A&A 641, A9 (2020)
https://doi.org/10.1051/0004-6361/201935891
© ESO 2020

Astronomy & Astrophysics
Special issue

*Planck 2018 results*

# *Planck* 2018 results

## IX. Constraints on primordial non-Gaussianity

Planck Collaboration: Y. Akrami[12,50,52], F. Arroja[54], M. Ashdown[59,4], J. Aumont[86], C. Baccigalupi[71], M. Ballardini[19,36], A. J. Banday[86,7], R. B. Barreiro[55], N. Bartolo[24,56,⋆], S. Basak[77], K. Benabed[49,81], J.-P. Bernard[86,7], M. Bersanelli[27,40], P. Bielewicz[68,71], J. R. Bond[6], J. Borrill[10,84], F. R. Bouchet[49,81], M. Bucher[2,5], C. Burigana[39,25,42], R. C. Butler[36], E. Calabrese[74], J.-F. Cardoso[49,81], B. Casaponsa[55], A. Challinor[51,59,9], H. C. Chiang[22,5], L. P. L. Colombo[27], C. Combet[62], B. P. Crill[57,8], F. Cuttaia[36], P. de Bernardis[26], A. de Rosa[36], G. de Zotti[37], J. Delabrouille[2], J.-M. Delouis[61], E. Di Valentino[58], J. M. Diego[55], O. Doré[57,8], M. Douspis[48], A. Ducout[60], X. Dupac[30], S. Dusini[56], G. Efstathiou[59,51], F. Elsner[65], T. A. Enßlin[65], H. K. Eriksen[52], Y. Fantaye[3,17], J. Fergusson[9], R. Fernandez-Cobos[55], F. Finelli[36,42], M. Frailis[38], A. A. Fraisse[22], E. Franceschi[36], A. Frolov[79], S. Galeotta[38], S. Galli[49,81], K. Ganga[2], R. T. Génova-Santos[53,13], M. Gerbino[33], J. González-Nuevo[14], K. M. Górski[57,88], S. Gratton[59,51], A. Gruppuso[36,42], J. E. Gudmundsson[85,22], J. Hamann[78], W. Handley[59,4], F. K. Hansen[52], D. Herranz[55], E. Hivon[49,81], Z. Huang[75], A. H. Jaffe[47], W. C. Jones[22], G. Jung[24], E. Keihänen[21], R. Keskitalo[10], K. Kiiveri[21,35], J. Kim[65], N. Krachmalnicoff[71], M. Kunz[11,48,3], H. Kurki-Suonio[21,35], J.-M. Lamarre[80], A. Lasenby[4,59], M. Lattanzi[43,25], C. R. Lawrence[57], M. Le Jeune[2], F. Levrier[80], A. Lewis[20], M. Liguori[24,56], P. B. Lilje[52], V. Lindholm[21,35], M. López-Caniego[30], Y.-Z. Ma[70,73,67], J. F. Macías-Pérez[62], G. Maggio[38], D. Maino[27,40,44], N. Mandolesi[36,25], A. Marcos-Caballero[55], M. Maris[38], P. G. Martin[6], E. Martínez-González[55], S. Matarrese[24,56,32], N. Mauri[42], J. D. McEwen[66], P. D. Meerburg[59,9,87], P. R. Meinhold[23], A. Melchiorri[26,45], A. Mennella[27,40], M. Migliaccio[29,46], M.-A. Miville-Deschênes[1,48], D. Molinari[25,36,43], A. Moneti[49,81], L. Montier[86,7], G. Morgante[36], A. Moss[76], M. Münchmeyer[49], P. Natoli[25,83,43], F. Oppizzi[24], L. Pagano[25,43,48], D. Paoletti[36,42], B. Partridge[34], G. Patanchon[2], F. Perrotta[71], V. Pettorino[1], F. Piacentini[26], G. Polenta[83], J.-L. Puget[48,49], J. P. Rachen[15], B. Racine[52], M. Reinecke[65], M. Remazeilles[58], A. Renzi[56], G. Rocha[57,8], J. A. Rubiño-Martín[53,13], B. Ruiz-Granados[53,13], L. Salvati[48], M. Savelainen[21,35,64], D. Scott[18], E. P. S. Shellard[9], M. Shiraishi[24,56,16], C. Sirignano[24,56], G. Sirri[42], K. Smith[69], L. D. Spencer[74], L. Stanco[56], R. Sunyaev[65,82], A.-S. Suur-Uski[21,35], J. A. Tauber[31], D. Tavagnacco[38,28], M. Tenti[41], L. Toffolatti[14,36], M. Tomasi[27,40], T. Trombetti[39,43], J. Valiviita[21,35], B. Van Tent[63], P. Vielva[55], F. Villa[36], N. Vittorio[29], B. D. Wandelt[49,81], I. K. Wehus[52], A. Zacchei[38], and A. Zonca[72]

