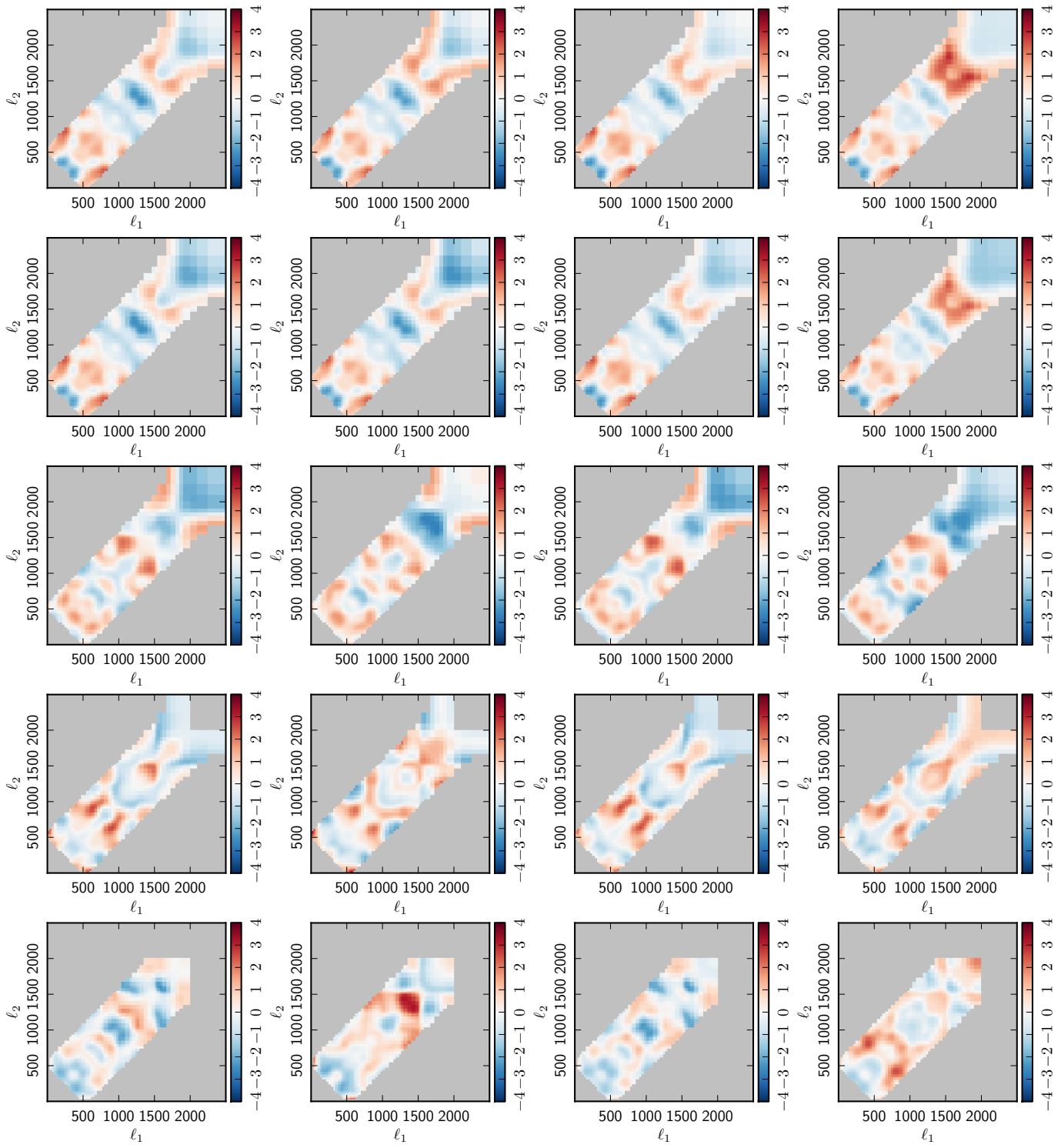

Fig. 8. Smoothed binned signal-to-noise bispectrum $\mathcal{B}$ for the *Planck* 2015 cleaned sky map, as determined with the binned estimator, as a function of ℓ_1 and ℓ_2 for a fixed ℓ_3-bin [518, 548]. *From left to right* results are shown for the four component separation methods `SMICA`, `SEVEM`, `NILC`, and `Commander`. *From top to bottom* are shown: *TTT, TTT* cleaned from radio and CIB point sources; *T2E, TE2*; and *EEE*. The colour range is in signal-to-noise from −4 to +4. The light grey regions are where the bispectrum is not defined, either because it is outside the triangle inequality or because of the cut $\ell^E_{\rm max} = 2000$.

contamination in the cleaned *TTT* bispectra, in agreement with the results of Table 3. However, after removing it we do not see a clear signal of any other residual NG. Of course this is for the moment only a qualitative statement; more quantitative tools for studying the amount of NG in these smoothed bispectra are in development.

Looking at the polarized bispectra in the high-ℓ_3 slices, in particular *TE2* and *EEE*, we do see some bluer and redder regions that might indicate residual NG. This agrees with statements made earlier, and discussed in greater detail in the next section, that the *Planck* polarized bispectrum is for the moment not as clean and well-understood as the temperature one. We also

