . . . | $(0.07 \pm 0.02)$ | | |
| $N_{eff}$ . . . . . . . . . . . . | [0.05, 10.0] | 3.046 | Effective number of neutrino-like relativistic degrees of freedom (see text) |
| $Y_P$ . . . . . . . . . . . . . | [0.1, 0.5] | BBN | Fraction of baryonic mass in helium |
| $A_L$ . . . . . . . . . . . . . | [0.0, 10] | 1 | Amplitude of the lensing power relative to the physical value |
| $n_s$ . . . . . . . . . . . . . | [0.8, 1.2] | . . . | Scalar spectrum power-law index ($k_0 = 0.05$ Mpc$^{-1}$) |
| $\ln(10^{10}A_s)$ . . . . . . . | [2, 4.0] | . . . | Log power of the primordial curvature perturbations ($k_0 = 0.05$ Mpc$^{-1}$) |
| $\Omega_\Lambda$ . . . . . . . . . . . . | . . . | . . . | Dark energy density divided by the critical density today |
| Age . . . . . . . . . . . . | . . . | . . . | Age of the Universe today (in Gyr) |
| $\Omega_m$ . . . . . . . . . . . . | . . . | . . . | Matter density (inc. massive neutrinos) today divided by the critical density |
| $z_{re}$ . . . . . . . . . . . . . | . . . | . . . | Redshift at which Universe is half reionized |
| $H_0$ . . . . . . . . . . . . | [20, 100] | . . . | Current expansion rate in km s$^{-1}$ Mpc$^{-1}$ |
| $100\theta_D$ . . . . . . . . . . | . . . | . . . | $100 \times$ angular extent of photon diffusion at last scattering |
| $100\theta_{eq}$ . . . . . . . . . . | . . . | . . . | $100 \times$ angular size of the comoving horizon at matter-radiation equality |

**Notes.** The columns indicate the cosmological parameter symbol, their uniform prior ranges in square brackets, or between parenthesis for a Gaussian prior, the baseline values if fixed for the standard ΛCDM model, and their definition. These parameters are the same as for the previous release. The top block lists the estimated parameters, while the lower block lists derived parameters.

resulting best-fit values are very close to one[13], with the greatest calibration refinement being less than 0.2%, in line with the accuracy expected from the description of the data processing in Planck Collaboration VIII (2016). This verifies that the maps produced by the HFI DPC and used for the half-mission-based likelihood come from the aggregation of well-calibrated and consistent data.

#### 4.1.2. Impact of Galactic mask and dust modelling

We have tested the robustness of our results with respect to our model of the Galactic dust contribution in various ways.

Galactic masks. We have examined the impact of retaining a smaller fraction of the sky, less contaminated by Galactic emission. The baseline $TT$ likelihood uses the G70, G60, and G50 masks (see Appendix A) at 100, 143, and 217 GHz, respectively. We have tested the effects of using G50, G41, and G41 (corresponding to $f_{sky}^{noap} = 0.60, 0.50$, and 0.50 before apodization, case "M605050" in Fig. 35), and of the priors on the Galactic dust amplitudes relative to these masks described in Table 11. We find stable results as we vary these sky cuts, with the greatest shift in $\theta_{MC}$ of $0.5\sigma$, compatible with the expected shift of $0.57\sigma$ calculated using Eq. (53). Going to higher sky fraction is more difficult. Indeed, the improvement in the parameter determination from increasing the sky fraction at 143 GHz and 217 GHz would be modest, as we would only gain information in the small-scale regime, which is not probed by 100 GHz. Increasing the sky fraction at 100 GHz is also more difficult because our estimates have shown that adding as little as 5% of the sky closer to the Galactic plane requires a change in the dust template and more than doubles the dust contamination at 100 GHz.

[13] The fitted values are 1.0000, 0.9999, 1.0000, 1.0000, 0.9987, 0.9986, 0.9992, 0.9989, 0.9989, 0.9981, 0.9989, 1.0000, and 0.9999 for detsets 100-ds1, 100-ds2, 143-ds1, 143-ds2, 143-5, 143-6, 143-7, 217-1, 217-2, 217-3, 217-4, 217-ds1, and 217-ds2, respectively.

Amplitude priors. We have tested the impact of not using any prior (i.e., using arbitrarily wide, uniform priors) on the Galactic dust amplitudes (case "No gal. priors" in Fig. 35). Again, cosmological results are stable, with the greatest shifts in $\ln(10^{10}A_s)$ of $0.23\sigma$ and in $n_s$ of $0.20\sigma$. The values of the dust amplitude parameters, however, do change, and their best-fit values increase by about 15 $\mu$K$^2$ for all pairs of frequencies, while at the same time the error bars of the dust amplitude parameters increase very significantly. All of the amplitude levels obtained from the 545 GHz cross-correlation are within $1\sigma$ of this result. The dust levels from this experiment are clearly unphysically high, requiring 22 $\mu$K$^2$ ($\mathcal{D}_\ell, \ell = 200$) for the $100 \times 100$ pair. This level of dust contamination is clearly not allowed by the $545 \times 100$ cross-correlation, demonstrating that the prior deduced from it is informative. Nevertheless, the fact that cosmological parameters are barely modified in this test indicates that the values of the dust amplitudes are only weakly correlated with those of the cosmological parameters, consistent with the results of Figs. 44 and 45 below, which show the parameter correlations quantitatively.

Galactic dust template slope. We have allowed for a variation of the Galactic dust index $n$, defined in Eq. (24), from its default value $n = -2.63$, imposing a Gaussian prior of $-2.63 \pm 0.05$ ("GALINDEX" case in Fig. 35). We find no shift in cosmological parameters (smaller than $\sim 0.1\sigma$) and recover a value for the index of $n = -2.572 \pm 0.038$, consistent with our default choice.

Impact of $\ell \lesssim 500$ at 217 GHz. We have analysed the impact of excising the first 500 multipoles ("LMIN=505 at 217 GHz" in Fig. 35) in the $143 \times 217$ and $217 \times 217$ spectra, where the Galactic dust contamination is the strongest. We find very good stability in the cosmological parameters, with the greatest change being a $0.16\sigma$ increase in $n_s$. This is compatible with the expectations estimated from Eq. (53) of $0.14\sigma$. The inclusion of the first 500 multipoles at 217 GHz in the baseline `Plik` likelihood is one of the sources of the roughly $0.45\sigma$ difference

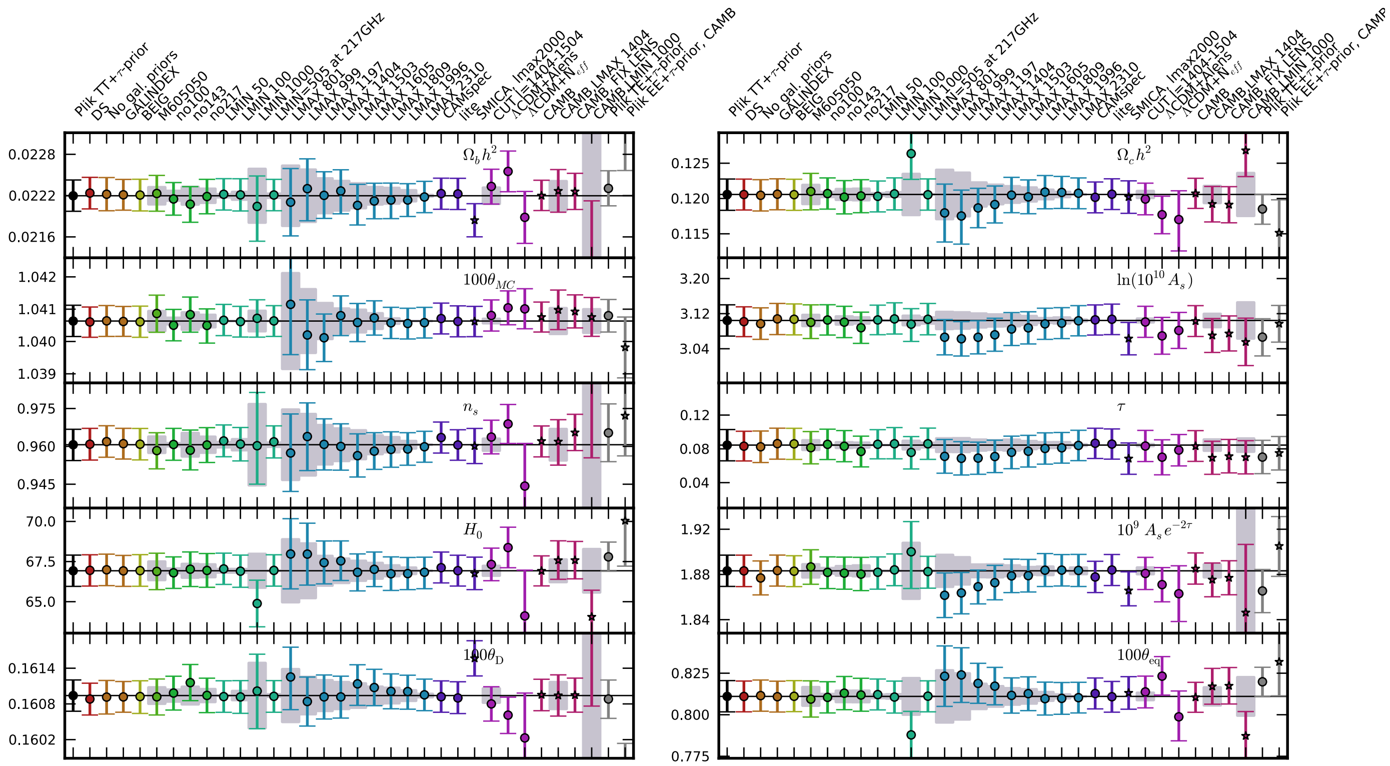

**Fig. 35.** Marginal mean and 68% CL error bars on cosmological parameters estimated with different data choices for the `Plik` likelihood, in comparison with results from alternate approaches or model. We assume a ΛCDM model and use variations of the `PlikTT` likelihood in most of the cases, in combination with a prior $\tau = 0.07 \pm 0.02$ (using neither low-$\ell$ temperature nor polarization data). The "`PlikTT`+tauprior" case (black dot and thin horizontal black line) indicates the baseline (HM, $\ell_{\rm min} = 30$, $\ell_{\rm max} = 2508$), while the other cases are described in Sect. 4.1 (and 4.2, 5.6, E.4). The grey bands show the standard deviation of the expected parameter shift, for those cases where the data used is a sub-sample of the baseline likelihood (see Eq. (53)). All the results were run with `PICO` except for few ones that were run with `CAMB`, as indicated in the labels.

in $n_{\rm s}$ observed when using the `CamSpec` code, since the latter excises that range of multipoles; for further discussion see Planck Collaboration XIII (2016, Table 1 and Sect. 3.1), as well as Sect. 4.2.

### 4.1.3. Impact of beam uncertainties

The case labelled "BEIG" in Fig. 35 corresponds to the exploration of beam eigenvalues with priors 10 times higher than indicated by the analysis of our MC simulation of beam uncertainties (which indicated by dotted lines in Fig. 21). This demonstrates that these beam uncertainties are so small in this data release that they do not contribute to the parameter posterior widths. They are therefore not enabled by default.

### 4.1.4. Inter-frequency consistency and redundancy

We have tested the effect of estimating parameters while excluding one frequency channel at a time. In Figs. 33 and 35, the "no100" case shows the effect of excluding the $100 \times 100$ frequency spectrum, the "no143" of excluding the $143 \times 143$ and $143 \times 217$ spectra, and the "no217" of excluding the $143 \times 217$ and $217 \times 217$ spectra.

We obtain the greatest deviations in the "no217" case for $\ln(10^{10} A_{\rm s})$ and $\tau$, which shift to lower values by $0.53\sigma$ and $0.47\sigma$, about twice the expected shift calculated using Eq. (53), $0.25\sigma$ and $0.23\sigma$ respectively (in units of standard deviations of the "no217" case). The value of $\Omega_{\rm c}h^2$ decreases by only $-0.1\sigma$. Figure 37 further shows the $217 \times 217$ spectrum conditioned on the $100 \times 100$ and $143 \times 143$ ones. This conditional deviates significantly in two places, at $\ell = 200$ and $\ell = 1450$. The $\ell = 1450$ case was already discussed in Sect. 3.8 and is further analysed in Sect. 4.1.6. Around $\ell = 200$, we see some excess scatter (both positive and negative) in the data around a jump between two consecutive bins of the conditional. This corresponds to the two bins around the first peak (one right before and the other almost at the location of the first peak), as can be seen in Fig. 28. All of the frequencies exhibit a similar behaviour (see Fig. 32); however, it is most pronounced in the 217 GHz case. This multipole region is also near the location of the bump in the effective dust model. The magnitude of this excess power in the model is not big enough or sharp enough to explain this excess scatter (see Fig. 17). Finally, note that the best-fit CMB solution at large scales is dominated by the $100 \times 100$ data, which are measured on a greater sky fraction (see Fig. 14).