*(Affiliations can be found after the references)*

Received 15 May 2019 / Accepted 5 November 2019

**ABSTRACT**

We analyse the *Planck* full-mission cosmic microwave background (CMB) temperature and $E$-mode polarization maps to obtain constraints on primordial non-Gaussianity (NG). We compare estimates obtained from separable template-fitting, binned, and optimal modal bispectrum estimators, finding consistent values for the local, equilateral, and orthogonal bispectrum amplitudes. Our combined temperature and polarization analysis produces the following final results: $f_{\rm NL}^{\rm local} = -0.9 \pm 5.1$; $f_{\rm NL}^{\rm equil} = -26 \pm 47$; and $f_{\rm NL}^{\rm ortho} = -38 \pm 24$ (68% CL, statistical). These results include low-multipole ($4 \leq \ell < 40$) polarization data that are not included in our previous analysis. The results also pass an extensive battery of tests (with additional tests regarding foreground residuals compared to 2015), and they are stable with respect to our 2015 measurements (with small fluctuations, at the level of a fraction of a standard deviation, which is consistent with changes in data processing). Polarization-only bispectra display a significant improvement in robustness; they can now be used independently to set primordial NG constraints with a sensitivity comparable to WMAP temperature-based results and they give excellent agreement. In addition to the analysis of the standard local, equilateral, and orthogonal bispectrum shapes, we consider a large number of additional cases, such as scale-dependent feature and resonance bispectra, isocurvature primordial NG, and parity-breaking models, where we also place tight constraints but do not detect any signal. The non-primordial lensing bispectrum is, however, detected with an improved significance compared to 2015, excluding the null hypothesis at $3.5\sigma$. Beyond estimates of individual shape amplitudes, we also present model-independent reconstructions and analyses of the *Planck* CMB bispectrum. Our final constraint on the local primordial trispectrum shape is $g_{\rm NL}^{\rm local} = (-5.8 \pm 6.5) \times 10^4$ (68% CL, statistical), while constraints for other trispectrum shapes are also determined. Exploiting the tight limits on various bispectrum and trispectrum shapes, we constrain the parameter space of different early-Universe scenarios that generate primordial NG, including general single-field models of inflation, multi-field models (e.g. curvaton models), models of inflation with axion fields producing parity-violation bispectra in the tensor sector, and inflationary models involving vector-like fields with directionally-dependent bispectra. Our results provide a high-precision test for structure-formation scenarios, showing complete agreement with the basic picture of the ΛCDM cosmology regarding the statistics of the initial conditions, with cosmic structures arising from adiabatic, passive, Gaussian, and primordial seed perturbations.

**Key words.** cosmic background radiation – cosmology: observations – cosmology: theory – early Universe – inflation – methods: data analysis

## 1. Introduction

This paper, one of a set associated with the 2018 release (also known as "PR3") of data from the *Planck*[1] mission (Planck Collaboration I 2020), resents the data analysis and constraints on primordial non-Gaussianity (NG), which are obtained using the Legacy *Planck* cosmic microwave background (CMB)

⋆ Corresponding author: Nicola Bartolo,
e-mail: nicola.bartolo@pd.infn.it

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

maps. It also includes some implications for inflationary models driven by the 2018 NG constraints. This paper updates the earlier study based on the temperature data from the nominal *Planck* operations period, including the first 14 months of observations (Planck Collaboration XXIV 2014, hereafter PCNG13), and a later study that used temperature data and a first set of polarization maps from the full *Planck* mission–29 and 52 months of observations for the High Frequency Instrument (HFI) and the Low Frequency Instrument (LFI), respectively (Planck Collaboration XVII 2016, hereafter PCNG15). The analysis described in this paper sets the most stringent constraints to date on primordial NG, which are near what is ultimately possible from using only CMB temperature data. The results of this paper are mainly based on the measurements of the CMB angular bispectrum, which are complemented with the next higher-order NG correlation function, that is the trispectrum. For notations and conventions relating to (primordial) bispectra and trispectra we refer the reader to the two previous *Planck* papers on primordial NG (PCNG13; PCNG15). This paper also complements the precise characterization of inflationary models (Planck Collaboration X 2020) and cosmological parameters (Planck Collaboration VI 2020), with specific statistical estimators that go beyond the constraints on primordial power spectra. It also complements the statistical and isotropy tests on CMB anisotropies of Planck Collaboration VII (2020), focusing on the interpretation of specific, well motivated, non-Gaussian models of inflation. These models span from the irreducible minimal amount of primordial NG, predicted by standard single-field models of slow-roll inflation, to various classes of inflationary models that constitute the prototypes of extensions of the standard inflationary picture and of physically motivated mechanisms that are able to generate a higher level of primordial NG measurable in the CMB anisotropies. This work establishes the most robust constraints on some of the most well-known and studied types of primordial NG, namely the local, equilateral, and orthogonal shapes. Moreover, this 2018 analysis includes a better characterization of the constraints coming from CMB polarization data. Besides focusing on these major goals, we reanalyse a variety of other NG signals and also investigate some new aspects of primordial NG. For example, we perform for the first time an analysis of the running of NG using *Planck* data in the context of some well defined inflationary models. Additionally, we constrained primordial NG predicted by theoretical scenarios on which much attention has been focused recently, such as, bispectrum NG generated in the tensor (gravitational wave) sector. For a detailed analysis of oscillatory features that combines power spectrum and bispectrum constraints see Planck Collaboration X (2020). As in the last data release ("PR2"), besides extracting the constraints on NG amplitudes for specific shapes, we also provide a model-independent reconstruction of the CMB angular bispectrum by using various methods. Such a reconstruction can help to pin down interesting features in the CMB bispectrum signal beyond those captured by existing parameterizations.