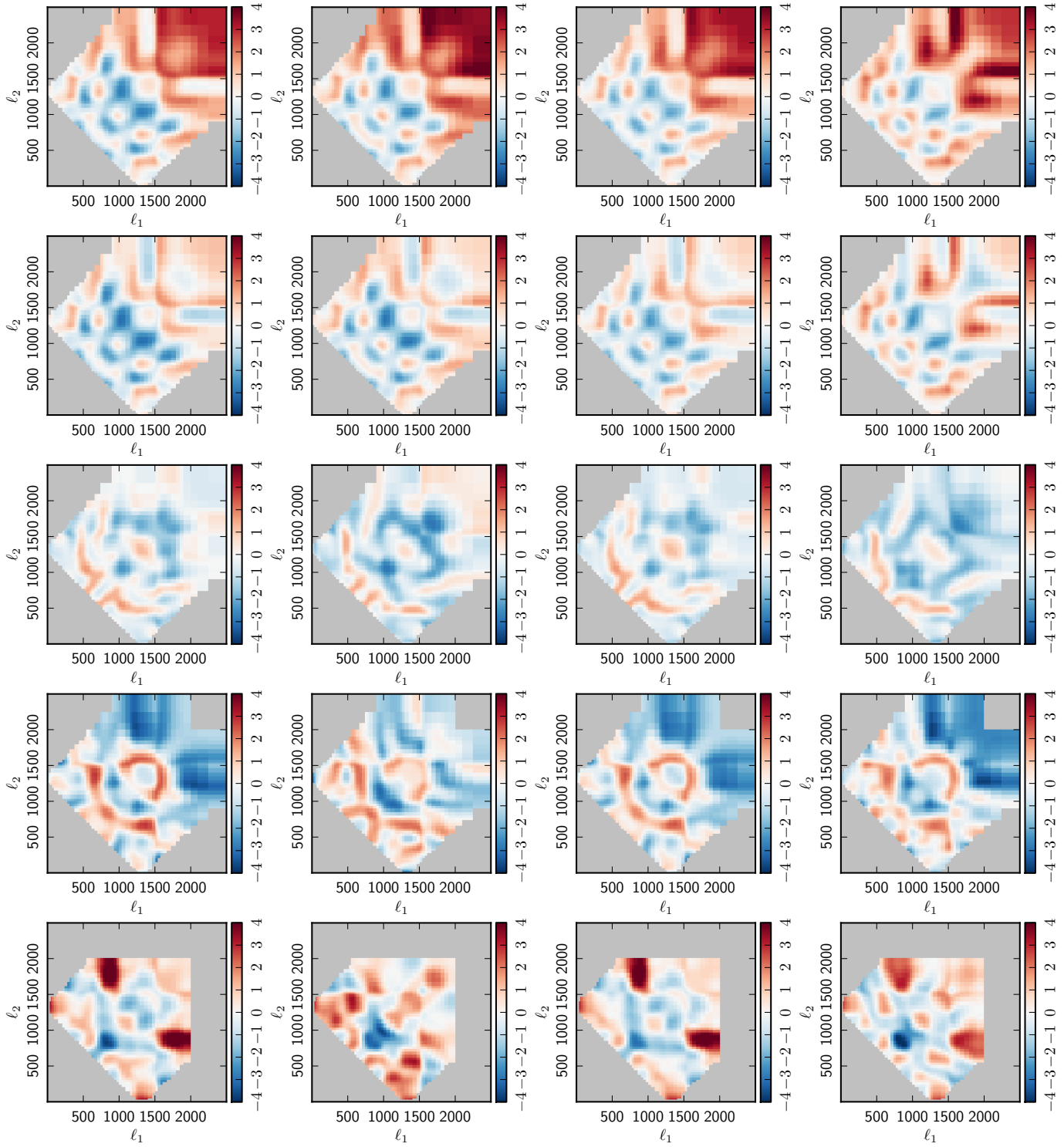

Fig. 9. Similar to Fig. 8, but with $\ell_3 \in [1291, 1345]$.

see a very good qualitative agreement between the four component separation methods in temperature, which worsens somewhat when mixing in more and more polarization; in particular SMICA and NILC give very similar results.

6.3. Primordial curvature reconstruction

In this section, we compress the information in T and E maps into maps of projected primordial curvature fluctuations, ζ. To the extent that the primary CMB temperature and polarization are Gaussian fields and the early universe ζ fluctuations are presumed Gaussian, all of the information in T, E is encoded in the mean field (Wiener-filtered map) plus fluctuations, characterized by a covariance matrix; it is just a different compression/re-expression of the original statistics. Thus, one could look for non-Gaussian deviations by directly evaluating the 3-point and higher statistics of this re-expression. Since fluctuations as well as the mean are given, a full description of the errors follows. That a Gaussian assumption is made does not mean that the Wiener-filtered map is in fact Gaussian, since the constraining

data may drive it from a map consistent with Gaussianity. Hence one can search in the statistics of the Wiener-filtered map for evidence for non-Gaussianity.

The weighting of the temperature and E-mode polarization associated with the Wiener filter is an optimal one (inverse-total-covariance weighted), and estimators for non-Gaussianity constructed using this expression are nearly optimal. However, since a (fiducial) primordial power spectrum is assumed for the Wiener filter (usually a uniform n_s ζ-spectrum, using the best fit *Planck* 2015 scalar spectral index and power amplitude A_s), this approach is best suited for perturbative non-Gaussianity of the sort we treat in this paper. In practice, the estimators used here for non-Gaussianity act directly on temperature and polarization data, rather than through the intermediary of optimal ζ-maps and their fluctuations.

The scalar fluctuations can be expressed in terms of the curvature variable $\zeta(\boldsymbol{x}) = \ln a(\boldsymbol{x})$ on uniform total density hypersurfaces, where a is the inhomogeneous expansion factor, or, equivalently, by its wavenumber transform $\zeta(\boldsymbol{k})$. In turn, $\zeta(\boldsymbol{k})$ can be expanded in multipoles, $\zeta_{LM\kappa}(k)$. Instead of the magnitude of the wavenumber k, a mixed representation gives a multipole expansion at each comoving distance from our location, χ, $\zeta_{LM\kappa}(\chi)$, $M = 0, ..., L$. Here $\kappa = c$ or s, with $\kappa = c$ referring to the real part of ζ_{LM} and $\kappa = s$ to the imaginary part. For $M = 0$ there is no s component, only c. The mean of the ζ field given the temperature field T and its covariance describing allowed fluctuations about that mean are

$$\langle \zeta_{LM\kappa}(\chi)\,|\,a^T_{LM\kappa}\rangle = C^{\zeta T}_L(\chi)\left[C^{TT}_L\right]^{-1} a^T_{LM\kappa} \tag{52}$$

$$\times\,\langle \delta\zeta_{LM\kappa}(\chi)\delta\zeta_{L'M'\kappa'}(\chi')\,|\,T\rangle = \delta_{LL'}\delta_{MM'}\delta_{\kappa\kappa'} \tag{53}$$

$$\times\left\{C^{\zeta\zeta}_L(\chi,\chi') - C^{\zeta T}_L(\chi)\left[C^{TT}_L\right]^{-1} C^{T\zeta}_L(\chi')\right\}.$$

In these expressions, C^{TT} is the total covariance for the temperature, including both signal and noise variances. The specific forms assume the C^{TT} matrix is diagonal in multipole space, depending only on C^{TT}_L, but if the noise is inhomogeneous it will have off-diagonal components and the equations become matrix equations. In this section, we assume the noise is diagonal. Equation (52) shows the unsurprising result that for each $LM\kappa$ only one mode is determined by T. Replacing T by E gives $\langle \zeta_{LM\kappa}(\chi)\,|\,a^E_{LM\kappa}\rangle$ and $\langle \delta\zeta\delta\zeta^\dagger|E\rangle$. Although T and E are correlated, the uncorrelated part of E delivers a different mode for ζ from the one given by T. Thus when ζ is constrained by both T and E, the two modes deliver substantially more information than for T and E alone, and the fluctuations about the mean are thereby diminished. This is further helped by the acoustic oscillations of polarization being out of phase with those for T, so when T is down, E is not, and vice versa. The interplay of the two modes is quantified by signal-to-noise in Fig. 10.