This test shows that the parameters of the ΛCDM model do not rely on any specific frequency map, except for a weak pull of the higher resolution 217 GHz data towards higher values of both $A_{\rm s}$ and $\tau$ (but keeping $A_{\rm s}\exp(-2\tau)$ almost constant).

### 4.1.5. Changes of parameters with $\ell_{\rm min}$

We have checked the stability of the results when changing $\ell_{\rm min}$ from the baseline value of $\ell_{\rm min} = 30$ to $\ell_{\rm min} = 50$ and 100 (and $\ell_{\rm min} = 1000$, which is discussed in Sect. 4.1.6). These correspond to the cases labelled "LMIN 50" and "LMIN 100" in Fig. 35 (to be compared to the reference case "`PlikTT`+tauprior"). This check is important, since the

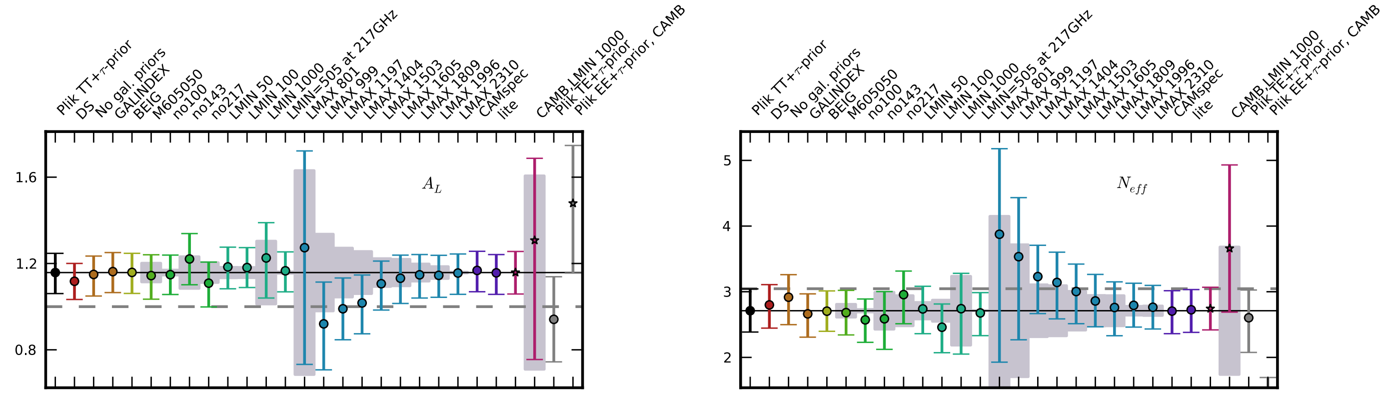


**Fig. 36.** Marginal mean and 68% CL error bars on the parameters $A_L$ (*left*) and $N_{\rm eff}$ (*right*) in ΛCDM extensions, estimated with different data choices for the PlikTT likelihood in comparison with results from alternate approaches or model, combined with a Gaussian prior on $\tau = 0.07 \pm 0.02$ (i.e., neither low-$\ell$ temperature nor polarization data). The "PlikTT+tauprior" case indicates the baseline (HM, $\ell_{\rm min} = 30$, $\ell_{\rm max} = 2508$), while the other cases are described in subsections of Sect. 4.1. The thin horizontal black line shows the baseline result and the thick dashed grey line displays the ΛCDM value ($A_L = 1$ and $N_{\rm eff} = 3.04$). The grey bands show the standard deviation of the expected parameter shift, for those cases where the data used is a sub-sample of the baseline likelihood (see Eq. (53)).

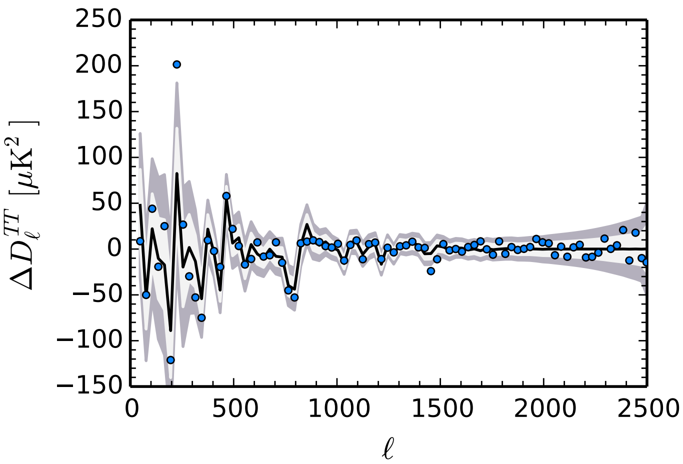


**Fig. 37.** $217 \times 217$ spectrum conditioned on the joint result from the $100 \times 100$ and $143 \times 143$ spectra. The most extreme outliers are at $\ell = 200$ and $\ell = 1450$.

Gaussian approximation assumed in the likelihood is bound to fail at very low $\ell$ (for further discussion, see Sect. 3.6).

The results are in good agreement, with shifts in parameters smaller than $0.2\sigma$, well within expectations calculated from Eq. (53). This is also confirmed in Fig. 42, where the $TT$ hybridization scale of the full likelihood is varied (i.e., the multipole where the low-$\ell$ and high-$\ell$ likelihoods are joined).

#### 4.1.6. Changes of parameters with $\ell_{\rm max}$

We have tested the stability of our results against changes in the maximum multipole $\ell_{\rm max}$ considered in the analysis. We test the restriction to $\ell_{\rm max}$ in the range $\ell_{\rm max} = 999$–$2310$, with the baseline likelihood having $\ell_{\rm max} = 2508$. For each frequency power spectrum we choose $\ell_{\rm max}^{\rm freq} = \min(\ell_{\rm max}, \ell_{\rm max}^{\rm freq,\,base})$, where $\ell_{\rm max}^{\rm freq,\,base}$ is the baseline $\ell_{\rm max}$ at each frequency as reported in Table 16. The results shown in Fig. 35 use the same settings as the baseline likelihood (in particular, we leave the same nuisance parameters free to vary) and always use a prior on $\tau$.

The results in Fig. 35 suggest there is a shift in the mean values of the parameters when using low $\ell_{\rm max}$; e.g., for $\ell_{\rm max} = 999$, $\ln(10^{10}A_{\rm s})$, $\tau$, and $\Omega_{\rm c}h^2$ are lower by 1.0, 0.8, and $0.8\sigma$ with respect to the baseline parameters. These parameters then converge to the baseline values for $\ell_{\rm max} \gtrsim 1500$. Following the arguments given earlier (Eq. (53)), when using these nested subsamples of the baseline data we expect shifts of the order of 0.5, 0.4, and $0.8\sigma$ respectively, in units of the standard deviation of the $\ell_{\rm max} = 999$ results. We further note that the value of $\theta$ for $\ell_{\rm max} \lesssim 1197$ is lower compared to the baseline value. In particular, at $\ell_{\rm max} = 1197$, its value is $0.8\sigma$ low, while the expected shift is of the order of $0.7\sigma$, in units of the standard deviation of the $\ell_{\rm max} = 1197$ results. The value of $\theta$ then rapidly converges to the baseline for $\ell_{\rm max} \gtrsim 1300$. Figure C.8 in Appendix C.3.3 also shows that these shifts are related to a change in the amplitude of the foreground parameters. In particular, the overall level of foregrounds at each frequency decreases with increasing $\ell_{\rm max}$, partially compensating for the increase in $\ln(10^{10}A_{\rm s})$ and $\Omega_{\rm c}h^2$. Although all these shifts are compatible with expectations within a factor of 2, we performed some further investigations in order to understand the origin of these changes. In the following, we provide a tentative explanation.

Table 18 shows the difference in $\chi^2$ between the best-fit model obtained using $\ell_{\rm max} = 999$ (or $\ell_{\rm max} = 1404$) and the baseline PlikTT+tauprior best-fit solution in different multipole intervals. For this test, we ran the $\ell_{\rm max}$ cases fixing the nuisance parameters to the baseline best-fit solution. This is required in order to be able to "predict" the power spectra at multipoles higher than $\ell_{\rm max}$, since otherwise the foreground parameters, which are only weakly constrained by the low-$\ell$ likelihood, can converge to unreasonable values. We note that fixing the foregrounds has an impact on cosmological parameters, which can differ from the ones shown in Fig. 35 (see Appendix C.3.4 for a direct comparison). Nevertheless, since the overall behaviour with $\ell_{\rm max}$ is similar, we use this simplified scenario to study the origin of the shifts.

The $\chi^2$ differences in Table 18 indicate that the cosmology obtained using $\ell_{\rm max} = 999$ is a better fit in the region between $\ell = 630$ and 829. In particular, the low value of $\theta$ preferred by the $\ell_{\rm max} = 999$ data set shifts the position of the third peak to smaller scales. This enables a better fit to the low points at $\ell \approx 700$–$850$ (before the third peak), followed by the high points at $\ell \approx 850$–$950$ (after the third peak). This is also clear from the residuals and the green solid line in Fig. 32, which shows

**Table 18.** Difference of $\chi^2$ values between pairs of best-fit models in different $\ell$−ranges for the co-added $TT$ power spectrum.

| Multipole range | $\Delta_{\ell_{\max}=999}$ | $\Delta_{\ell_{\max}=1404}$ | $\Delta_{A_L}$ |
|---|---|---|---|
| 30–129 . . . . . . | 0.1 | 0.31 | 0.4 |
| 130–229 . . . . . . | 0.07 | 0.05 | 0.3 |
| 230–329 . . . . . . | −0.4 | −0.22 | −0.45 |
| 330–429 . . . . . . | 0.34 | −0.09 | 0.22 |
| 430–529 . . . . . . | −0.01 | 0.17 | 0.26 |
| 530–629 . . . . . . | 0.61 | −0.26 | −0.2 |
| 630–729 . . . . . . | −1.66 | −0.8 | −0.8 |
| 730–829 . . . . . . | −1.15 | −0.13 | −0.79 |
| 830–929 . . . . . . | −0.45 | 0.01 | 0.91 |
| 930–1029 . . . . . | −0.87 | 0.41 | 0.58 |
| 1030–1129 . . . . . | 2.17 | −0.94 | −0.24 |
| 1130–1229 . . . . . | 1.65 | 1.47 | −0.17 |
| 1230–1329 . . . . . | 0.87 | 0.17 | −0.08 |
| 1330–1429 . . . . . | 6.21 | −1.46 | −0.64 |
| 1430–1529 . . . . . | −0.2 | 3.35 | −0.62 |
| 1530–1629 . . . . . | 0.78 | 0.27 | −0.44 |
| 1630–1729 . . . . . | 0.73 | 0.9 | 0.06 |
| 1730–1829 . . . . . | 0 | 1.18 | −0.01 |
| 1830–1929 . . . . . | 0.59 | −0.08 | −0.31 |
| 1930–2029 . . . . . | 0.21 | 0.04 | −0.04 |
| 2030–2129 . . . . . | 0 | 0.57 | −0.12 |
| 2130–2229 . . . . . | 0.11 | 0.19 | −0.18 |
| 2230–2329 . . . . . | −0.17 | 0.25 | −0.2 |
| 2330–2429 . . . . . | 0.06 | −0.16 | 0.09 |
| 2430–2508 . . . . . | 2.63 | 2.66 | −0.19 |

**Notes.** The first column shows the $\ell$-range, the second shows the difference $\Delta_{\ell_{\max}=999}$ between the $\chi^2$ values for a ΛCDM best-fit model obtained using either a likelihood with $\ell_{\max} = 999$ or the baseline, i.e., $\Delta_{999} \equiv \left(\chi^2_{\ell_{\max}=999} - \chi^2_{\rm BASE}\right)_{\Lambda\rm CDM}$. The $\ell_{\max} = 999$ case was run fixing the foreground parameters to the best fit of the baseline case. The third column is the same as the second, but for $\ell_{\max} = 1404$. The fourth column shows the difference $\Delta_{A_L}$ between the $\chi^2$ values obtained in the ΛCDM+$A_L$ and the ΛCDM frameworks. In this case, all the foreground and nuisance parameters were free to vary in the same way as in the baseline case.

the difference in best-fit models between the $\ell_{\max} = 999$ case and the reference case. However, the values in Table 18 also show that the $\ell_{\max} = 999$ cosmology is disfavoured by the multipole region between $\ell \approx 1330$–$1430$, before the fifth peak. The $\ell_{\max} = 999$ model predicts too little power in this multipole range, which can be better fit if the position of the fifth peak moves to lower multipoles. As a consequence, $\theta$ shifts to higher values when including $\ell_{\max} \gtrsim 1400$.