The paper is organized as follows. In Sect. 2 we recall the main primordial NG models tested in this paper. Section 3 briefly describes the bispectrum estimators that we use, as well as details about the data set and our analysis procedures. In Sect. 4 we discuss detectable non-primordial contributions to the CMB bispectrum, namely those arising from lensing and point sources. In Sect. 5 we constrain $f_{\rm NL}$ for the local, equilateral, and orthogonal bispectra. We also report the results for scale-dependent NG models and other selected bispectrum shapes, including NG in the tensor (primordial gravitational wave) sector; in this section reconstructions and model-independent analyses of the CMB bispectrum are also provided. In Sect. 6 these results are validated through a series of null tests on the data, with the goal of assessing the robustness of the results. This includes, in particular, a first analysis of Galactic dust and thermal SZ residuals. *Planck* CMB trispectrum limits are obtained and discussed in Sect. 7. In Sect. 8 we derive the main implications of *Planck*'s constraints on primordial NG for some specific early Universe models. We conclude in Sect. 9.

## 2. Models

Primordial NG comes in with a variety of shapes, corresponding to well motivated classes of inflationary model. For each class, a common physical mechanism is responsible for the generation of the corresponding type of primordial NG. Below we briefly summarize the main types of primordial NG that are constrained in this paper, providing the precise shapes that are used for data analysis. For more details about specific realizations of inflationary models within each class, see the previous two *Planck* papers on primordial NG (PCNG13; PCNG15) and reviews (e.g. Bartolo et al. 2004a; Liguori et al. 2010; Chen 2010b; Komatsu 2010; Yadav & Wandelt 2010). We only give a more expanded description of those shapes of primordial NG analysed here for the first time with *Planck* data (e.g. running of primordial NG).

### 2.1. General single-field models of inflation

The parameter space of single-field models is well described by the so called equilateral and orthogonal templates (Creminelli et al. 2006; Chen et al. 2007b; Senatore et al. 2010). The equilateral shape is

$$B_\Phi^{\rm equil}(k_1,k_2,k_3) = 6A^2 f_{\rm NL}^{\rm equil} \times\left\{-\frac{1}{k_1^{4-n_s}k_2^{4-n_s}}-\frac{1}{k_2^{4-n_s}k_3^{4-n_s}}-\frac{1}{k_3^{4-n_s}k_1^{4-n_s}}-\frac{2}{(k_1k_2k_3)^{2(4-n_s)/3}} +\left[\frac{1}{k_1^{(4-n_s)/3}k_2^{2(4-n_s)/3}k_3^{4-n_s}}+5\,\text{perms.}\right]\right\}, \quad (1)$$

while the orthogonal NG is described by

$$B_\Phi^{\rm ortho}(k_1,k_2,k_3) = 6A^2 f_{\rm NL}^{\rm ortho} \times\left\{-\frac{3}{k_1^{4-n_s}k_2^{4-n_s}}-\frac{3}{k_2^{4-n_s}k_3^{4-n_s}}-\frac{3}{k_3^{4-n_s}k_1^{4-n_s}}-\frac{8}{(k_1k_2k_3)^{2(4-n_s)/3}} +\left[\frac{3}{k_1^{(4-n_s)/3}k_2^{2(4-n_s)/3}k_3^{4-n_s}}+5\,\text{perms.}\right]\right\}. \quad (2)$$

Here the potential $\Phi$ is defined in relation to the comoving curvature perturbation $\zeta$ by $\Phi \equiv (3/5)\zeta$ on superhorizon scales (thus corresponding to Bardeen's gauge-invariant gravitational potential (Bardeen 1980) during matter domination on superhorizon scales). $P_\Phi(k) = A/k^{4-n_s}$ is the Bardeen gravitational potential power spectrum, with normalization $A$ and scalar spectral index $n_s$. A typical example of this class is provided by models of inflation where there is a single scalar field driving inflation and generating the primordial perturbations, characterized by a non-standard kinetic term or more general higher-derivative interactions. In the first case the inflaton Lagrangian is $\mathcal{L} = P(X,\phi)$,