When both T and E are included as constraints, a two by two matrix appears which includes the C^{TE} correlation in the off-diagonal as well as the C^{TT} and C^{EE} along the diagonal:

$$\langle \zeta_{LM\kappa}(\chi)\,|\,a^T_{LM\kappa}, a^E_{LM\kappa}\rangle = \left[C^{\zeta T}_L(\chi)\quad C^{\zeta E}_L(\chi)\right]^\dagger \begin{pmatrix} C^{TT}_L & C^{TE}_L \\ C^{ET}_L & C^{EE}_L \end{pmatrix}^{-1} \begin{bmatrix} a^T_{LM\kappa} \\ a^E_{LM\kappa} \end{bmatrix}. \tag{54}$$

$$\langle \delta\zeta_{LM\kappa}(\chi)\delta\zeta_{L'M'\kappa'}(\chi')\,|\,T,E\rangle = \delta_{LL'}\delta_{MM'}\delta_{\kappa\kappa'}\left\{C^{\zeta\zeta}_L(\chi,\chi') - \left[C^{\zeta T}_L(\chi)\quad C^{\zeta E}_L(\chi)\right]^\dagger \begin{pmatrix} C^{TT}_L & C^{TE}_L \\ C^{ET}_L & C^{EE}_L \end{pmatrix}^{-1} \begin{bmatrix} C^{T\zeta}_L(\chi') \\ C^{E\zeta}_L(\chi') \end{bmatrix}\right\}. \tag{55}$$

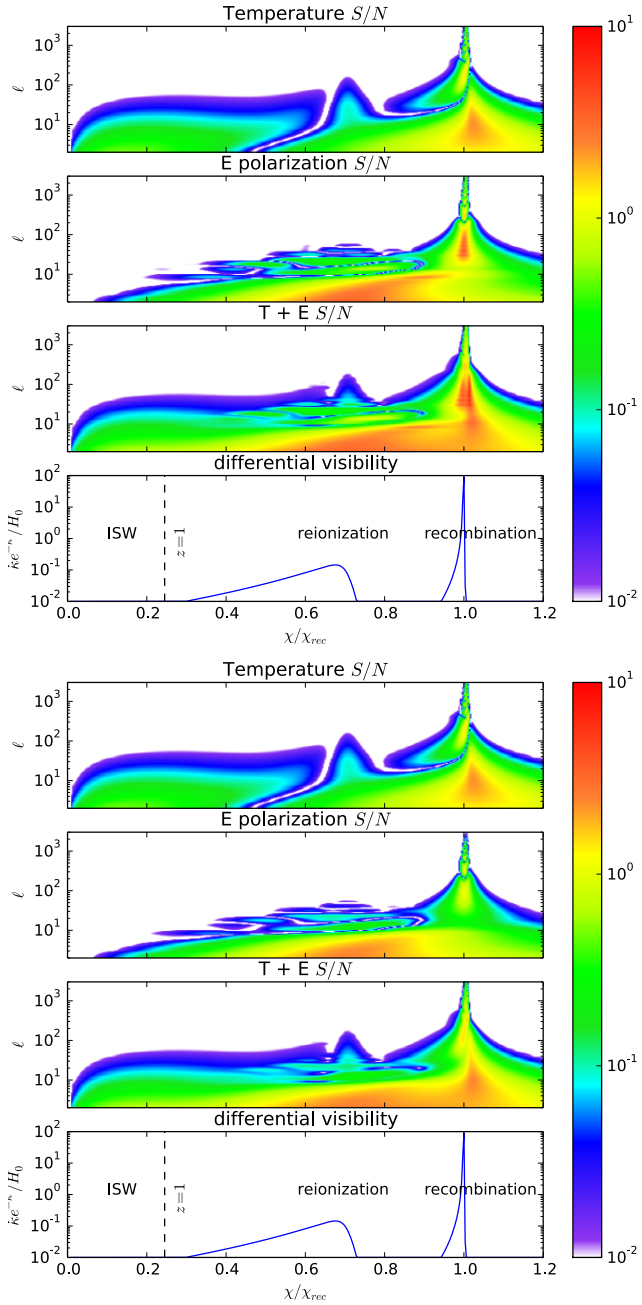

Fig. 10. Signal-to-noise for T and E as a function of multipole L and comoving distance χ for an ideal cosmic-variance limited experiment (*upper panels*) and for the Planck experiment, with noise determined from FFP8 simulations (*lower panels*). The complementary nature of the information provided by T and E is evident. The differential visibility is shown for comparison.

Figure 11 shows an all-sky mean field (Wiener-filtered) map of the curvature variable ζ constructed in (densely-packed) shells from the multipoles $\zeta_{LM\kappa}(\chi)$. The figure actually shows ζ projected onto differential visibility, $w = \mathrm{d}e^{-\tau}/\mathrm{d}\chi$. The map $\zeta_w(\theta,\phi)$ is the spherical transform of $\zeta_{LM\kappa w} = \int w(\chi)\mathrm{d}\chi\zeta_{LM\kappa}(\chi)$. The top left map is $\langle\zeta_w|T\rangle$ using the all-sky maps SMICA DX11. The middle left map is $\langle\zeta_w|E\rangle$, based on just the polarization information. (The E-maps did not have high pass filtering imposed, so that an indication of the full results obtainable with E alone can be seen.) The bottom left panel shows $\langle\zeta_w|T,E\rangle$. Visually the combined two-mode constraint has more detail than either