Concerning the shifts in $\Omega_c h^2$, $A_s$ and $\tau$, Fig. 35 shows that these parameters converge to the full baseline solution between $\ell_{\max} = 1404$ and $\ell_{\max} = 1505$. The $\Delta\chi^2$ values in Table 18 between the best-fit $\ell_{\max} = 1404$ case and the baseline suggest that the $\ell_{\max} = 1404$ cosmology is disfavoured by the multipole region $\ell = 1430$–$1530$ (fifth peak), and – at somewhat lower significance – by the regions close to the fourth peak ($\ell \approx 1130$–$1230$) and the sixth peak ($\ell \approx 1730$–$1829$). The pink line in Fig. 32 shows the differences between the $\ell_{\max} = 1404$ best-fit model and the baseline, and it suggests that the $\ell_{\max} = 1404$ cosmology predicts an amplitude of the high-$\ell$ peaks that is too large.

This effect can be compensated by more lensing, which can be obtained with greater values of $\Omega_c h^2$ and $\ln(10^{10} A_s)$, as well as a greater value of $\tau$ to compensate for the increase in $A_s$ in the normalization of the spectra, as observed when considering $\ell_{\max} \gtrsim 1500$. This also explains why the baseline ($\ell_{\max} = 2508$) best-fit solution prefers a value of the optical depth which is $0.8\sigma$ higher than the mean value of the Gaussian prior ($\tau = 0.07 \pm 0.02$), $\tau = 0.085 \pm 0.018$. In order to verify this interpretation, we performed the following test (using the `CAMB` code instead of `PICO`). We fixed the theoretical lensing power spectrum to the best-fit parameters preferred by the $\ell_{\max} = 1404$ cosmology, and estimated cosmological parameters using the baseline likelihood. This is the "CAMB, FIX LENS" case in Fig. 35, which shows that cosmological parameters shift back to the values preferred at $\ell_{\max} = 1404$ ("CAMB, $\ell_{\max} = 1404$") if they cannot alter the amount of lensing in the model.

Since the $\ell \approx 1400$−$1500$ region is also affected by the deficit at $\ell = 1450$ (described in Sect. 3.8), we tested whether excising this multipole region from the baseline likelihood (with $\ell_{\max} = 2508$) has an impact on the determination of cosmological parameters. The results in Fig. 35 (case "CUT $\ell = 1404$−$1504$") show that the parameter shifts are at the level of 0.47, −0.29, 0.38, and $0.45\sigma$ on $\Omega_b h^2$, $\Omega_c h^2$, $\theta$, and $n_s$, respectively (0.39, 0.09, 0.24, and $0.29\sigma$ expected from Eq. (53)), confirming that this multipole region has some impact on the parameters, although it cannot completely account for the shift between the $\ell_{\max} \approx 1400$ case and the baseline.

We also estimated cosmological parameters including only multipoles $\ell > 1000$ ("LMIN 1000" case), and compared them to the "LMAX 999" case[14] (see also Appendix C.5). The two-dimensional posterior distributions in Fig. 33 show the complementarity of the information from $\ell \leq 999$ and $\ell \geq 1000$, with degeneracy directions between pairs of parameters changing in these two multipole regimes. The $\ell_{\min} = 1000$ likelihood sets constraints on the amplitude of the spectra $A_s e^{-2\tau}$ and on $n_s$ that are almost a factor of 2 weaker than the ones obtained with the baseline likelihood, and somewhat higher than the ones obtained with $\ell_{\max} = 999$. The value of $\tau$ is thus more effectively determined by its prior and shifts downward by $0.59\sigma$ with respect to the baseline. The value of $\Omega_c h^2$ shifts upward by $1.7\sigma$ (cf. $0.8\sigma$ expected from Eq. (53)). Whether this change is just due to a statistical fluctuation is still a matter of investigation.

However, since parameter shifts are correlated, we evaluated whether the ensemble of the shifts in all cosmological parameters between the $\ell_{\max} = 999$ and $\ell_{\min} = 1000$ cases are compatible with statistical expectations. In order to do so, we computed the $\chi^2_\Delta$ statistic of the shift as

$$\chi^2_\Delta = \sum_{ij} \Delta_i \Sigma^{-1}_{ij} \Delta_j, \tag{54}$$

where $\Delta_i$ is the difference in best-fit value of the $i$th parameter between the $\ell_{\max} = 999$ and $\ell_{\min} = 1000$ cases and $\Sigma$ is the covariance matrix of the expected shifts, calculated as the sum

[14] During the revision of this paper, we noticed that the $\ell > 1000$ case explores regions of parameter space that are outside the optimal `PICO` interpolation region, as also remarked by Addison et al. (2016). This inaccuracy mainly affected this particular test for constraints on $n_s$ and $\Omega_b h^2$: the error bars for these parameters were underestimated by a factor of about 2 while the mean values were misestimated by about $0.8\sigma$ with respect to runs performed with `CAMB`. Nevertheless, we found that for all other parameters, and in all other likelihood tests presented in this section, this problem did not arise, since the explored parameter space was entirely contained in the `PICO` interpolation region so as to guarantee accurate results, as also detailed in Sect. C.5. Furthermore, this inaccuracy does not change any of the conclusions of this paper. We therefore decided to keep in Fig. 35 the results obtained with `PICO` but we have added results for the $\ell > 1000$ case obtained with `CAMB` (case "CAMB, $l_{\min} = 1000$").

of the parameter covariance matrices obtained in each of the two cases, ignoring correlations between the two datasets. We include in this calculation the ΛCDM parameters ($\Omega_b h^2$, $\Omega_c h^2$, $\theta$, $n_s$, $A_s \exp(-2\tau)$), excluding $\tau$, since the constraints on this parameter are dominated by the same prior in both cases, and using $A_s \exp(-2\tau)$ instead of $\ln(10^{10} A_s)$, since the latter is very correlated with $\tau$ and the $TT$ power spectrum is mostly sensitive to the combination $A_s \exp(-2\tau)$. Finally, we estimate the $\chi^2_\Delta$ both in the case where we leave the foregrounds free to vary or in the case where we fix them to the best fit of the baseline `PlikTT` + taupriorsolution. Assuming that $\chi^2_\Delta$ has a $\chi^2$ distribution for 5 degrees of freedom, we find that the shifts observed in the data are consistent with simulations at the $1.2\sigma$ ($1.1\sigma$ with fixed foregrounds) level for the case where we do not include the low-$\ell$ $TT$ likelihood at $\ell < 30$ to the $\ell_{max} = 999$ case, and at the $1.5\sigma$ ($1.4\sigma$ with fixed foregrounds) level for the case where we include the low-$\ell$ $TT$ likelihood. We also find that the use of $A_s \exp(-2\tau)$ instead of $\ln(10^{10} A_s)$ changes these significances only in the case where we include the low-$\ell$ $TT$ likelihood to the $\ell_{max} = 999$ case and leave the foregrounds free to vary, in which case we find consistency at the level of $1.8\sigma$, in agreement with the findings of Addison et al. (2016; although in this case the use of $\ln(10^{10} A_s)$ and the exclusion of $\tau$ makes this test less indicative of the true significance of the shifts). In all cases, we do not find evidence for a discrepancy between the two datasets. A more precise and extended evaluation and discussions of these shifts, based on numerical simulations, will be presented in a future publication.

#### 4.1.7. Impact of varying $A_L$

Figure 36 (left) displays the impact of various choices on the value of the lensing parameter $A_L$ in the ΛCDM+$A_L$ framework. The baseline likelihood prefers a value of $A_L$ that is about $2\sigma$ greater than the physical value, $A_L = 1$. It is clear that this preference only arises when data with $\ell_{max} \gtrsim 1400$ are included, and it is caused by the same effects as we proposed in Sect. 4.1.6 to explain the shifts in parameters at $\ell_{max} \gtrsim 1400$ in the ΛCDM case. More lensing helps to fit the data in the $\ell \approx 1300-1500$ region, as indicated by the $\chi^2$ differences between the ΛCDM+$A_L$ best-fit and the ΛCDM one in Table 18. This drives the value of $A_L$ to $1.159 \pm 0.090$ with `PlikTT`+tauprior, $1.8\sigma$ higher than expected. The case "ΛCDM+$A_L$" of Fig. 35 also shows that opening up this unphysical degree of freedom shifts the other cosmological parameters at the $1\sigma$ level; e.g., $\Omega_c h^2$ and $A_s$ shift closer to the values preferred in the ΛCDM case when using $\ell_{max} \lesssim 1400$. While in the ΛCDM case high values of these parameters allow increasing lensing, in the ΛCDM+$A_L$ case this is already ensured by a high value of $A_L$, so $\Omega_c h^2$ and $A_s$ can adopt values that better fit the $\ell \lesssim 1400$ range. When using `PlikTT` in combination with the lowTEB likelihood, the deviation increases to $2.4\sigma$, $A_L = 1.204 \pm 0.086$,[15] due to the fact that more lensing allows smaller values of $\Omega_c h^2$ and $A_s$ and a greater value of $n_s$, better fitting the deficit at $\ell \approx 20$ in the temperature power spectrum (see Planck Collaboration XIII 2016, Sect. 5.1.2 and Fig. 13).

#### 4.1.8. Impact of varying $N_{eff}$

We have investigated the effect of opening up the $N_{eff}$ degree of freedom in order to assess the robustness of the constraints on the ΛCDM extensions, which rely heavily on the high-$\ell$ tail of the data. Figure 36 (right) shows that $N_{eff}$ departs from the standard 3.04 value by about $1\sigma$ when using `PlikTT`+tauprior, $N_{eff} = 2.7 \pm 0.33$. The $\chi^2$ improvement for this model over ΛCDM is only $\Delta\chi^2 = 1.5$. We note that when the lowTEB likelihood (or alternatively, the low-$\ell$ $TT$ likelihood plus the prior on $\tau$) is used in combination with `PlikTT`, the value of $N_{eff}$ shifts higher by about $1\sigma$, $N_{eff} = 3.09 \pm 0.29$. This shift is about a factor 2 more than the one expected from Eq. (53), $0.5\sigma$, between the `PlikTT`+tauprior and `PlikTT`+tauprior+low-$\ell$ $TT$ cases. This shift is due to the fact that the deficit at $\ell \approx 20$ is better fit by higher $n_s$ and, as a consequence, an increase in $N_{eff}$ helps decreasing the enhanced power at high $\ell$.

Figure 36 also shows that, not surprisingly, the most extreme variations as compared to the reference case (less than $1\sigma$) arise when the high-resolution data are dropped (by reducing $\ell_{max}$ or by removing the 217 GHz channel), owing to the strong dependence of the $N_{eff}$ constraints on the damping tail.

Having opened up this degree of freedom, the standard parameters are now about $1\sigma$ away (see case "ΛCDM+$N_{eff}$" of Fig. 35), and such a model would prefer quite a low value of $H_0$, which would then be at odds with priors derived from direct measurements (see Planck Collaboration XIII 2016, for an in-depth analysis).

### 4.2. Intercomparison of likelihoods

In addition to the baseline high-$\ell$ `Plik` likelihood, we have developed four other high-$\ell$ codes, `CamSpec`, `Hillipop`, `Mspec`, and `Xfaster`. `CamSpec` and `Xfaster` have been described in separate papers (Planck Collaboration XV 2014; Rocha et al. 2011), and brief descriptions of `Mspec` and `Hillipop` are given in Appendix D. These codes have been used to perform data consistency tests, to examine various analysis choices, and to cross-check each other by comparing their results and ensuring that they are the same. In general, we find good agreement between the codes, with only minor differences in cosmological parameters.