where $X = g^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi$, with at most one derivative on $\phi$ (Chen et al. 2007b). Different higher-derivative interactions of the inflaton field characterize, ghost inflation (Arkani-Hamed et al. 2004) or models of inflation based on Galileon symmetry (e.g., Burrage et al. 2011). The two amplitudes $f_{\rm NL}^{\rm equil}$ and $f_{\rm NL}^{\rm ortho}$ usually depend on the sound speed $c_{\rm s}$ at which the inflaton field fluctuations propagate and on a second independent amplitude measuring the inflaton self-interactions. The Dirac-Born-Infeld (DBI) models of inflation (Silverstein & Tong 2004; Alishahiha et al. 2004) are a string-theory-motivated example of the $P(X,\phi)$ models, predicting an almost equilateral type NG with $f_{\rm NL}^{\rm equil} \propto c_{\rm s}^{-2}$ for $c_{\rm s} \ll 1$. More generally, the effective field theory (EFT) approach to inflationary perturbations (Cheung et al. 2008; Senatore et al. 2010; Bartolo et al. 2010a) yields NG shapes that can be mapped into the equilateral and orthogonal template basis. The EFT approach allows us to draw generic conclusions about single-field inflation. We discuss them using one example in Sect. 8. Nevertheless, we shall also explicitly search for such EFT shapes, analysing their exact non-separable predicted shapes, $B^{\rm EFT1}$ and $B^{\rm EFT2}$, along with those of DBI, $B^{\rm DBI}$, and ghost inflation, $B^{\rm ghost}$ (Arkani-Hamed et al. 2004).

### 2.2. Multi-field models

The bispectrum for multi-field models is typically of the local type[2]

$$\begin{aligned} B_\Phi^{\rm local}(k_1,k_2,k_3) &= 2f_{\rm NL}^{\rm local}\Big[P_\Phi(k_1)P_\Phi(k_2) + P_\Phi(k_1)P_\Phi(k_3) \\ &\quad + P_\Phi(k_2)P_\Phi(k_3)\Big] \\ &= 2A^2 f_{\rm NL}^{\rm local}\left[\frac{1}{k_1^{4-n_{\rm s}}k_2^{4-n_{\rm s}}} + {\rm cycl.}\right]. \end{aligned} \tag{3}$$

This usually arises when more scalar fields drive inflation and give rise to the primordial curvature perturbation ("multiple-field inflation"), or when extra light scalar fields, different from the inflaton field driving inflation, determine (or contribute to) the final curvature perturbation. In these models initial isocurvature perturbations are transferred on super-horizon scales to the curvature perturbations. Non-Gaussianities if present are transferred too. This, along with non-linearities in the transfer mechanism itself, is a potential source of significant NG (Bartolo et al. 2002; Bernardeau & Uzan 2002; Vernizzi & Wands 2006; Rigopoulos et al. 2006, 2007; Lyth & Rodriguez 2005; Tzavara & van Tent 2011; Jung & van Tent 2017). The bispectrum of Eq. (3) mainly correlates large- with small-scale modes, peaking in the "squeezed" configurations $k_1 \ll k_2 \approx k_3$. This is a consequence of the transfer mechanism taking place on super-horizon scales and thus generating a localized point-by-point primordial NG in real space. The curvaton model (Mollerach 1990; Linde & Mukhanov 1997; Enqvist & Sloth 2002; Lyth & Wands 2002; Moroi & Takahashi 2001) is a clear example where local NG is generated in this way (e.g., Lyth & Wands 2002; Lyth et al. 2003; Bartolo et al. 2004c). In the minimal adiabatic curvaton scenario $f_{\rm NL}^{\rm local} = (5/4r_{\rm D}) - 5r_{\rm D}/6 - 5/3$ (Bartolo et al. 2004c,d), in the case when the curvaton field potential is purely quadratic (Lyth & Wands 2002; Lyth et al. 2003; Lyth & Rodriguez 2005; Malik & Lyth 2006; Sasaki et al. 2006). Here $r_{\rm D} = [3\rho_{\rm curv}/(3\rho_{\rm curv} + 4\rho_{\rm rad})]_{\rm D}$ represents the "curvaton decay fraction" at the epoch of the curvaton decay, employing the sudden decay approximation. Significant NG can be produced (Bartolo et al. 2004c,d) for low values of $r_{\rm D}$; a different modelling of the curvaton scenario has been discussed by Linde & Mukhanov (2006) and Sasaki et al. (2006). We update the limits on both models in Sect. 8, using the local NG constraints. More general models with a curvaton-like spectator field have also been intensively investigated recently (see, e.g., Torrado et al. 2018). Notice that through a similar mechanism to the curvaton mechanism, local bispectra can be generated from non-linear dynamics during the preheating and reheating phases (Enqvist et al. 2005; Chambers & Rajantie 2008; Barnaby & Cline 2006; Bond et al. 2009) or due to fluctuations in the decay rate or interactions of the inflaton field, as realized in modulated (p)reheating and modulated hybrid inflationary models (Kofman 2003; Dvali et al. 2004a,b; Bernardeau et al. 2004; Zaldarriaga 2004; Lyth 2005; Salem 2005; Lyth & Riotto 2006; Kolb et al. 2006; Cicoli et al. 2012). We also explore whether there is any evidence for dissipative effects during warm inflation, with a signal which changes sign in the squeezed limit (see e.g., Bastero-Gil et al. 2014).