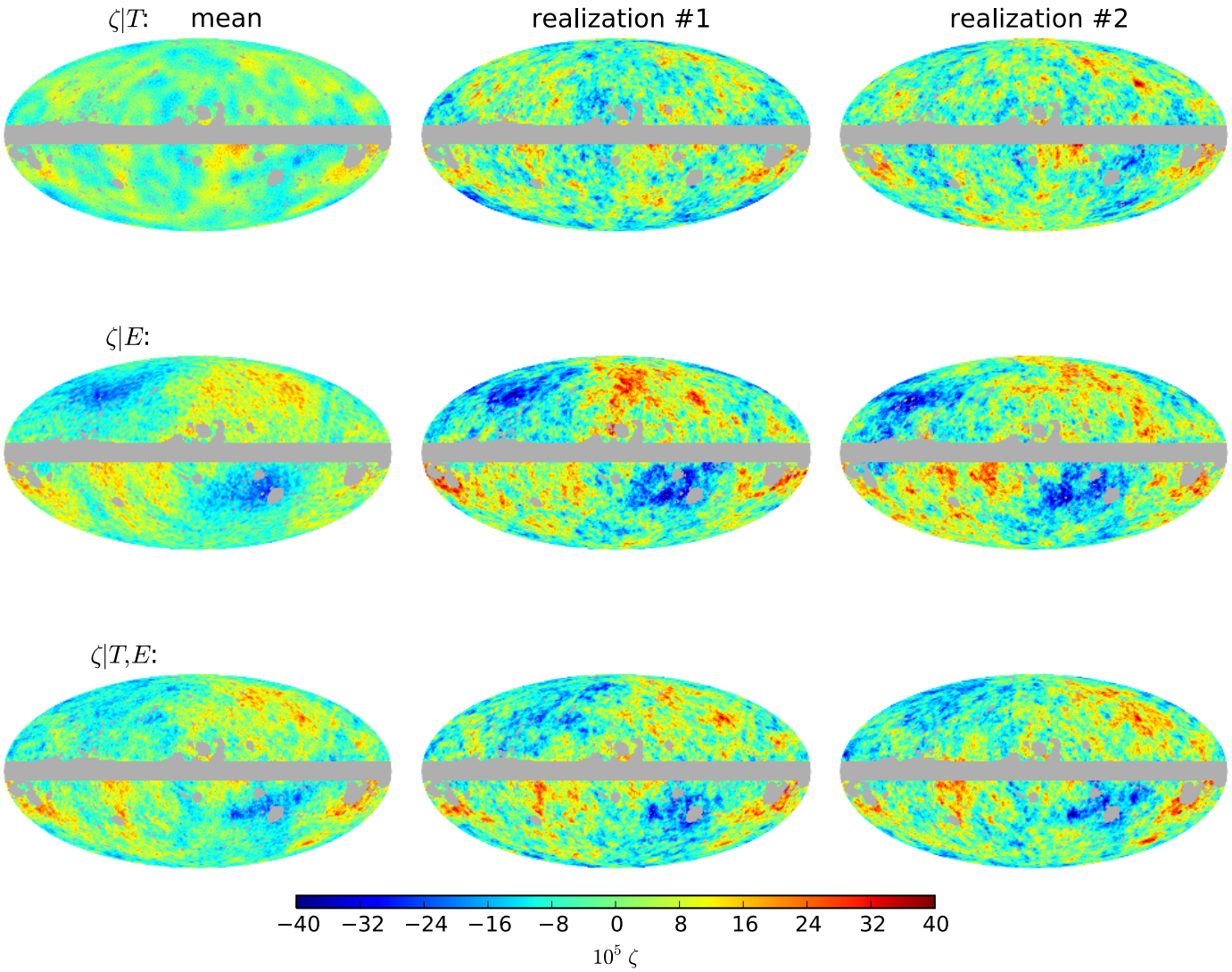

Fig. 11. Mean field and fluctuations for differential visibility projected ζ, as described in the text. The filter used in these maps is a 40′ FWHM Gaussian. The grey regions are masked out.

of the one-mode constrained maps, enhancing the estimation of ζ_w. In each of the three cases, two sample realizations are shown which include allowed fluctuations about the mean map. These illustrate the level of uncertainty in the patterns found in the mean maps, necessary ingredients to the full statistical description. Note that the fluctuations about the mean are larger for T or E alone than for the combined T, E.

In all cases, the fluctuations become very large indeed if we allow for arbitrary projections, not conditioned by the visibility. Another approach similar to the differential visibility projection is to make a slice at $\chi = \chi_{\rm dec}$; i.e., making the projection a delta-function in χ at the peak where $\mathrm{d}w/\mathrm{dln}\,\chi = 0$, which defines $\chi_{\rm dec}$. This single slice case is shown in Fig. 12. The figures look quite similar in structure to the projected differential visibility plots. Although $\chi \sim \chi_{\rm dec}$ is where most information is, the fluctuations of a single slice are somewhat larger than for the differential visibility projection. The differential visibility includes a subdominant "reionization bump", with a small weight focused at low L where sample variance is large; restricting the projection of ζ to be just over the reionization region results in a low L mean map that is largely swamped by the allowed fluctuations about it. The fluctuations about the mean become much larger if we project over all χ, since there are vast terrains in χ in which there is very little CMB temperature or polarization information; in particular between recombination and reionization. For that uncharted territory, a realization of the reconstructed ζ-map reverts to a realization of the fiducial ζ-power.

Figure 10 quantifies where the information resides in L–χ space. Here, the signal variance is the ensemble average of the square of the mean field, assuming the ζ are drawn from a Gaussian with fiducial covariance,

$$\langle\langle\zeta_{LM\kappa}(\chi)\,|\,T\rangle\langle\zeta_{LM\kappa}(\chi')\,|\,T\rangle\rangle = C_L^{\zeta T}(\chi)\left[C_L^{TT}\right]^{-1} C_L^{T\zeta}(\chi'), \qquad (56)$$

and the noise (i.e., fluctuation) variance is the covariance matrix, as given by Eq. (54) for T and the equivalent for E. Figure 10 plots the signal-to-noise at each individual χ slice (hence ignores the components off-diagonal in χ):

$$[S/N]_L^T(\chi) = \frac{\rho}{\sqrt{(1-\rho^2)}}\,;\ \rho \equiv \frac{C_L^{\zeta T}(\chi)}{\sqrt{C_L^{TT} C_L^{\zeta\zeta}(\chi,\chi)}} \qquad (57)$$

for T, with an equivalent expression for E. The signal-to-noise structure in χ can be contrasted with the differential visibility