The `CamSpec`, `Hillipop`, and `Mspec` codes are, like `Plik`, based on pseudo-$C_\ell$ estimators and an analytic calculation of the covariance (Efstathiou 2004, 2006), with some differences in the approximations used to calculate this covariance. The `Xfaster` code (Rocha et al. 2011) is an an approximation to the iterative, maximum likelihood, quadratic band-power estimator based on a diagonal approximation to the quadratic Fisher matrix estimator (Rocha et al. 2011, 2010), with noise bias estimated using difference maps, as described in Planck Collaboration IX (2016). For temperature, all of the codes use the same Galactic masks, but they differ in point-source masking: `Hillipop` uses a mask based on a combination of $S/N > 7$ and cuts based on flux, while the others use the baseline $S/N > 5$ mask described in Appendix A. The codes also differ in foreground modelling, in the choice of data combinations, and in the $\ell$-range. For the comparison presented here, all make use of half-mission maps.

Figure 38 shows a comparison of the power spectra residuals and error bars from each code, while Fig. E.5 in Appendix E.4 compares the combined spectra with the best-fit model. In temperature, the main feature visible in these plots is an overall nearly constant shift, up to $10\,\mu\mathrm{K}^2$ in some cases. This represents a real difference in the best-fit power each code attributes to foregrounds. For context, it is useful to note the statistical uncertainty on the foregrounds; for example, the $1\sigma$ error on the total foreground power at 217 GHz at $\ell = 1500$ is $2.5\,\mu\mathrm{K}^2$ (calculated here with `Mspec`, but similar for the other codes). Shifts

[15] These results were obtained with the `PICO` code, and are thus close to but not identical to those obtained with `CAMB` and reported in Planck Collaboration XIII (2016).

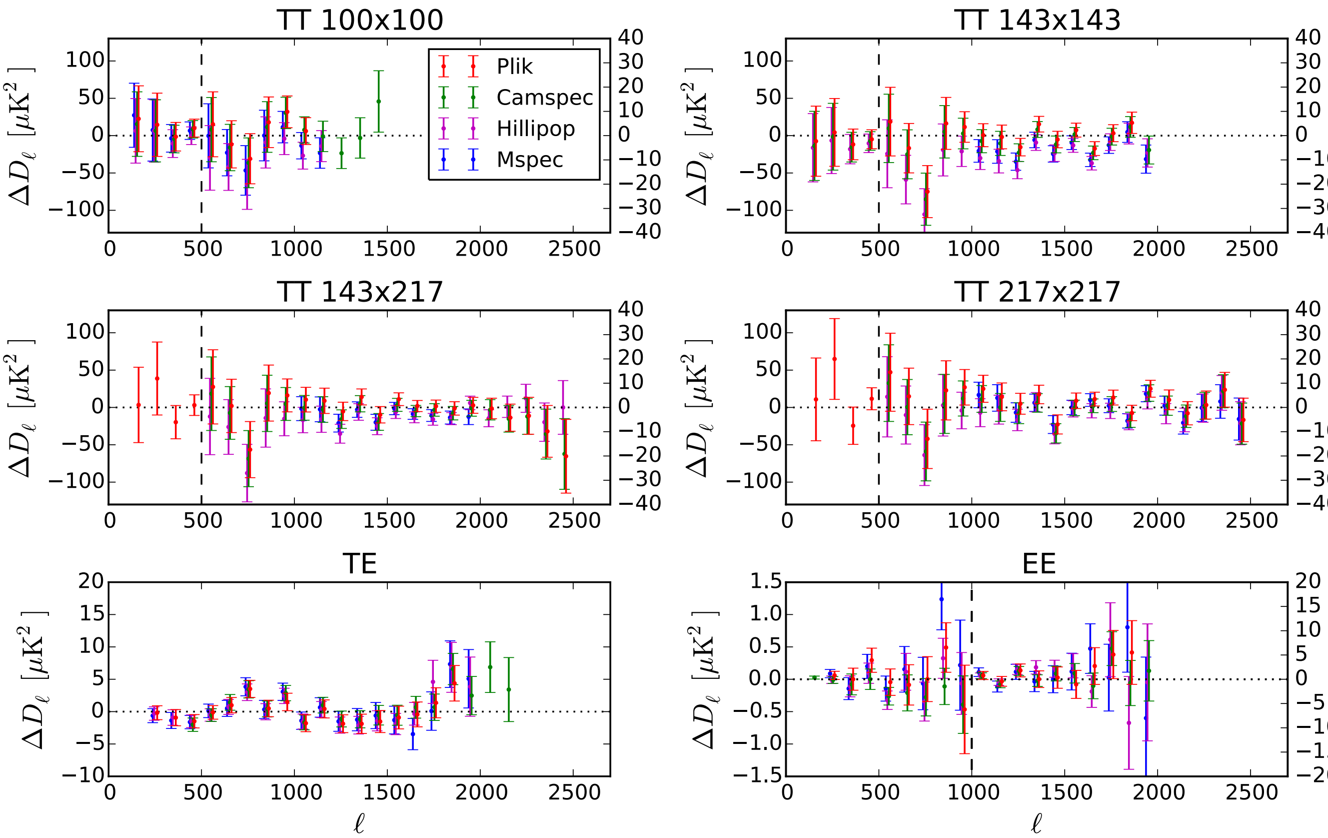


**Fig. 38.** Comparison of power spectra residuals from different high-$\ell$ likelihood codes. The figure shows "data/calib − FG − $\mathtt{Plik}_{\rm CMB}$", where "data" stands for the empirical cross-frequency spectra, "FG" and "calib" are the best-fit foreground model and recalibration parameter for each individual code at that frequency, and the best-fit model $\mathtt{Plik}_{\rm CMB}$ is subtracted for visual presentation. These plots thus show the difference in the amount of power each code attributes to the CMB. The power spectra are binned in bins of width $\Delta\ell = 100$. The $y$-axis scale changes at $\ell = 500$ for $TT$ and $\ell = 1000$ for $EE$ (vertical dashes).

of this level do not lead to very large differences in cosmological parameters except in a few cases that we discuss.

For easier visual comparison of error bars, we show in Fig. 39 the ratios of each code's error bars to those from `Plik`. These have been binned in bins of width $\Delta\ell = 100$, and are thus sensitive to the correlation structure of each code's covariance matrix, up to 100 multipoles into the off-diagonal. For all the codes and for both temperature and polarization, the correlation between multipoles separated by more than $\Delta\ell = 100$ is less than 3%, so Fig. 39 contains the majority of the relevant information about each code's covariance.

A few differences are visible, mostly at high frequency, when the 217 GHz data are used. First, the `Hillipop` error bars in $TT$ for $143 \times 217$ become increasingly tighter than the other codes at $\ell > 1700$. This is because `Hillipop`, unlike the other codes, gives non-zero weight to $143 \times 217$ spectra when both the 143 and the 217 GHz maps come from the same half-mission. This leads to a slight increase in power at high $\ell$ compared to `Plik`, as can be seen in Fig. 38. Conversely, the `Hillipop` error bars are slightly larger by a few percent at $\ell < 1700$; however the source of this difference is not understood. Second, the `Mspec` error bars in temperature are increasingly tighter towards higher frequency, as compared to other codes; for $217 \times 217$, `Mspec` uncertainties are smaller by 5–10% for $\ell$ between 1000 and 2000. This arises from the `Mspec` map-based Galactic cleaning procedure, which removes excess variance due to CMB–foreground correlations by subtracting a scaled 545 GHz map. However, for polarization, where one must necessarily clean with the noisier 353 GHz maps, the `Mspec` error bars for $TE$ and $EE$ become larger. `CamSpec`, which also performs a map cleaning for low-$\ell$ polarization, switches to a power-spectrum cleaning at higher $\ell$ to mitigate this effect.

The differences in ΛCDM parameters from $TT$ are shown in Table 19. Generally, parameters agree to within a fraction of$\sigma$, but with some differences we discuss. One thing to keep in mind in interpreting this comparison is that these differences are not necessarily indicative of systematic errors. Some of the differences are expected due to statistical fluctuations because different codes weight the data differently.

One of the biggest differences with respect to the baseline code is in $n_{\rm s}$, which is higher by about $0.45\sigma$ for `CamSpec`, with a related downward shift of $A_{\rm s}{\rm e}^{-2\tau}$. To put these shifts into perspective, we refer to the whisker plots of Figs. 35 and 36 which compare `CamSpec` $TT$ results with `Plik` in the ΛCDM case (base and extended). A difference in $n_{\rm s}$ of about $0.16\sigma$ between `Plik` and `CamSpec` can be attributed to the inclusion in `Plik` of the first 500 multipoles for $143 \times 217$ and $217 \times 217$; these multipoles are excluded in `CamSpec` (see also Sect. 4.1.2). Indeed, cutting out those multipoles in `Plik` brings $n_{\rm s}$ closer by $0.16\sigma$ to the `CamSpec` value and slightly degrades the constraint on $n_{\rm s}$ compared to the full `Plik` result. Using Eq. (53), we see that the shift and degradation in constraining power are consistent with expectations. A similar $0.16\sigma$ shift can be attributed to different dust templates. `CamSpec` uses a steeper power law index (−2.7). Using the `CamSpec` template in `Plik` brings $n_{\rm s}$ closer to the `CamSpec` value. Allowing the power law index of the galactic

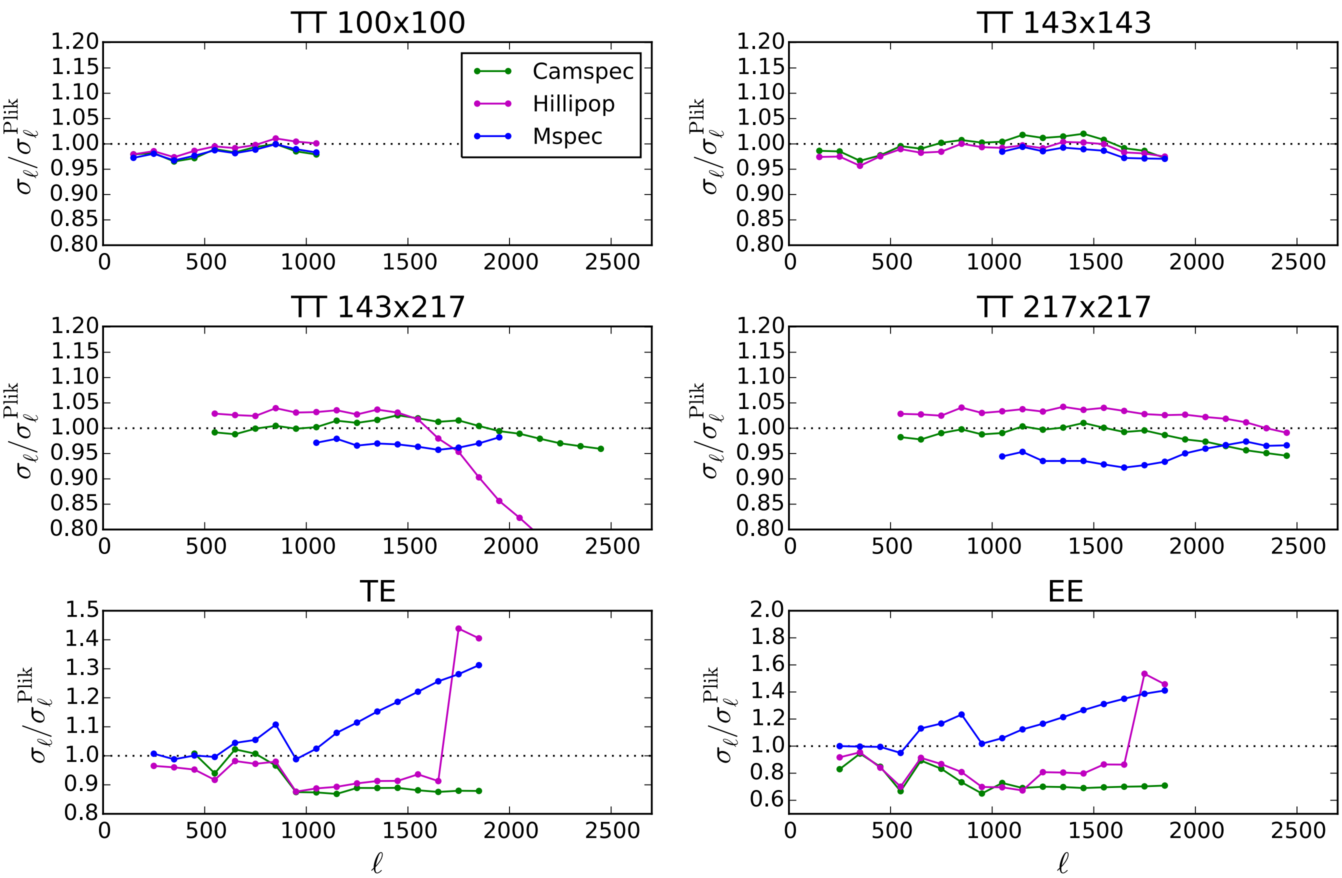


**Fig. 39.** Comparison of error bars from the different high-$\ell$ likelihood codes. The quantities plotted are the ratios of each code's error bars to those from `Plik`, and are for bins of width $\Delta\ell = 100$. Results are shown only in the $\ell$ range common to `Plik` and the code being compared.

template to vary when exploring cosmological parameters yields a slightly shallower slope (see Sect. 4.1.2). The slope of the dust template is mainly determined at relatively high $\ell$, i.e., in the regime where it is hardest to determine the template accurately since the dust contribution is only a small fraction of the CIB and point-source contributions (see the $\ell \gtrsim 1000$ parts of Figs. 19 and 20). The remaining difference of $0.13\sigma$ arises from differences in data preparation (maps, calibration, binning) and covariance estimates. We therefore believe that a $0.2\sigma$ is a conservative upper bound of the systematic error in $n_{\rm s}$ associated with the uncertainties in the modelling of foregrounds, which is the biggest systematic uncertainty in $TT$.