### 2.3. Isocurvature non-Gaussianity

In most of the models mentioned in this section the focus is on primordial NG in the adiabatic curvature perturbation $\zeta$. However, in inflationary scenarios with multiple scalar fields, isocurvature perturbation modes can be produced as well. If they survive until recombination, these will then contribute not only to the power spectrum, but also to the bispectrum, producing in general both a pure isocurvature bispectrum and mixed bispectra because of the cross-correlation between isocurvature and adiabatic perturbations (Komatsu 2002; Bartolo et al. 2002; Komatsu et al. 2005; Kawasaki et al. 2008, 2009; Langlois et al. 2008; Hikage et al. 2009, 2013a,b; Langlois & Lepidi 2011; Langlois & van Tent 2011; Kawakami et al. 2012; Langlois & van Tent 2012).

In the context of the ΛCDM cosmology, there are at the time of recombination four possible distinct isocurvature modes (in addition to the adiabatic mode), namely the cold-dark-matter (CDM) density, baryon-density, neutrino-density, and neutrino-velocity isocurvature modes (Bucher et al. 2000). However, the baryon isocurvature mode behaves identically to the CDM isocurvature mode, once rescaled by factors of $\Omega_{\rm b}/\Omega_{\rm c}$, so we only consider the other three isocurvature modes in this paper. Moreover, we only investigate isocurvature NG of the local type, since this is the most relevant case in multi-field inflation models, which we require in order to produce isocurvature modes. We also limit ourselves to studying just one type of isocurvature mode (considering each of the three types separately) together with the adiabatic mode, to avoid the number of free parameters becoming so large that no meaningful limits can be derived. Finally, for simplicity we assume the same spectral index for the primordial isocurvature power spectrum and the adiabatic-isocurvature cross-power spectrum as for the adiabatic power spectrum, again to reduce the number of free parameters. As shown by Langlois & van Tent (2011), under these assumptions

[2] See, e.g. Byrnes & Choi (2010) for a review on this type of model in the context of primordial NG. Early papers discussing primordial local bispectra given by Eq. (3) include Falk et al. (1993), Gangui et al. (1994), Gangui & Martin (2000), Verde et al. (2000), Wang & Kamionkowski (2000), and Komatsu & Spergel (2001).

we have in general six independent $f_{\rm NL}$ parameters: the usual purely adiabatic one; a purely isocurvature one; and four correlated ones.

The primordial isocurvature bispectrum templates are a generalization of the local shape in Eq. (3):

$$B^{IJK}(k_1,k_2,k_3) = 2f_{\rm NL}^{I,JK}P_\Phi(k_2)P_\Phi(k_3) + 2f_{\rm NL}^{J,KI}P_\Phi(k_1)P_\Phi(k_3) + 2f_{\rm NL}^{K,IJ}P_\Phi(k_1)P_\Phi(k_2), \quad (4)$$

where $I, J, K$ label the different adiabatic and isocurvature modes. The invariance under the simultaneous exchange of two of these indices and the corresponding momenta means that $f_{\rm NL}^{I,JK} = f_{\rm NL}^{I,KJ}$, which reduces the number of independent parameters from eight to six in the case of two modes (and explains the notation with the comma). The different CMB bispectrum templates derived from these primordial shapes vary most importantly in the different types of radiation transfer functions that they contain. For more details, see in particular Langlois & van Tent (2012).

An important final remark is that, unlike the case of the purely adiabatic mode, polarization improves the constraints on the isocurvature NG significantly, up to a factor of about 6 as predicted by Langlois & van Tent (2011, 2012) and confirmed by the 2015 *Planck* analysis (PCNG15). The reason for this is that while the isocurvature temperature power spectrum (to which the local bispectrum is proportional) becomes very quickly negligible compared to the adiabatic one as $\ell$ increases (already around $\ell \approx 50$ for CDM), the isocurvature polarization power spectrum remains comparable to the adiabatic one to much smaller scales (up to $\ell \approx 200$ for CDM). Hence there are many more polarization modes than temperature modes that are relevant for determining these isocurvature $f_{\rm NL}$ parameters. For more details, again see Langlois & van Tent (2012).