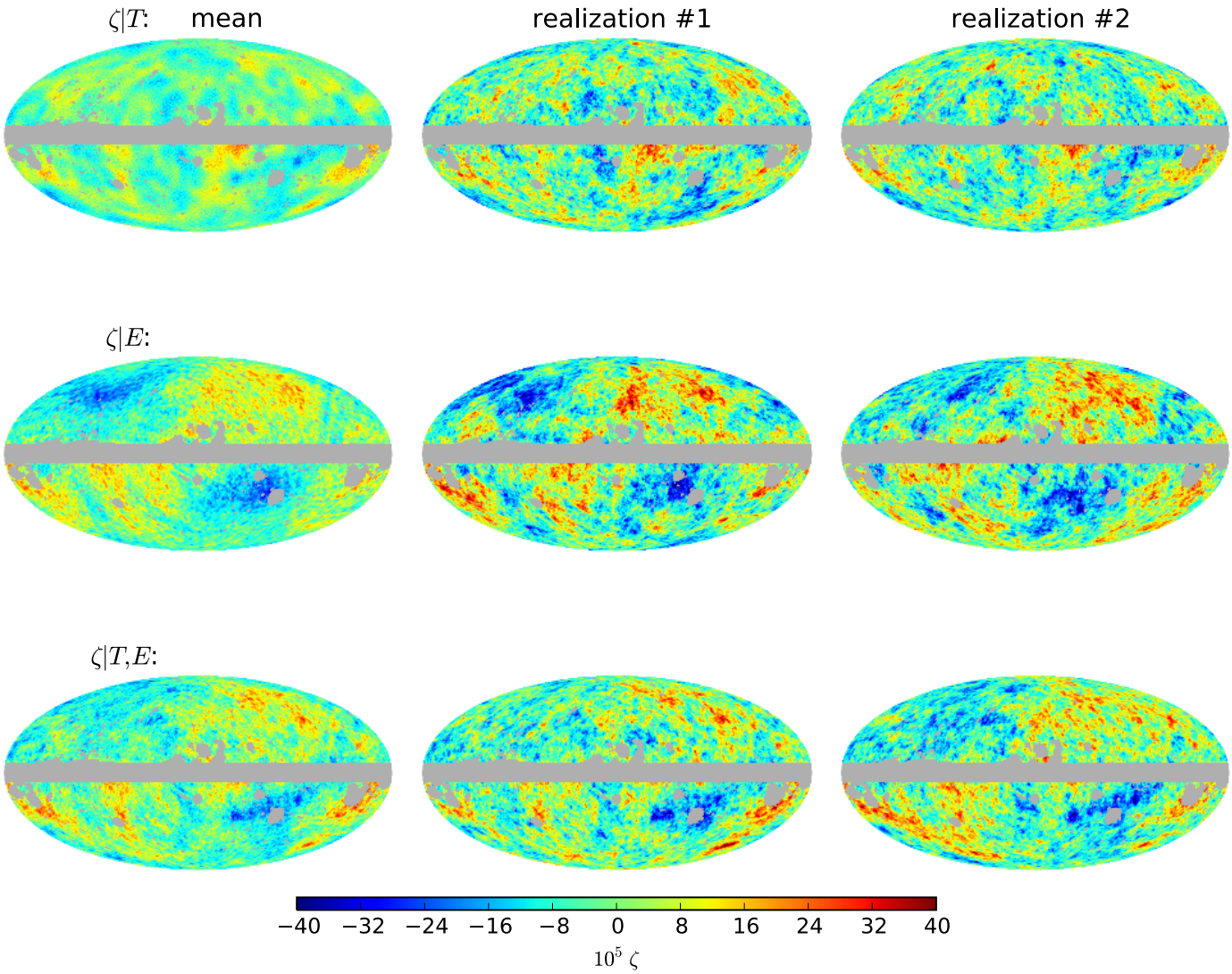

Fig. 12. Mean field and fluctuations for a $\chi_{\rm dec}$ slice of ζ, as described in the text. The filter used in these maps is a 40′ FWHM Gaussian. The grey regions are masked out.

plotted in the bottom panel: Fig. 10 shows that the L–χ information has a reach beyond the differential visibility structure, especially for low multipoles, because the associated waves can straddle the last scattering surface. The "reionization bump" in signal-to-noise seen in Fig. 10 is not that prominent. The Integrated Sachs Wolfe impact on the S/N is evident, but is relatively low, with the consequence that we cannot draw out high significance results from the ISW effect for ζ reconstruction (or equivalently for gravitational potential reconstruction).

The top panels in Fig. 10 are for an ideal experiment, with no noise (apart from the cosmic variance "noise" in the Wiener map fluctuations). The bottom panels of Fig. 10 include realistic *Planck* noise, as estimated from the FFP8 simulations: the high signal to noise in E in the top panels is noticeably diminished around $L \sim 100$ over what a cosmic variance limited experiment would give.

Figures 11 and 12 are Gaussian-filtered on a relatively large $L \sim 200$ scale (40′ FWHM). To see what happens at higher resolution, Figs. 13 and 14 zoom in on a typical 20° × 20° patch, with long waves removed using a filter $W_L = \sin^2(L-L_{\rm c})/\Delta L$ for $L_{\rm c} < L < L_{\rm c}+\Delta L$. The specific choices are $L_{\rm c} = 20$ and $\Delta L = 20$, the values used for the *Planck* high pass for polarization maps. To allow direct comparison, we have just done the same high pass filtering for the T map. We also removed the means of the maps. The resolution is $L \sim 400$ (20′ FWHM). Figures 13 and 14 illustrate that the fluctuations play a larger role in this higher resolution regime. Note that the $\chi_{\rm dec}$ slice has about the same fluctuation level as the differential visibility projection.

Lensing effects are not taken into account in these ζ-maps. In principle one could de-lens the temperature and polarization maps before forming the Wiener filter. In practice this would be a highly noisy operation with current *Planck* data. Hence the contaminating influence of lensing on these ζ-reconstructions would be treated by comparing simulated ζ-maps with and without lensing. Such corrections are expected to be a subdominant bias, as it has been modelled in the rest of the paper.

7. Validation of *Planck* results

In this section, we perform a battery of tests aimed at verifying the robustness of the results obtained in the previous section. Table 10 shows excellent agreement with our 2013 analysis of nominal mission data (Planck Collaboration XXIV 2014). The agreement using different component separation methods in

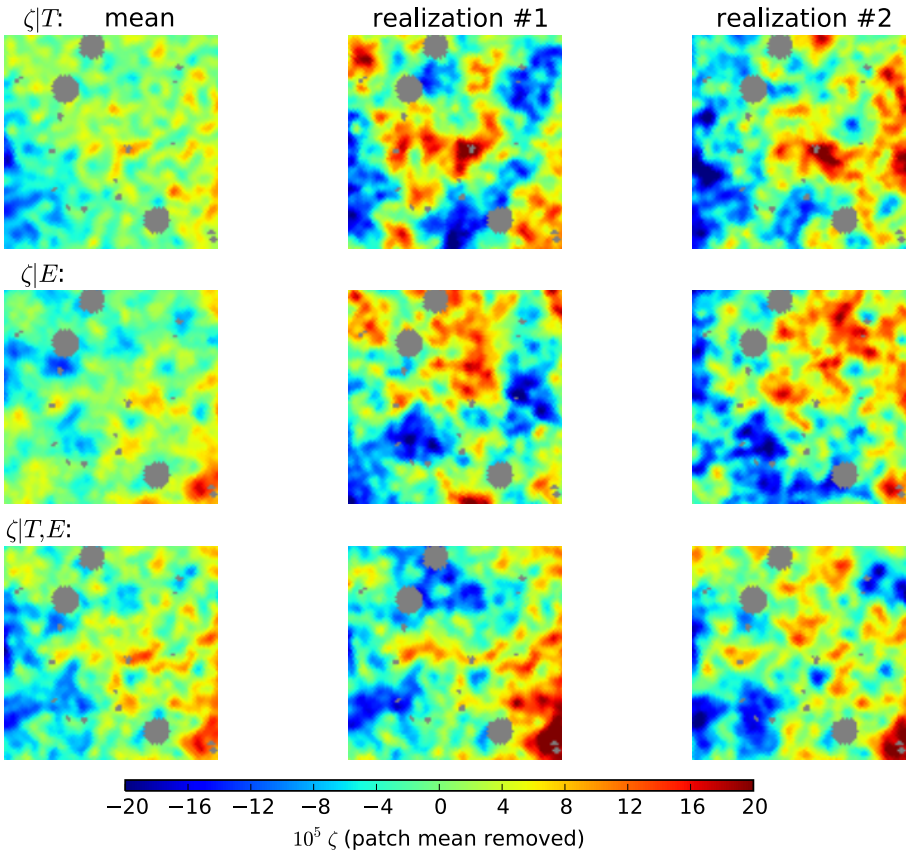

Fig. 13. Mean field and fluctuations for differential visibility projected ζ, for a 20 deg by 20 deg patch, as described in the text. The filter used in these maps is a 20′ FWHM Gaussian. Sources in grey are masked out.