A shift that is less well understood is the $\approx 1\sigma$ shift in $A_{\rm s}{\rm e}^{-2\tau}$ between `Plik` and `Hillipop`. The preference for a lower amplitude from `Hillipop` is sourced by the lower power attributed to the CMB, seen in Fig. 38. With $\tau$ partially fixed by the prior, this implies lower $A_{\rm s}$ and hence a smaller lensing potential envelope, explaining the somewhat lower value of $A_{\rm L}$ found by `Hillipop`. Tests performed with the same code suggest that $1\sigma$ is too great a shift to be explained simply by the different foreground models, so some part of it must be due to the different data weighting; as can be seen in Fig. 39, `Hillipop` gives less weight to $500 \lesssim \ell \lesssim 1500$, and slightly more outside of this region.

This comparison also shows the stability of the results with respect to the Galactic cleaning procedure. `Mspec` and `Plik` use different procedures, yet their parameter estimates agree to better than $0.5\sigma$ (see Appendix D.1). But we note that the `Plik`–`CamSpec` differences are higher in the polarization case, and can reach $1\sigma$, as can be judged from the whisker plot in polarization of Fig. C.10.

### *4.3. Consistency of Poisson amplitudes with source counts*

The Poisson component of the foreground model is sourced by shot-noise from astrophysical sources. In this section we discuss the consistency between the measured Poisson amplitudes and other probes and models of the source populations from which they arise. The Poisson amplitude priors that we calculate are not used in the main analysis, because they improve uncertainties on the cosmological parameters by at most 10%, and only for a few extensions; instead they serve as a self-consistency check.

This type of check was also performed in Like13, which we update here by:

1. developing a new method for calculating these priors that is accurate enough to give realistic uncertainties on Poisson predictions (for the first time);
2. including a comparison of more theoretical models;
3. taking into account the 2015 point-source masks.

In Like13 the Poisson power predictions were calculated via

$$C_\ell = \int_0^{S_{\rm cut}} {\rm d}S\, S^2 \frac{{\rm d}N}{{\rm d}S}, \qquad (55)$$

where ${\rm d}N/{\rm d}S$ is the differential number count, $S_{\rm cut}$ is an effective flux-density cut above which sources are masked, and the integral was evaluated independently at each frequency. Although it is adequate for rough consistency checks, Eq. (55) ignores the facts that the 2013 point-source mask was built from a union of sources detected at different frequencies, and that the *Planck* flux-density cut varies across the sky, and it also ignores the effect of Eddington bias. In order to accurately account for all of

**Table 19.** Comparison between the parameter estimates from different high-$\ell$ codes.

| Parameter | `Plik` | `CamSpec` | `Hillipop` | `Mspec` | `Xfaster` (`SMICA`) |
|---|---|---|---|---|---|
| $\Omega_b h^2$ . . . . . . . . . . | $0.02221 \pm 0.00023$ | $0.02224 \pm 0.00023$ | $0.02218 \pm 0.00023$ | $0.02218 \pm 0.00024$ | $0.02184 \pm 0.00024$ |
| $\Omega_c h^2$ . . . . . . . . . . | $0.1203 \pm 0.0023$ | $0.1201 \pm 0.0023$ | $0.1201 \pm 0.0022$ | $0.1204 \pm 0.0024$ | $0.1202 \pm 0.0023$ |
| $100\theta_{MC}$ . . . . . . . . | $1.0406 \pm 0.00047$ | $1.0407 \pm 0.00048$ | $1.0407 \pm 0.00046$ | $1.0409 \pm 0.00050$ | $1.041 \pm 0.0005$ |
| $\tau$ . . . . . . . . . . . . . | $0.085 \pm 0.018$ | $0.087 \pm 0.018$ | $0.075 \pm 0.019$ | $0.075 \pm 0.018$ | $0.069 \pm 0.019$ |
| $10^9 A_s e^{-2\tau}$ . . . . . . | $1.888 \pm 0.014$ | $1.877 \pm 0.014$ | $1.870 \pm 0.011$ | $1.878 \pm 0.012$ | $1.866 \pm 0.015$ |
| $n_s$ . . . . . . . . . . . . | $0.962 \pm 0.0063$ | $0.965 \pm 0.0066$ | $0.961 \pm 0.0072$ | $0.959 \pm 0.0072$ | $0.960 \pm 0.0071$ |
| $\Omega_m$ . . . . . . . . . . . | $0.3190 \pm 0.014$ | $0.3178 \pm 0.014$ | $0.3164 \pm 0.014$ | $0.3174 \pm 0.015$ | $0.3206 \pm 0.015$ |
| $H_0$ . . . . . . . . . . . | $67.0 \pm 1.0$ | $67.1 \pm 1.0$ | $67.1 \pm 1.0$ | $67.1 \pm 1.1$ | $66.8 \pm 1.0$ |

**Notes.** Each column gives the results for various high-$\ell$ $TT$ likelihoods at $\ell > 50$ when combined with a prior of $\tau = 0.07 \pm 0.02$. The `SMICA` parameters were obtained for $\ell_{max} = 2000$.

these effects, we now calculate the Poisson power as

$$C_\ell^{ij} = \int_0^\infty \mathrm{d}S_1 \ldots \mathrm{d}S_n \, S_i S_j \frac{\mathrm{d}N(S_1, \ldots, S_n)}{\mathrm{d}S_1 \ldots \mathrm{d}S_n} I(S_1, \ldots, S_n), \quad (56)$$

where the frequencies are labelled $1 \ldots n$, the differential source count model, $\mathrm{d}N/\mathrm{d}S$, is now a function of the flux densities at *each* frequency, and $I(S_1, \ldots, S_n)$ is the joint "incompleteness" of our catalogue for the particular cut that was used to build the point-source mask.

The joint incompleteness was determined by injecting simulated point sources into the *Planck* sky maps, using the procedure described in Planck Collaboration XXVI (2016). The same point-source detection pipelines that were used to produce the Second *Planck* Catalogue of Compact Sources (PCCS2) were run on the injected maps, producing an ensemble of simulated *Planck* sky catalogues with realistic detection characteristics. The joint incompleteness is defined as the probability that a source would not be included in the mask as a function of the source flux density, given the specific masking thresholds being considered. The raw incompleteness is a function of sky location, because the *Planck* noise varies across the sky. The incompleteness that appears in Eq. (56) is integrated over the region of the sky used in the analysis; the injection pipeline estimates this quantity by injecting sources only into these regions.

Equation (56) can be applied to any theoretical model which makes a prediction for the multi-frequency $\mathrm{d}N/\mathrm{d}S$. We have adopted the following models.

1. For radio galaxies we have two models. The first is the Tucci et al. (2011) model, updated to include new source-count measurements from Mocanu et al. (2013). We also consider a phenomenological model that is a power law in flux density and frequency, and assumes that the sources' spectral indices are Gaussian-distributed with mean $\bar{\alpha}$ and standard deviation $\sigma_\alpha$; we use different values for $\bar{\alpha}$ and $\sigma_\alpha$ above and below 143 GHz. We shall refer to this second model as the "power-law" model, and the differential source counts are given by

$$\frac{\mathrm{d}N(S_1, S_2, S_3)}{\mathrm{d}S_1 \mathrm{d}S_2 \mathrm{d}S_3} = \frac{A(S_1 S_2 S_3)^{\gamma-1}}{2\pi\sigma_{12}\sigma_{23}} \times \exp\left[-\frac{(\alpha(S_1, S_2) - \bar{\alpha}_{12})^2}{2\sigma_{12}^2} - \frac{(\alpha(S_2, S_3) - \bar{\alpha}_{32})^2}{2\sigma_{32}^2}\right], \quad (57)$$

   where labels 1–3 refer to *Planck* 100, 143, and 217 GHz and $\alpha(S_i, S_j) = \ln(S_j/S_i)/\ln(\nu_j/\nu_i)$. Both radio models are excellent fits to the available source-count data, and we take the difference between them as an estimate of model uncertainty. With the power-law model we are additionally able to propagate uncertainties in the source count data to the final Poisson estimate via MCMC.
2. For dusty galaxies we use the Béthermin et al. (2012) model, as in Planck Collaboration XXX (2014). The model is in good agreement with the number counts measured with te *Spitzer* Space Telescope and the *Herschel* Space Observatory. It also gives a reasonable CIB redshift distribution, which is important for cross-spectra, and is a very good fit to CIB power spectra (see Béthermin et al. 2013). In contrast to the radio-source case, the major contribution to the dusty galaxy Poisson power arises from sources with flux densities well below the cuts; for example, we note that decreasing the flux-density cuts by a factor of 2 decreases the Poisson power by less than 1% at the relevant frequencies. In this case, Eq. (55) is a sufficient and more convenient approximation, and we make use of it when calculating Poisson levels for dusty galaxies.

We give predictions for Poisson levels for three different masks: (1) the 2013 point-source mask, which was defined for sources detected at $S/N > 5$ at any frequency between 100 and 353 GHz; (2) the 2015 point-source mask, which is frequency-dependent and includes $S/N > 5$ sources detected only at each individual frequency (used by `Plik`, `CamSpec`, and `Mspec` in this work); and (3) the `Hillipop` mask, which is also frequency-dependent and involves both a S/N cut and a flux-density cut[16].

Table 20 summarizes the main results of this section. Generally, we find good agreement between the priors from source counts and the posteriors from chains, with the priors being much more constraining. The exception to the good agreement is at 100 GHz where the prediction is lower than the measured value by around $4\sigma$ for the baseline 2015 mask and $6\sigma$ for the `Hillipop` mask. This is a sign either of a foreground modelling error or (perhaps more likely) of a residual unmodelled systematic in the data. We note that this disagreement was not present in Like13, where the Poisson amplitude at 100 GHz was found to be smaller. We also note that removing the relative calibration prior (Eq. (35)) or increasing the $\ell_{max}$ at 100 GHz by a few hundred reduces the tension in the `Mspec` results. In any case, it is unlikely to affect parameter estimates at all, since very little

[16] We note that the `Hillipop` mask was constructed partly so that Eq. (55) would be an accurate approximation. We find that for the radio contribution is is accurate to 2%, or $1\sigma$, and for the dust contribution it is essentially exact.