### 2.4. Running non-Gaussianity

We briefly describe inflationary models that predict a mildly scale-dependent bispectrum, which is also known in the literature as the running of the bispectrum (see e.g., Chen 2005; Liguori et al. 2006; Kumar et al. 2010; Byrnes et al. 2010a,b; Shandera et al. 2011). In inflationary models this running is as natural as the running of the power spectrum, that is the spectral index $n_{\rm s}$. Other models with strong scale dependence, for example, oscillatory models, are discussed in Sect. 2.5. Further possibilities for strong scale dependence exist (see e.g., Khoury & Piazza 2009; Riotto & Sloth 2011), but we do not consider these in our study. The simplest model (single-field slow-roll, with canonical action and initial conditions) predicts that both the amplitude and scale dependence are of the order of the slow-roll parameters (this is true except in some very particular models, see, e.g., Chen et al. 2013), they are small and currently not observable. However, other elaborate but theoretically well-motivated models make different predictions and these can be used to confirm or exclude such models. Measuring the running of the non-Gaussianity parameters with scale is important because this running carries information about, for instance, the number of inflationary fields and their interactions. This information may not be accessible with the power spectrum alone. The first constraints on the running of a local model were obtained with WMAP7 data in Becker & Huterer (2012). Forecasts of what would be feasible with future data were performed by for instance LoVerde et al. (2008), Sefusatti et al. (2009), Becker et al. (2011, 2012), and Giannantonio et al. (2012).

#### 2.4.1. Local-type scale-dependent bispectrum

We start by describing models with a local-type mildly scale-dependent bispectrum. Assuming that there are multiple scalar fields during inflation with canonical kinetic terms, that their correlators are Gaussian at horizon crossing, and using the slow-roll approximation and the $\delta N$ formalism, Byrnes et al. (2010b) found a quite general expression for the power spectrum of the primordial potential perturbation:

$$P_\Phi(k) = \frac{2\pi^2}{k^3}\mathcal{P}_\Phi(k) = \frac{2\pi^2}{k^3}\sum_{ab}\mathcal{P}_{ab}(k), \quad (5)$$

where the indexes $a, b$ run over the different scalar fields.

The non-linearity parameter then reads (Byrnes et al. 2010b)

$$f_{\rm NL}(k_1,k_2,k_3) = \frac{B_\Phi(k_1,k_2,k_3)}{2\left[P_\Phi(k_1)P_\Phi(k_2) + 2\,\text{perms.}\right]} = \frac{\sum_{abcd}(k_1k_2)^{-3}\mathcal{P}_{ac}(k_1)\mathcal{P}_{bd}(k_2)f_{cd}(k_3) + 2\,\text{perms.}}{(k_1k_2)^{-3}\mathcal{P}(k_1)\mathcal{P}(k_2) + 2\,\text{perms.}}, \quad (6)$$

where the last line is the general result valid for any number of slow-roll fields. The functions $f_{cd}$ (as well as the functions $\mathcal{P}_{ab}$) can be parameterized as power laws. In the general case, $f_{\rm NL}$ can also be written as

$$f_{\rm NL}(k_1,k_2,k_3) = \sum_{ab} f_{\rm NL}^{ab}\frac{(k_1k_2)^{n_{{\rm multi},a}}k_3^{3+n_{f,ab}} + 2\,\text{perms.}}{k_1^3+k_2^3+k_3^3}, \quad (7)$$

where $n_{{\rm multi},a}$ and $n_{f,ab}$ are parameters of the models that are proportional to the slow-roll parameters. It is clear that in the general case there are too many parameters to be constrained. Instead we consider two simpler cases, which are among the three models of running non-Gaussianity that are analysed in Sect. 5.2.2.

Firstly, when the curvature perturbation originates from only one of the scalar fields (e.g. as in the simplest curvaton scenario) the bispectrum simplifies to (Byrnes et al. 2010b)

$$B_\Phi(k_1,k_2,k_3) \propto (k_1k_2)^{n_{\rm s}-4}k_3^{n_{\rm NG}} + 2\,\text{perms.} \quad (8)$$

In this case

$$f_{\rm NL}(k_1,k_2,k_3) = f_{\rm NL}^p\frac{k_1^{3+n_{\rm NG}}+k_2^{3+n_{\rm NG}}+k_3^{3+n_{\rm NG}}}{k_1^3+k_2^3+k_3^3}, \quad (9)$$

where $n_{\rm NG}$ is the running parameter which is sensitive to the third derivative of the potential. If the field producing the perturbations is not the inflaton field but an isocurvature field subdominant during inflation, then neither the spectral index measurement nor the running of the spectral index are sensitive to the third derivative. Therefore, those self-interactions can uniquely be probed by the running of $f_{\rm NL}$.

The second class of models are two-field models where both fields contribute to the generation of the perturbations but the running of the bispectrum is still given by one parameter only (by choosing some other parameters appropriately) as (Byrnes et al. 2010b)

$$B_\Phi(k_1,k_2,k_3) \propto (k_1k_2)^{n_{\rm s}-4+(n_{\rm NG}/2)} + 2\,\text{perms.} \quad (10)$$

Comparing the two templates (8) and (10) one sees that there are multiple ways to generalize (with one extra parameter) the constant local $f_{\rm NL}$ model, even with the same values for $f_{\rm NL}$ and $n_{\rm NG}$. If one is able to distinguish observationally between these two shapes then one could find out whether the running originated from single or multiple field effects for example.