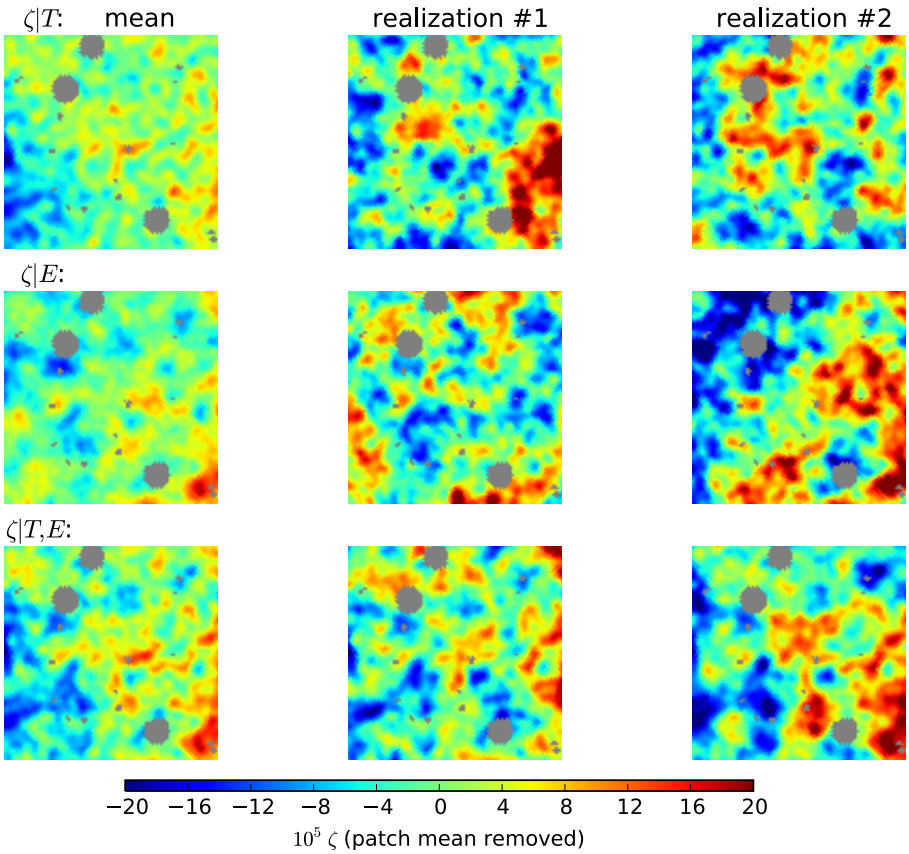

Fig. 14. Mean field and fluctuations for a $\chi_{\rm dec}$ slice of ζ, for a 20 deg by 20 deg patch, as described in the text. The filter used in these maps is a 20′ FWHM Gaussian. Sources in grey are masked out.

temperature is also generally very good. Our focus here is thus on polarization bispectra. Redundancy is perhaps the most important element in our analysis, as far as robustness is concerned. We devote considerable attention to comparing the outcomes of different estimators and component separation pipelines, and assess their level of internal consistency. We also verify the stability of our results in the harmonic and pixel domains, by considering different sky cuts and multipole intervals. Given the large computational requirements of these tests, and since results from different optimal estimators agree very well, as shown in previous sections, we will variously use the KSW, binned, or modal pipeline for different tests. In doing this we will also exploit the complementarity of different decompositions, which might make some of them more suited for different tests than others (for example, the binned pipeline directly works with a harmonic space decomposition of the bispectrum, thus making it perfectly suited for tests of ℓ-dependence, the modal pipeline can perform quick model-independent tests by working on a relatively small subset of bispectrum modes, and so on).

7.1. Dependence on foreground cleaning method

7.1.1. Comparison between $f_{\rm NL}$ measurements

In Table 10 we show $f_{\rm NL}$ results for the local, equilateral and orthogonal shapes, using four different optimal estimators, and four different foreground cleaning pipelines. The agreement between different estimators, on a given map, is within a fraction of a standard deviation, in line with theoretical expectations and simulations, as reported in Sect. 5. This level of agreement applies to all of the T+E, (TTT), and (EEE) bispectra.

The overall picture becomes more complex when comparing outputs across different foreground cleaning methods and estimators. Whereas for TTT and T+E results the agreement also seems quite good in this case (being at the level of half a sigma or better, for nearly all combinations of cleaned maps and shapes), larger discrepancies are present in the EEE bispectrum measurements. The most notable differences are found for the equilateral shape, where SMICA and NILC find values of $f_{\rm NL}$ consistent with 0 within 1σ, while SEVEM and Commander measure a roughly 2σ deviation from Gaussianity. The largest discrepancy is found for the pair Commander–NILC, using the binned pipeline (see Table 10). This estimator recovers $f_{\rm NL}^{\rm equil} = 369 \pm 160$ using the Commander E-map and $f_{\rm NL}^{\rm equil} = 97 \pm 141$ using the NILC E-map. Other pipelines, and different choices of component separation methods, show slightly smaller but similar discrepancies, at a level of about 1.5σ. The same shape and estimator analysis of temperature maps shows good agreement: Commander recovers $f_{\rm NL}^{\rm equil} = -36 \pm 73$, while NILC gives $f_{\rm NL}^{\rm equil} = -45 \pm 71$. The combined T+E measurement, still for the same modal pipeline and equilateral shape, yields $f_{\rm NL}^{\rm equil} = 26 \pm 50$ for Commander and $f_{\rm NL}^{\rm equil} = -4 \pm 46$ for NILC, corresponding to about half a standard deviation difference. This general trend is seen for other shapes and estimators.

Simulations were used to give insight into the expected level of disagreement. For each of the four component separation methods, we generated 100 FFP8-based Gaussian simulations with realistic beams and noise levels. These simulations start from the same initial single frequency realizations, and are processed through the four different foreground cleaning pipelines. The starting maps *do not* include any foreground component (the same map generation procedure is used in the Monte Carlo determination of error bars). The differences in final simulations are thus generated only by the different data filtering and coadding operations performed either in pixel, harmonic, or needlet domains by the various foreground cleaning methods, and by additional manipulations of the maps which are required for $f_{\rm NL}$ estimation, such as inpainting. Therefore, the average scattering in $f_{\rm NL}$, measured from these realizations, provides us with a baseline assessment of the expected discrepancies between different methods when foreground residuals and other spurious sources of NG are negligible. We can then compare them with differences observed on the data to establish whether they are consistent with expectations, or are too large. The latter would raise the concern that foreground contamination, or other systematics, might be affecting the results.

Results are shown in Table 12, for EEE and T+E and two different sky coverages. The scatter between $f_{\rm NL}$ values from simulations is about 0.5σ for both T+E and EEE. This is smaller than the differences in the *Planck* $f_{\rm NL}$ values obtained from EEE analysis of different foreground-cleaned maps, especially for the equilateral shape. However, for the final combined T+E measurement, observed differences are in good agreement with expectations from simulations for the majority of cases. Another important point is that the consistency shown in Table 12 for T+E measurements is stable to the change of sky coverage (in polarization) from $f_{\rm sky} = 0.74$ to $f_{\rm sky} = 0.64$. This will be confirmed by additional tests later in this section. For the SMICA–SEVEM pair we also verified stability using an even larger mask with $f_{\rm sky} = 0.52$.