**Table 20.** Priors on the Poisson amplitudes given a number of different point-source masks and models.

| Mask | Type | Model | Power spectrum 100 × 100 | 143 × 143 | 143 × 217 | 217 × 217 |
|---|---|---|---|---|---|---|
| Baseline 2013 ....... | Radio | Power-law | 84 ± 3 | 29 ± 1 | 16 ± 1 | 9 ± 1 |
| | Dusty | Bethermin | 4 ± 1 | 13 ± 3 | 41 ± 8 | 129 ± 25 |
| Baseline 2015 ....... | Radio | Power-law | 148 ± 7 | 40 ± 1 | 16 ± 1 | 10 ± 1 |
| | | Tucci | 139 | 40 | 16 | 11 |
| | Dusty | Bethermin | 4 ± 1 | 13 ± 3 | 41 ± 8 | 129 ± 25 |
| | Plik | | 260 ± 28 | 44 ± 8 | 39 ± 10 | 97 ± 11 |
| | Mspec | | 317 ± 46 | 22 ± 13 | 12 ± 7 | 21 ± 9 |
| Hillipop 2015 ...... | Radio | Power-law | 150 ± 7 | 47 ± 2 | 18 ± 1 | 11 ± 1 |
| | | Tucci | 141 | 47 | 18 | 12 |
| | Dusty | Bethermin | 4 ± 1 | 13 ± 3 | 41 ± 8 | 129 ± 25 |
| | Hillipop | | 372 ± 38 | 58 ± 21 | 53 ± 24 | 105 ± 18 |

**Notes.** Entries are $\mathcal{D}_\ell$ at $\ell = 3000$ in $\mu\mathrm{K}^2$ and are given at the effective band centre for each component. Uncertainties on the "power-law" model are statistical errors propagated from uncertainties in the Mocanu et al. (2013) source-count data. Priors on the dust component have formally been calculated only for the `Hillipop` mask, but they are repeated for the other masks, for which they are accurate to better than 1%. The results from different codes, to which these predictions should be compared, use $TT$ $\ell > 50$ data with a prior of $\tau = 0.07 \pm 0.02$. For `Mspec` about 90% of the dusty contribution is cleaned out at the map level, hence the measured values above are in some cases far less than prior value.

cosmological information comes from the multipole range at 100 GHz that constrains the Poisson amplitude.

### 4.4. TE and EE test results

#### 4.4.1. Residuals per frequency and inter-frequency differences

Figure 40 shows the residuals for each frequency and Fig. 41 shows the differences between frequencies of the $TE$ and $EE$ power spectra (the procedure is explained in Appendix C.3.2). The residuals are calculated with respect to the best-fit cosmology as preferred by `PlikTT`+tauprior, although we use the best-fit solution of the `PlikTT`, TE, EE+tauprior run to subtract the polarized Galactic dust contribution.

The binned inter-frequency residuals show deviations at the level of a few $\mu\mathrm{K}^2$ from the best-fit model. These deviations do not necessarily correspond to high values of the $\chi^2$ calculated on the unbinned data (see Table 16). This is because some of the deviations are relatively small for the unbinned data and correctly follow the expected $\chi^2$ distribution. However, if the deviations are biased (e.g., have the same sign) in some $\ell$ range, they can result in larger deviations (and large $\chi^2$) after binning. Thus, the $\chi^2$ calculated on unbinned data is not always sufficient to identify these type of biases. We therefore also use a second quantity, $\chi$, defined as the weighted linear sum of residuals, to diagnose biased multipole regions or frequency spectra:

$$\chi = \boldsymbol{w}^{\mathsf{T}}(\hat{\boldsymbol{C}} - \boldsymbol{C}) \quad \text{with} \quad \boldsymbol{w} = (\operatorname{diag} \mathsf{C})^{-1/2}, \tag{58}$$

where $\hat{\boldsymbol{C}}$ is the unbinned vector of data in the multipole region or frequency spectrum of interest, $\boldsymbol{C}$) is the corresponding model, and $\boldsymbol{w}$ is a vector of weights, equal to the inverse standard deviation evaluated from the diagonal of the corresponding covariance matrix $\mathsf{C}$. The $\chi$ statistic is distributed as a Gaussian with zero mean and standard deviation equal to

$$\sigma_\chi = \sqrt{\boldsymbol{w}^{\mathsf{T}}\mathsf{C}\boldsymbol{w}}. \tag{59}$$

We then define the normalized $\chi_{\mathrm{norm}}$ as the $\chi$ in units of standard deviation,

$$\chi_{\mathrm{norm}} = \chi/\sigma_\chi. \tag{60}$$

The $\chi_{\mathrm{norm}}$ values that we obtain for different frequency power spectra are given in Table 16.

For $EE$, the worst-behaved spectra from the $\chi_{\mathrm{norm}}$ point of view are 143 × 143 (3.7$\sigma$ deviation) and 100 × 217 (−3.0$\sigma$), while from the $\chi^2$ point of view, the worst is 100 × 143 (PTE = 3.9%). For $TE$, the worst from the $\chi_{\mathrm{norm}}$ point of view are 100 × 217 (5$\sigma$), 100 × 100 (3.7$\sigma$), and 143 × 143 (−2.2$\sigma$), while from the $\chi^2$ point of view the worst is 100 × 100 (PTE = 0.43%). The extreme deviations from the expected distributions show that the frequency spectra are not described very accurately by our data model. This is also clear from Fig. 41, which shows that there are differences of up to 5$\sigma$ between pairs of foreground-cleaned spectra.

However, as the co-added residuals in Fig. 29 show, systematic effects in the different frequency spectra appear to average out, leaving relatively small residuals with respect to the `PlikTT`+tauprior best-fit cosmology. In other words, these effects appear not to be dominated by common modes between detector sets or across frequencies. This is also borne out by the good agreement between the data and the expected polarization power spectra conditioned on the temperature ones, as shown in the conditional plots of Fig. 34.

#### 4.4.2. TE and EE robustness tests

For $TE$ and $EE$, we ran tests of robustness similar to those applied earlier to $TT$. These are presented in Appendix C.3.5, and the main conclusions are the following. We find that the `Plik` cosmological results are affected by less than 1$\sigma$ when using detset cross-spectra instead of half-mission ones. This is also the case when we relax the dust amplitude priors, when we marginalize over beam uncertainties, or when we change $\ell_{\mathrm{min}}$

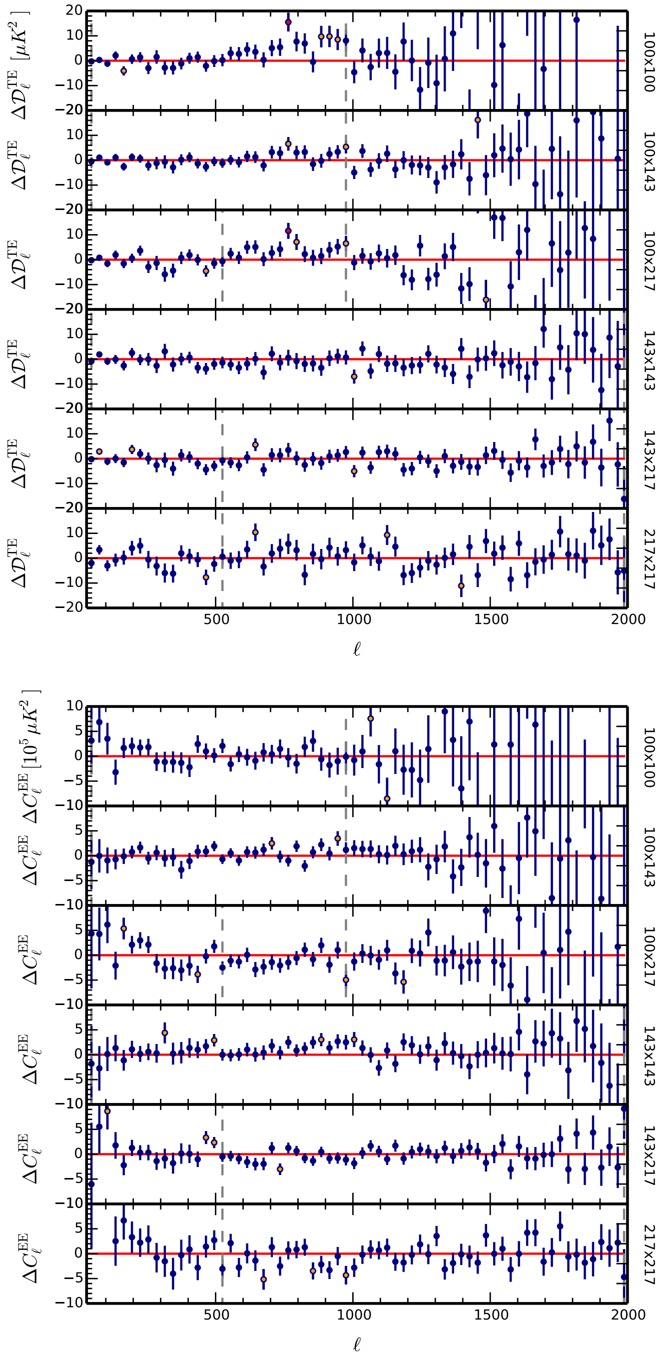


**Fig. 40.** Residual frequency power spectra after subtraction of the PlikTT+tauprior best-fit model. We clean Galactic dust from the spectra from using the best-fit solution of PlikTT, TE, EE+tauprior. The residuals are relative to the baseline HM power spectra (blue points, except for those that deviate by at least 2 or $3\sigma$, which are shown in orange or red, respectively). The vertical dashed lines delimit the $\ell$ ranges retained in the likelihood. *Upper*: *TE* power spectra. *Lower*: *EE* power spectra.

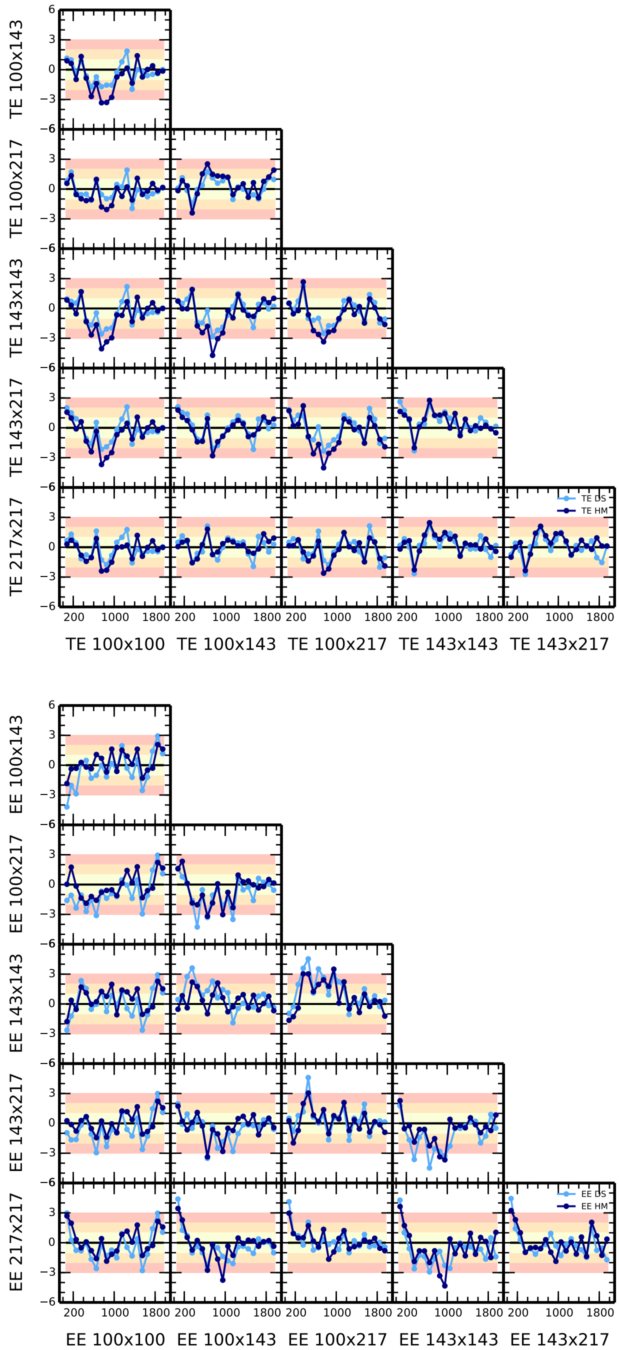


**Fig. 41.** Inter-frequency foreground-cleaned power-spectra differences. Each panel shows the difference of two frequency power spectra, that indicated on the left axis minus that on the bottom axis, after subtracting foregrounds using the best-fit PlanckTT+lowP foreground solutions. Differences are shown for both the HM power spectra (dark blue) and the DS power spectra (light blue).

or $\ell_{\max}$. The alternative CamSpec likelihood has larger shifts, but still smaller than $1\sigma$ in *TE* and $0.5\sigma$ in *EE*. However, we also see larger shifts (more than $2\sigma$ in *TE*) with Plik when some frequency channels are dropped; and, when they are varied, the beam leakage parameters adopt much higher values than expected from the prior, while still leaving some small discrepancies between individual cross-spectra that have yet to be explained.