Byrnes et al. (2010b) further assumed that $|n_{f_{\rm NL}} \ln(k_{\rm max}/k_{\rm min})| \ll 1$. In our case, $\ln(k_{\rm max}/k_{\rm min}) \lesssim 8$ and $n_{\rm NG}$ can be at most of order 0.1. If the observational constraints on $n_{\rm NG}$ using the previous theoretical templates turn out to be weaker, then one cannot use those constraints to limit the fundamental parameters of the models because the templates are being used in a region where they are not applicable. However, from a phenomenological point of view, we wish to argue that the previous templates are still interesting cases of scale-dependent bispectra, even in that parameter region. Byrnes et al. (2010b) also computed the running of the trispectra amplitudes $\tau_{\rm NL}$ and $g_{\rm NL}$. For general single-source models they showed that $n_{\tau_{\rm NL}} = 2n_{\rm NG}$, analogous to the well-known consistency relation $\tau_{\rm NL}(k) = \left(\frac{6}{5} f_{\rm NL}(k)\right)^2$, providing a useful consistency check (see, e.g., Smith et al. 2011).

#### 2.4.2. Equilateral type scale-dependent bispectrum

General single-field models that can produce large bispectra having a significant correlation with the equilateral template also predict a mild running non-Gaussianity. A typical example is DBI-inflation, as studied, for example by Chen (2005) and Chen et al. (2007b), with a generalization within the effective field theory of inflation in Bartolo et al. (2010d). Typically in these models a running NG arises of the form

$$f_{\rm NL} \to f_{\rm NL}^{*} \left(\frac{k_1 + k_2 + k_3}{3k_{\rm piv}}\right)^{n_{\rm NG}}, \tag{11}$$

where $n_{\rm NG}$ is the running parameter and $k_{\rm piv}$ is a pivot scale needed to constrain the amplitude. For example, in the case where the main contribution comes from a small sound speed of the inflaton field, $n_{\rm NG} = -2s$, where $s = \dot{c}_{\rm s}/(Hc_{\rm s})$, and therefore running NG allows us to constrain the time dependence of the sound speed[3]. The equilateral NG with a running of the type given in Eq. (11) is a third type of running NG analysed in Sect. 5.2.2 (together with the local single-source model of Eq. (8) and the local two-field model of Eq. (10)). We refer the reader to Sect. 3.1.2 for the details on the methodology adopted to analyse these models.

### 2.5. Oscillatory bispectrum models

Oscillatory power spectrum and bispectrum signals are possible in a variety of well-motivated inflationary models, including those with an imposed shift symmetry or if there are sharp features in the inflationary potential. Our first *Planck* temperature-only non-Gaussian analysis (PCNNG13) included a search for the simplest resonance and feature models, while the second *Planck* temperature with polarization analysis (PCNG15) substantially expanded the frequency range investigated, while also encompassing a much wider class of oscillatory models. These phenomenological bispectrum shapes had free parameters designed to capture the main properties of the key extant oscillatory models, thus surveying for any oscillatory signals present in the data at high significance. Our primary purpose here is to use the revised *Planck* 2018 data set to determine the robustness of our second analysis, so we only briefly introduce the models studied, referring the reader to our previous work (PCNG15) for more detailed information.

#### 2.5.1. Resonance and axion monodromy

Motivated by the UV completion problem facing large-field inflation, effective shift symmetries can be used to preserve the flat potentials required by inflation, with a prime example being the periodically modulated potential of axion monodromy models. This periodic symmetry can cause resonances during inflation, imprinting logarithmically-spaced oscillations in the power spectrum, bispectrum and beyond (Chen et al. 2008; Flauger et al. 2010; Hannestad et al. 2010; Flauger & Pajer 2011). For the bispectrum, to a good approximation, these models yield the simple oscillatory shape (see e.g., Chen 2010b)

$$B_{\Phi}^{\rm res}(k_1, k_2, k_3) = \frac{6A^2 f_{\rm NL}{}^{\rm res}}{(k_1 k_2 k_3)^2} \sin\left[\omega \ln(k_1 + k_2 + k_3) + \phi\right], \tag{12}$$

where the constant $\omega$ is an effective frequency associated with the underlying periodicity of the model and $\phi$ is a phase. The units for the wavenumbers, $k_i$, are arbitrary as any specific choice can be absorbed into the phase which is marginalized over in our results. There are more general resonance models that naturally combine properties of inflation inspired by fundamental theory, notably a varying sound speed $c_{\rm s}$ or an excited initial state. These tend to modulate the oscillatory signal on $K = k_1 + k_2 + k_3$ constant slices, with either equilateral or flattened shapes, respectively (see e.g., Chen 2010a), which we take to have the form

$$S^{\rm eq}(k_1, k_2, k_3) = \frac{\tilde{k}_1 \tilde{k}_2 \tilde{k}_3}{k_1 k_2 k_3}, \qquad S^{\rm flat} = 1 - S^{\rm eq}, \tag{13}$$

where $\tilde{k}_1 \equiv k_2 + k_3 - k_1$. Note that $S^{\rm eq}$ correlates closely with the equilateral shape in (1) and $S^{\rm flat}$ with the orthogonal shape in (2), since the correction from the spectral index $n_{\rm s}$ is small. The resulting generalized resonance shapes for which we search are then

$$\begin{aligned} B^{\rm res-eq}(k_1, k_2, k_3) &\equiv S^{\rm eq}(k_1, k_2, k_3) \times B^{\rm res}(k_1, k_2, k_3), \\ B^{\rm res-flat}(k_1, k_2, k_3) &\equiv S^{\rm flat}(k_1, k_2, k_3) \times B^{\rm res}(k_1, k_2, k_3). \end{aligned} \tag{14}$$