Residual foregrounds may be responsible for at least some of the observed excess of scatter in EEE-derived $f_{\rm NL}$ between different cleaning algorithms. This is supported by the fact that several EEE results for this test change significantly for different masks, and that discrepancies are alleviated by using a larger mask, especially for equilateral shapes (see e.g., SEVEM–NILC and SMICA–SEVEM in Table 12). However, modal coefficients and their correlations are stable to a change of mask (see below), as are values of $f_{\rm NL}$ for a given component separation method (see Table 13).

Another possible contributor is a mismatch between the noise model (used to build the estimator normalization, weights and linear term), and the actual noise in the data. Polarization data are very noisy, and it is a known problem that the model assumed underestimates the true noise. This means that the error bars for EEE $f_{\rm NL}$ results, quoted in Table 10, are somewhat underestimated, which does not seem to be a problem for the final T+E results, since the weight of the EEE bispectrum in the final combined measurement is very low. This is confirmed by the results of this test. Indeed, we investigate EEE in detail because it is a useful and sensitive indicator of various systematics in the polarized maps (which could eventually leak into the TTE and TEE bispectra), rather than for its statistical weight in the final measurement. It is then fair to say that issues in the EEE bispectra, and related $f_{\rm NL}$ measurements are not yet fully understood and will require further investigation. Even though the T+E are consistent, we recommend that results that include polarization data are regarded as *preliminary* at this stage.

7.1.2. Comparison between reconstructed bispectra

It is important to stress that the conclusions reached at the end of the previous subsection refer to the three main bispectra in our analysis, defined by the standard scale-invariant local, equilateral, and orthogonal primordial shapes. These shapes select a specific subset of configurations in the overall bispectrum domain (essentially squeezed, equilateral, and flattened triangles).

Table 12. Comparison between local, equilateral, and orthogonal $f_{\rm NL}$ results, obtained using the four different component separation pipelines.

	$f_{\rm NL}$ (method$_1$) − $f_{\rm NL}$ (method$_2$)					
		$f_{\rm sky}$ = 0.74			$f_{\rm sky}$ = 0.64	
Methods	Local	Equilateral	Orthogonal	Local	Equilateral	Orthogonal
SMICA–SEVEM						
T	−1.2 ± 0.9	−6.0 ± 8.7	1.5 ± 4.8	. . .	. . .	. . .
E	−19 ± 21	−155 ± 86	34 ± 57	5 ± 22	−82 ± 90	−11 ± 66
T+E	−2.4 ± 1.6	−10 ± 18	13.5 ± 9.4	−1.5 ± 1.7	−12 ± 18	13 ± 10
SMICA–NILC						
T	0.4 ± 1.0	14.5 ± 8.9	2.5 ± 4.7	. . .	. . .	. . .
E	26 ± 11	83 ± 52	−59 ± 27	26 ± 13	32 ± 56	−96 ± 28
T+E	−0.7 ± 0.9	20.0 ± 8.2	−3.3 ± 3.8	0.6 ± 0.9	18.4 ± 8.4	−4.5 ± 4.0
SMICA–Commander						
T	0.4 ± 3.5	−14 ± 23	1.7 ± 14	. . .	. . .	. . .
E	−3 ± 16	−130 ± 77	−81 ± 42	−13 ± 17	−117 ± 100	−59 ± 40
T+E	−1.3 ± 3.2	−25 ± 18	9 ± 10	−1.4 ± 3.3	−26 ± 18	13 ± 10
SEVEM–NILC						
T	1.6 ± 1.0	20 ± 12	1.0 ± 4.5	. . .	. . .	. . .
E	45 ± 26	239 ± 94	−94 ± 69	30 ± 29	114 ± 105	−86 ± 79
T+E	3.1 ± 1.8	30 ± 18	−17 ± 10	2.2 ± 1.9	30 ± 18	−18 ± 10
SEVEM–Commander						
T	1.6 ± 3.4	−8 ± 22	0 ± 14	. . .	. . .	. . .
E	16 ± 22	25 ± 112	−116 ± 59	−18 ± 25	−35 ± 121	−48 ± 64
T+E	1.2 ± 3.3	−14 ± 21	−5 ± 11	0.2 ± 3.4	−14 ± 20	0 ± 11
NILC–Commander						
T	0.0 ± 3.0	−28 ± 22	−1 ± 12	. . .	. . .	. . .
E	−29 ± 21	−213 ± 84	−22 ± 54	−39 ± 23	−149 ± 108	38 ± 55
T+E	−1.9 ± 3.1	−45 ± 18	12 ± 11	−2.0 ± 3.2	−44 ± 17	18 ± 11

Notes. For each pair of cleaning methods, and for each NG model, we compute the difference in the measured $f_{\rm NL}$. The quoted error bar is the standard deviation of the same difference, extracted from a set of 100 realistic Gaussian simulations per method, not including foregrounds. These results have been obtained using the low resolution modal pipeline. See main text for comments and further details.

Therefore, testing consistency between methods for these shapes does not guarantee that results for the many other NG models considered in this work (such as, for example, the oscillatory bispectra of Sect. 8) will display the same level of agreement. For this reason we decided to perform a model-independent test of consistency between methods, based on comparisons between the β_n eigenmodes used for bispectrum reconstruction in Sect. 6.2. We also reconstruct the bispectrum starting from a binned ℓ-decomposition, and this will be used in Sect. 7.4 to study stability of the results in the harmonic domain. For the β_n study we consider a simple test based on measuring the correlation coefficient between modes extracted from different foreground-cleaned maps. The correlation is defined, as usual, by

$$r_{ij}^2 = \frac{\mathrm{cov}\left(\beta_n^i, \beta_n^j\right)^2}{(\sigma_n^2)_i (\sigma_n^2)_j}, \tag{58}$$

and we measure it for each combination of the SMICA, SEVEM, NILC, and Commander maps, labelled by the indices i, j. Results are given in Tables 14 and 15 for the two modal pipelines and are illustrated in Fig. 15. These results show an excellent degree of correlation between different maps in temperature (especially for SMICA, SEVEM, and NILC), which is reduced when polarization is considered. In fact the correlation for polarization is not much lower than temperature for SMICA and NILC, while it reduces the correlation for the pairs SMICA–SEVEM, and NILC–SEVEM, and for Commander when paired with any other method. This is consistent with previous findings of our $f_{\rm NL}$-based test. To test if these results are due to foreground residuals (or other effects that are not included in the simulations), we evaluate the same mode-mode correlations on the same sets of 100 realistic, foreground-free, Gaussian simulations as previously used, and processed through each of the different component separation pipelines. For this analysis we consider *TTT* and *EEE* bispectra, expanded via the low-resolution modal estimator. Our results are reported in Table 14, in the simulation column, and they clearly show that the trend in the simulations is consistent with what we see in the *Planck* data. In particular, *EEE* results show a lower degree of correlation in simulated maps, for the same pairs of methods. The observed loss of correlation in polarization does not seem to come from unresolved foregrounds or other unaccounted systematics, but rather something intrinsic to the foreground-removal algorithms. They are substantially different, as SMICA and NILC both perform the cleaning in harmonic space, at the level of E and B multipoles, whereas SEVEM is essentially a pixel-space template fitting method, performing the subtraction on Q and U maps, or inpainted before $f_{\rm NL}$ estimation. These issues will be studied in greater detail in future work, using Wiener-filtered, as well as inpainted maps for $f_{\rm NL}$ estimation. However, we have already seen that the larger scatter between modes from different foreground cleaning methods does not have a serious impact on $f_{\rm NL}$ estimation, at least for the standard local, equilateral, and orthogonal shapes. The non-standard