These results shows that our data model leaves residual instrumental systematic errors and is not yet sufficient to take advantage of the full potential of the HFI polarization information. Indeed, the current data model and likelihood code do not account satisfactorily for deviations at the $\mu K^2$ level, even if they can be captured in part by our beam leakage modelling. Nevertheless, the results for the $\Lambda$CDM model obtained from the PlikTE+tauprior and PlikEE+tauprior runs are in good agreement with the results from PlikTT+tauprior (see Appendix C.3.6). This agreement between temperature and polarization results within ΛCDM is not a proof of the accuracy of the co-added polarization spectra and their data model, but rather a check of consistency at the $\mu K^2$ level. This consistency is, of course, a very interesting result in itself. But this comparison of probes cannot yet be pushed further to check for the potential presence of a physical inconsistency within the base model that the data could in principle detect or constrain.

## 5. The full *Planck* spectra and likelihoods

This section discusses the results that are obtained by using the full *Planck* likelihood. Section 5.1 first addresses the question of robustness with respect to the choice of the hybridization scale (the multipole at which we transition from the low-$\ell$ likelihood to the high-$\ell$ likelihood). Sections 5.2 and 5.3 then present the full results for the power spectra and the baseline cosmological parameters. Section 5.4 summarizes the full systematic error budget. Section 5.5 concentrates on the significance of the possibly anomalous structure around $\ell \approx 20$ in this new release. We then introduce in Sect. 5.6 a useful compressed *Planck* high-$\ell$ temperature and polarization CMB-only likelihood, `Plik_lite`, which, when applicable, enables faster parameter exploration. Finally, in Sect. 5.7, we compare the *Planck* 2015 results with the previous results from WMAP, ACT, and SPT.

### 5.1. Insensitivity to hybridization scale

Before we use the low-$\ell$ and high-$\ell$ likelihoods together, we address the question of the hybridization scale, $\ell_{\rm hyb}$, at which we switch from one to the other (neglecting correlations between the two regimes, as we did and checked in Like13). To that end, we focus on the $TT$ case and use a likelihood based on the Blackwell-Rao estimator and the `Commander` algorithm (Chu et al. 2005; Rudjord et al. 2009) as described in Sect. 2.2, since this likelihood can be used to much higher $\ell_{\rm max}$ than the full pixel-based $T, E, B$ one. For this test without polarization data, we assume the same $\tau = 0.07 \pm 0.02$ prior as before.

The whisker plot of Fig. 42 shows the marginal mean and the 68% CL error bars for base-ΛCDM cosmological parameters when $\ell_{\rm hyb}$ is varied from the baseline value of 30 (case "LOWL 30") to $\ell_{\rm hyb}$ =50, 100, 150, 200, and 250, and compared to the PlikTT+tauprior case. The difference between the "LOWL 30" and "PlikTT+tauprior" values shows the effect of the low-$\ell$ dip at $\ell \approx 20$, which reaches $0.5\sigma$ on $n_{\rm s}$. The plot shows that the effect of varying $\ell_{\rm hyb}$ from 30 to 150 is a shift in $n_{\rm s}$ by less than $0.1\sigma$. This is the result of the Gaussian approximation pushed to $\ell_{\rm min} = 30$, already discussed in the simulation section (Sect. 3.6). It would have been much too slow to run the full low-$\ell$ $TEB$ likelihood with $\ell_{\rm max}$ substantially greater than 30, and we decided against the only other option, to leave a gap in polarization between $\ell = 30$ and the hybridization scale chosen in $TT$.

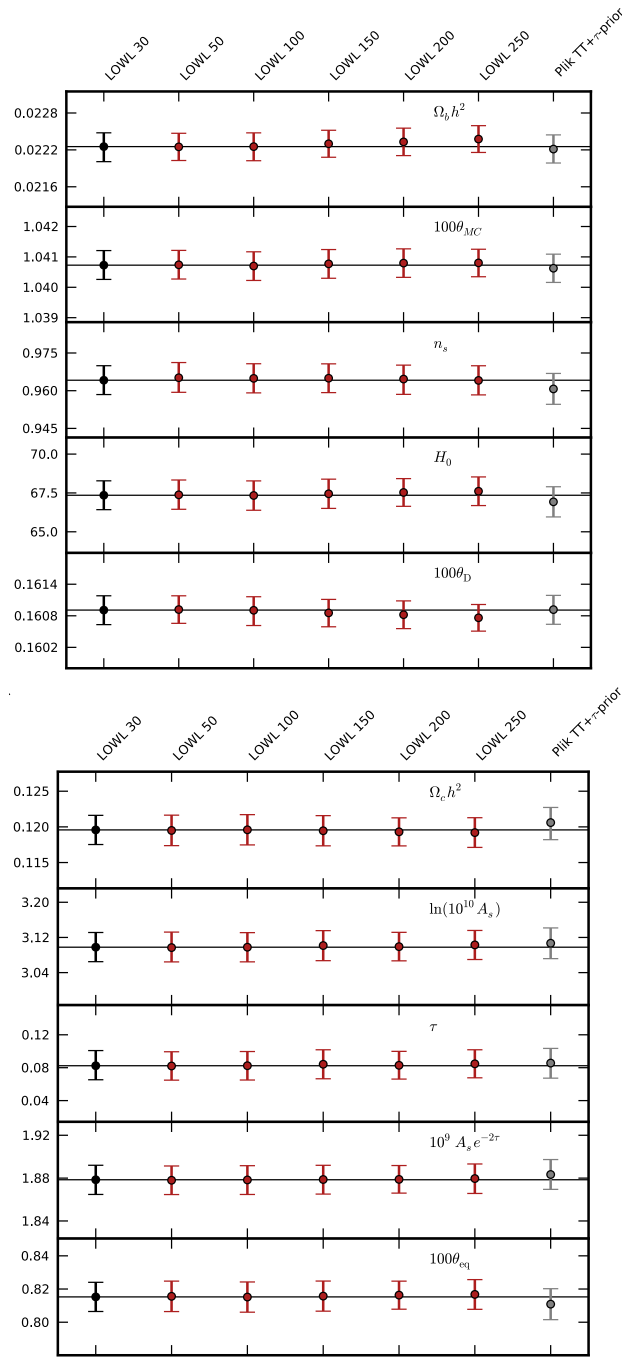


**Fig. 42.** Marginal mean and 68% CL error bars on cosmological parameters estimated with different multipoles for the transition between the low-$\ell$ and the high-$\ell$ likelihood. Here we use only the $TT$ power spectra and a Gaussian prior on the optical depth $\tau = 0.07 \pm 0.02$, within the base-ΛCDM model. "PlikTT+tauprior" refers to the case where we use the `Plik` high-$\ell$ likelihood only.

### 5.2. The Planck 2015 CMB spectra

The visual appearance of *Planck* 2015 CMB co-added spectra in $TT$, $TE$, and $EE$ can be seen in Fig. 50. Goodness-of-fit values can be found in Table E.1 of Appendix E. These differ somewhat from those given previously in Table 16 for `Plik` alone, because the inclusion of low $\ell$ in temperature brings in the $\ell \approx 20$ feature (see Sect. 5.5). Still, they remain acceptable, with PTEs all above 10% (16.8% for $TT$).

With this release, *Planck* now detects 36 extrema in total, consisting of 19 peaks and 17 troughs. Numerical values for the positions and amplitudes of these extrema may be found in Table E.2 of Appendix E.2, which also provides details of the steps taken to derive them. We provide in Appendix E.3 an alternate display of the correlation between temperature and ($E$-mode) polarization by showing their Pearson correlation coefficient and their decorrelation angle versus scale (Figs. E.2 and E.3).

### 5.3. *Planck* 2015 model parameters

Figure 43 compares constraints on pairs of parameters as well as their individual marginals for the base-ΛCDM model. The grey contours and lines correspond to the results of the 2013 release (Like13), which was based on $TT$ and WMAP polarization at low $\ell$ (denoted by WP), using only the data from the nominal mission. The blue contours and lines are derived from the 2015 baseline likelihood, PlikTT+lowTEB ("PlanckTT+lowP" in the plot), while the red contours and line are obtained from the full PlikTT, EE, TE+lowTEB likelihood ("PlanckTT, TE, EE+lowP" in the plot, see Appendix E.1 for the relevant robustness tests). In most cases the 2015 constraints are in quite good agreement with the earlier constraints, with the exception of the normalization $A_s$, which is higher by about 2%, reflecting the 2015 correction of the *Planck* calibration which was indeed revised upward by about 2% in power. The figure also illustrates the consistency and further tightening of the parameter constraints brought by adding the $E$-mode polarization at high $\ell$. The numerical values of the *Planck* 2015 cosmological parameters for base ΛCDM are given in Table 21.

As shown in Fig. 44, the degeneracies between foreground and calibration parameters generally do not affect the determination of the cosmological parameters. In the PlikTT+lowTEB case (top panel), the dust amplitudes appear to be nearly uncorrelated with the basic ΛCDM parameters. Similarly, the 100 and 217 GHz channel calibration is only relevant for the level of foreground emission. Cosmological parameters are, however, mildly correlated with the point-source and kinetic SZ amplitudes. Correlations are strongest (up to 30%) for the baryon density ($\Omega_b h^2$) and spectral index ($n_s$). We do not show correlations with the *Planck* calibration parameter ($y_P$), which is uncorrelated with all the other parameters except the amplitude of scalar fluctuations ($A_s$). The bottom panel shows the correlation for the PlikTE+lowTEB and PlikEE+lowTEB cases, which do not affect the cosmological parameters, except for 20% correlations in $EE$ between the spectral index ($n_s$) and the dust contamination amplitude in the 100 and 143 GHz maps.

We also display in Fig. 45 the correlations between the foreground parameters and the cosmological parameters in the PlikTT+lowTEB case when exploring classical extensions to the ΛCDM model. While $n_{run}$ seems reasonably insensitive to the foreground parameters, some extensions do exhibit a noticeable correlation, up to 40% in the case of $Y_{He}$ and the point-source level at 143 GHz.

Finally, we note that power spectra and parameters derived from CMB maps obtained by the component-separation methods described in Planck Collaboration IX (2016) are generally consistent with those obtained here, at least when restricted to the $\ell < 2000$ range in $TT$; this is detailed in Sect. E.4.

**Table 21.** Constraints on the basic six-parameter ΛCDM model using *Planck* angular power spectra.

| Parameter | PlanckTT+lowP 68% limits | PlanckTT, TE, EE+lowP 68% limits |
|---|---|---|
| $\Omega_b h^2$ | $0.02222 \pm 0.00023$ | $0.02225 \pm 0.00016$ |
| $\Omega_c h^2$ | $0.1197 \pm 0.0022$ | $0.1198 \pm 0.0015$ |
| $100\theta_{MC}$ | $1.04085 \pm 0.00047$ | $1.04077 \pm 0.00032$ |
| $\tau$ | $0.078 \pm 0.019$ | $0.079 \pm 0.017$ |
| $\ln(10^{10}A_s)$ | $3.089 \pm 0.036$ | $3.094 \pm 0.034$ |
| $n_s$ | $0.9655 \pm 0.0062$ | $0.9645 \pm 0.0049$ |
| $H_0$ | $67.31 \pm 0.96$ | $67.27 \pm 0.66$ |
| $\Omega_\Lambda$ | $0.685 \pm 0.013$ | $0.6844 \pm 0.0091$ |
| $\Omega_m$ | $0.315 \pm 0.013$ | $0.3156 \pm 0.0091$ |
| $\Omega_m h^2$ | $0.1426 \pm 0.0020$ | $0.1427 \pm 0.0014$ |
| $\Omega_m h^3$ | $0.09597 \pm 0.00045$ | $0.09601 \pm 0.00029$ |
| $\sigma_8$ | $0.829 \pm 0.014$ | $0.831 \pm 0.013$ |
| $\sigma_8 \Omega_m^{0.5}$ | $0.466 \pm 0.013$ | $0.4668 \pm 0.0098$ |
| $\sigma_8 \Omega_m^{0.25}$ | $0.621 \pm 0.013$ | $0.623 \pm 0.011$ |
| $z_{re}$ | $9.9^{+1.8}_{-1.6}$ | $10.0^{+1.7}_{-1.5}$ |
| $10^9 A_s$ | $2.198^{+0.076}_{-0.085}$ | $2.207 \pm 0.074$ |
| $10^9 A_s e^{-2\tau}$ | $1.880 \pm 0.014$ | $1.882 \pm 0.012$ |
| Age/Gyr | $13.813 \pm 0.038$ | $13.813 \pm 0.026$ |
| $z_*$ | $1090.09 \pm 0.42$ | $1090.06 \pm 0.30$ |
| $r_*$ | $144.61 \pm 0.49$ | $144.57 \pm 0.32$ |
| $100\theta_*$ | $1.04105 \pm 0.00046$ | $1.04096 \pm 0.00032$ |
| $z_{drag}$ | $1059.57 \pm 0.46$ | $1059.65 \pm 0.31$ |
| $r_{drag}$ | $147.33 \pm 0.49$ | $147.27 \pm 0.31$ |
| $k_D$ | $0.14050 \pm 0.00052$ | $0.14059 \pm 0.00032$ |
| $z_{eq}$ | $3393 \pm 49$ | $3395 \pm 33$ |
| $k_{eq}$ | $0.01035 \pm 0.00015$ | $0.01036 \pm 0.00010$ |
| $100\theta_{s,eq}$ | $0.4502 \pm 0.0047$ | $0.4499 \pm 0.0032$ |
| $f_{2000}^{143}$ | $29.9 \pm 2.9$ | $29.5 \pm 2.7$ |
| $f_{2000}^{143\times217}$ | $32.4 \pm 2.1$ | $32.2 \pm 1.9$ |
| $f_{2000}^{217}$ | $106.0 \pm 2.0$ | $105.8 \pm 1.9$ |

**Notes.** The top group contains constraints on the six primary parameters included directly in the estimation process. The middle group contains constraints on derived parameters. The last group gives a measure of the total foreground amplitude (in $\mu K^2$) at $\ell = 2000$ for the three high-$\ell$ temperature spectra used by the likelihood. These results were obtained using the CAMB code, and are identical to the ones reported in Table 3 in Planck Collaboration XIII (2016).