Table 13. For each of the four foreground cleaned maps, we compute $f_{\rm NL}$ for the local, equilateral, and orthogonal modes using two different polarization masks, one with $f_{\rm sky} = 0.74$ and the other with $f_{\rm sky} = 0.64$, while for temperature we use a single mask with $f_{\rm sky} = 0.76$.

	$f_{\rm sky} = 0.74$			$f_{\rm sky} = 0.64$			Difference		
	Local	Equilateral	Orthogonal	Local	Equilateral	Orthogonal	Local	Equilateral	Orthogonal
SMICA									
T	6.8 ± 5.4	-17 ± 66	-48 ± 33	...	...	...	...	...	...
E	25 ± 30	147 ± 159	-137 ± 73	48 ± 31	220 ± 168	-180 ± 81	-23 ± 16	-73 ± 68	43 ± 34
T+E ...	4.0 ± 4.8	5 ± 46	-30 ± 21	4.6 ± 5.2	19 ± 55	-37 ± 22	-0.7 ± 1.2	-14 ± 14	6.7 ± 7.7
SEVEM									
T	8.1 ± 5.8	-11 ± 75	-49 ± 34	...	...	...	...	...	...
E	44 ± 38	302 ± 183	-172 ± 91	43 ± 39	303 ± 191	-170 ± 96	1 ± 19	0 ± 76	-2 ± 49
T+E ...	6.4 ± 5.0	15 ± 52	-44 ± 23	6.2 ± 5.3	31 ± 54	-50 ± 25	0.2 ± 1.3	-16 ± 15	6.3 ± 8.8
NILC									
T	6.4 ± 5.8	-31 ± 76	-50 ± 33	...	...	...	...	...	...
E	-1 ± 30	64 ± 162	-78 ± 77	22 ± 30	190 ± 162	-84 ± 77	-23 ± 16	-124 ± 67	6 ± 37
T+E ...	3.3 ± 4.9	-15 ± 50	-27 ± 23	4.0 ± 5.3	1 ± 56	-33 ± 23	-0.7 ± 1.3	-16 ± 13	5.4 ± 7.5
Commander									
T	6.4 ± 6.6	-3 ± 77	-49 ± 36	...	...	...	...	...	...
E	28 ± 37	278 ± 178	-56 ± 81	61 ± 38	337 ± 188	-122 ± 91	-32 ± 20	-60 ± 92	66 ± 47
T+E ...	5.2 ± 5.4	30 ± 50	-39 ± 23	6.0 ± 5.7	45 ± 55	-51 ± 25	-0.7 ± 1.5	-15 ± 14	11.5 ± 8.9

Notes. We then calculate the difference between the two measurements and compare with expectations from simulations, obtained in the following way: firstly, we generate realistic Gaussian realizations for each component separation pipeline, not including foregrounds; then, for each simulated map and NG model, we measure $f_{\rm NL}$ using the two masks in turn; finally, we calculate the standard deviation on 100 realizations. See the main text for more details and a discussion of these results, which were obtained using the low resolution modal pipeline.

Table 14. Correlation coefficients between pairs of bispectrum modes, extracted using two different component-separated maps.

	$f_{\rm sky} = 0.74$				$f_{\rm sky} = 0.64$	
	TTT		*EEE*		*EEE*	
Methods	Data	Simul.	Data	Simul.	Data	Simul.
SMICA–SEVEM	0.97	0.97	0.61	0.62	0.60	0.61
SMICA–NILC	0.97	0.97	0.95	0.95	0.95	0.95
SMICA–Commander ...	0.78	0.81	0.70	0.70	0.73	0.73
SEVEM–NILC	0.96	0.97	0.54	0.55	0.54	0.54
SEVEM–Commander ...	0.81	0.83	0.69	0.67	0.70	0.70
NILC–Commander	0.85	0.86	0.64	0.63	0.66	0.66

Notes. For both *TTT* and *EEE* we compare correlations measured from data with averages over 100 Gaussian realizations. The simulations are processed through the different component separation pipelines in the same way as the data, but they do *not* include any foregrounds. The correlation is clearly lower for *EEE* bispectra than for *TTT*. However, this is seen not only in data but also in simulations, indicating that it is not due to foreground residual contamination or other unaccounted for systematics. The results presented in this table are obtained using the low resolution modal pipeline, with 610 modes; results on data have also been cross-checked with the high-resolution modal estimator, using 2001 modes, and they are stable.

shapes need to be analysed separately to check robustness of NG polarization results. This is the approach we will take for the various non-standard NG models.

7.2. Dependence on sky coverage

For each of the four component separation methods, we have used two different polarization masks, namely the same polarization mask as in Sect. 6, with $f_{\rm sky} = 0.74$ (defined as the polarization "common mask" in Sect. 3.4), and an extended mask with $f_{\rm sky} = 0.64$. The temperature mask is kept unchanged in this test, and it covers a sky fraction $f_{\rm sky} = 0.76$ (the temperature "common mask" of Sect. 3.4). We report the variation in $f_{\rm NL}$ for the three standard shapes in Table 13, which shows insensitivity to $f_{\rm sky}$, in agreement with earlier results on T+E. In this case, however, the *EEE* results also seem quite stable, supporting the view that foreground residuals are not affecting our local, equilateral, and orthogonal $f_{\rm NL}$ results, especially for the final, combined T+E measurements. Tests on FFP8 simulations including foregrounds (see Sect. 7.3) suggest that $f_{\rm NL}$ measurements obtained from the SMICA and SEVEM maps are the most accurate under the current choice of mask. As a further check of these two methods we consider a third polarization mask, with $f_{\rm sky} = 0.53$, and repeat the combined T+E $f_{\rm NL}$ measurement, also finding stable results. For SMICA we find $f_{\rm NL}^{\rm local} = 5.6 \pm 5.4$, $f_{\rm NL}^{\rm equil} = 65 \pm 58$, and $f_{\rm NL}^{\rm ortho} = -30 \pm 26$, while for SEVEM we obtain $f_{\rm NL}^{\rm local} = 9.4 \pm 5.4$, $f_{\rm NL}^{\rm equil} = 75 \pm 59$, $f_{\rm NL}^{\rm ortho} = -50 \pm 30$.