### 5.4. Overall systematic error budget assessment

The tests presented throughout this paper and its appendices documented our numerous tests of the *Planck* likelihood code and its outputs. Here, we summarize those results and attempt to isolate the dominant sources of systematic uncertainty. This assessment is of course a difficult task. Indeed, all known systematics are normally corrected for, and when relevant, the uncertainty on the correction is included in the error budget and thus in the error bar we report. In that sense, except for a very few cases where we decided to leave a known uncertainty in the data, this section tries to deal with the more difficult task of evaluating the unknown uncertainty!

This section summarizes the contribution of the known systematic uncertainties along with these potential unknown unknowns, specifically highlighting both internal consistency tests based on comparing subsets of the data, along with those using end-to-end instrumental simulations.

#### 5.4.1. Low-$\ell$ budget

The low-$\ell$ likelihood has been validated using both internal consistency tests and simulation-based, tests. Here we summarize only the main result of the analysis, which has been set forth in Sect. 2 above.

A powerful consistency test of the polarization data, described in Sect. 2.4, is derived by rotating some of the likelihood components by $\pi/4$. Specifically, the rotation is applied to the

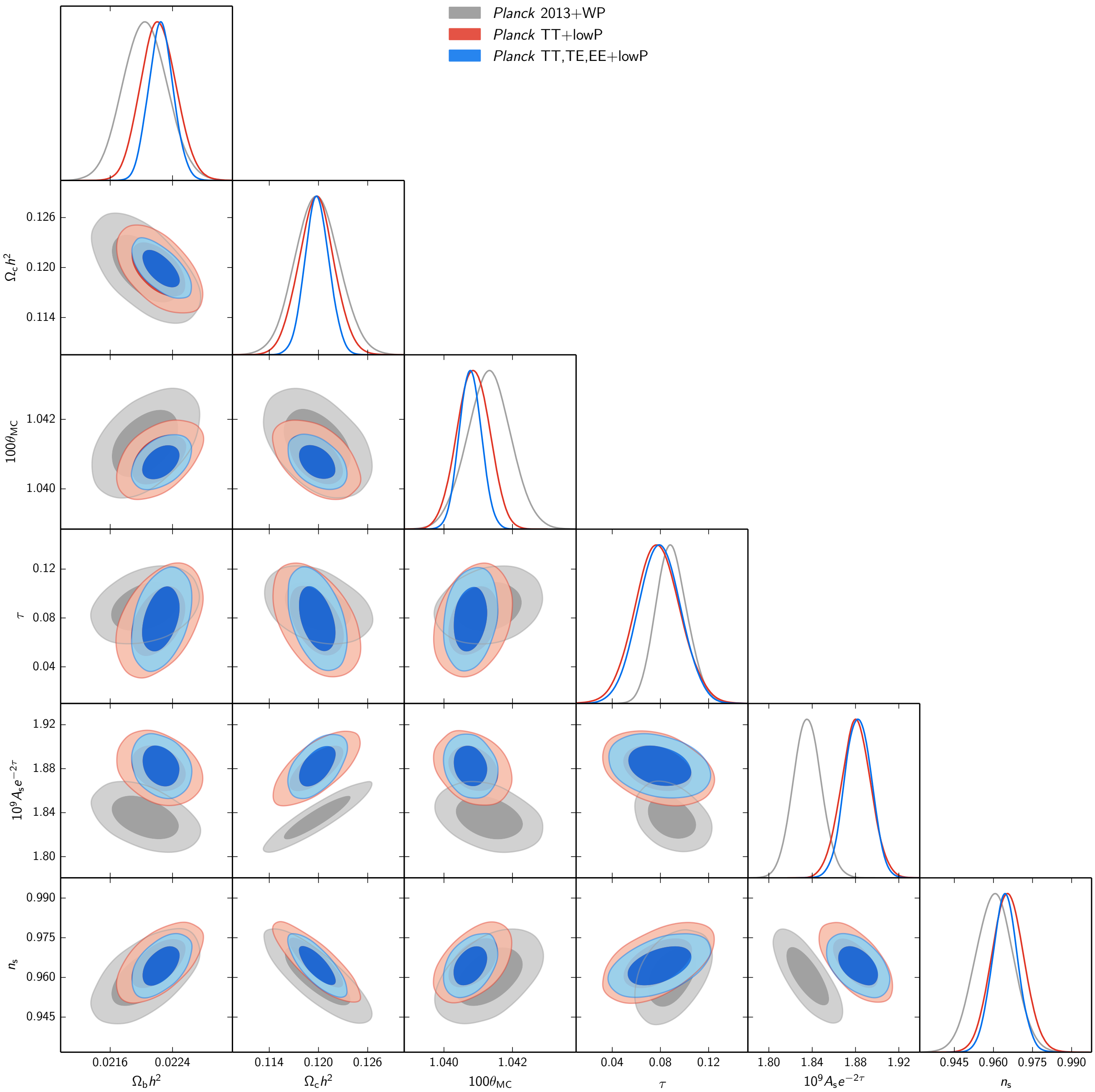


**Fig. 43.** ΛCDM parameter constraints. The grey contours show the 2013 constraints, which can be compared with the current ones, using either *TT* only at high $\ell$ (red) or the full likelihood (blue). Apart from further tightening, the main difference is in the amplitude, $A_s$, due to the overall calibration shift.

data maps and only to the noise covariance matrix (the likelihood being a scalar function, applying the same rotation to the signal matrix as well would be equivalent to not performing the rotation). The net effect is a conversion of $E \rightarrow -B$ and $B \rightarrow E$ for the signal, but leaving unaffected the Gaussian noise described in the covariance matrix. Under these circumstances, we do not expect to pick up any reionization signal, since it would then be present in *BB* or *TB*: the operation should result in a null $\tau$ detection. This is precisely what happens (see the blue dashed curve in Fig. 8). It is of course possible – though unlikely – that systematics are only showing up in the *E* channel, leaving *B* modes unaffected. Indeed, this possibility is further challenged by the fact that we do not detect anomalies in any of the six polarized power spectra; as detailed in Fig. 7, they are consistent with a ΛCDM signal and noise as described by the final 70 GHz covariance matrix.

These tests are specific to *Planck* and aimed at validating the internal consistency of the datasets employed to build the likelihood. As a further measure of consistency, we have carried out a null test employing the WMAP data, detailed in Sect. 2.6. In brief, we have taken WMAP's $K_a$, $Q$ and $V$ channels and cleaned them from any polarized foreground contributions using a technique analogous to the one used to clean the LFI 70 GHz maps, employing the *Planck* 353 GHz map to minimize any dust contribution, but relying on WMAP's K channel to remove any synchrotron contribution. The resulting LFI 70 GHz and WMAP maps separately lead to compatible $\tau$ detections; their half-difference noise estimates are compatible with

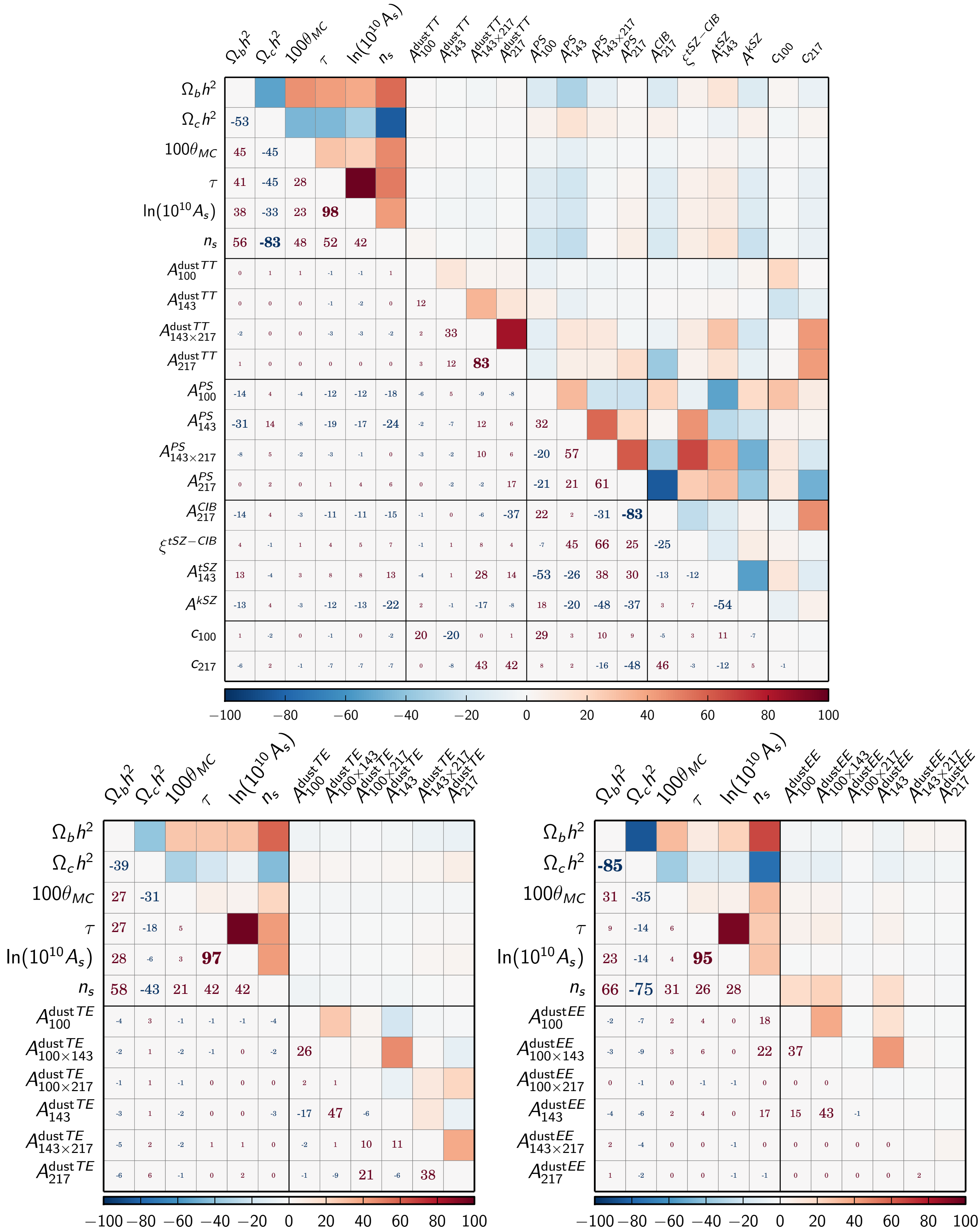


**Fig. 44.** Parameter correlations for `PlikTT`+lowTEB (*top*), `PlikTE`+lowTEB (*bottom left*), and `PlikEE`+lowTEB (*bottom right*). The degeneracies between foreground and calibration parameters do not strongly affect the determination of the cosmological parameters. In these figures the lower triangle gives the numerical values of the correlations in percent (with values below 10% printed at the smallest size), while the upper triangle represents the same values using a colour scale.