This analysis does not exhaust resonant models associated with a non-Bunch-Davies initial state, which can have a more sharply flattened shape ("enfolded" models), but it should help identify this tendency if present in the data. In addition, the discussion in Flauger et al. (2017) showed that the resonant frequency can "drift" slowly over time with a correction term to the frequency being proposed, but again we leave that for future analysis. Finally, we note that there are multifield models in which sharp corner-turning can result in residual oscillations with logarithmic spacing, thus mimicking resonance models (Achúcarro et al. 2011; Battefeld et al. 2013). However, these oscillations are more strongly damped and can be searched for by modulating the resonant shape (Eq. (12)) with a suitable envelope, as discussed for feature models below.

[3] This would help in further breaking (via primordial NG) some degeneracies among the parameters determining the curvature power spectrum in these modes. For a discussion and an analysis of this type, see Planck Collaboration XXII (2014) and Planck Collaboration XX (2016).

#### 2.5.2. Scale-dependent oscillatory features

Sharp features in the inflationary potential can generate oscillatory signatures (Chen et al. 2007a), as can rapid variations in the sound speed $c_s$ or fast turns in a multifield potential. Narrow features in the potential induce a corresponding signal in the power spectrum, bispectrum, and trispectrum; to a first approximation, the oscillatory bispectrum has a simple sinusoidal behaviour given by (Chen et al. 2007a)

$$\bar{B}^{\rm feat}(k_1,k_2,k_3) = \frac{6A^2 {f_{\rm NL}}^{\rm feat}}{(k_1k_2k_3)^2}\sin\Big[\omega(k_1+k_2+k_3)+\phi\Big], \tag{15}$$

where $\omega$ is a frequency determined by the specific feature properties and $\phi$ is a phase. The wavenumbers $k_i$ are in units of Mpc$^{-1}$. A more accurate analytic bispectrum solution has been found that includes a damping envelope taking the form (Adshead et al. 2012)

$$B^{K^2\cos}(k_1,k_2,k_3) = \frac{6A^2 {f_{\rm NL}}^{K^2\cos}}{(k_1k_2k_3)^2}K^2 D(\alpha\omega K)\cos(\omega K), \tag{16}$$

where $K = k_1 + k_2 + k_3$ and the envelope function is given by $D(\alpha\omega K) = \alpha\omega/(K\sinh(\alpha\omega K))$. Here, the model-dependent parameter $\alpha$ determines the large wavenumber cutoff, with $\alpha = 0$ for no envelope in the limit of an extremely narrow feature. Oscillatory signals generated instead by a rapidly varying sound speed $c_s$ take the form

$$B^{K\sin}(k_1,k_2,k_3) = \frac{6A^2 {f_{\rm NL}}^{K\sin}}{(k_1k_2k_3)^2}K\,D(\alpha\omega K)\sin(\omega K). \tag{17}$$

In order to encompass the widest range of physically-motivated feature models, we modulated the predicted signal (Eq. (15)) with equilateral and flattened shapes, as defined in Eq. (13), they are

$$B^{\rm feat-eq}(k_1,k_2,k_3) \equiv S^{\rm eq}(k_1,k_2,k_3)\times B^{\rm feat}(k_1,k_2,k_3)\,, \tag{18}$$

$$B^{\rm feat-flat}(k_1,k_2,k_3) \equiv S^{\rm flat}(k_1,k_2,k_3)\times B^{\rm feat}(k_1,k_2,k_3). \tag{19}$$

Like our survey of resonance models, this allows the feature signal to have arisen in inflationary models with (slowly) varying sound speeds or with excited initial states. In the latter case, it is known that very narrow features can mimic non-Bunch Davies bispectra with a flattened or enfolded shape (Chen et al. 2007a).

### 2.6. Non-Gaussianity from excited initial states

Inflationary perturbations generated by "excited" initial states (non-Bunch-Davies vacuum states) generically create non-Gaussianities with a distinct enfolded shape (see e.g., Chen et al. 2007b; Holman & Tolley 2008; Meerburg et al. 2009), that is, where the bispectrum signal is dominated by flattened configurations with $k_1 + k_2 \approx k_3$ (and cyclic permutations). In the present analysis, we investigate all the non-Bunch-Davies (NBD) models discussed in the previous *Planck* non-Gaussian papers, where explicit equations can be found for the bispectrum shape functions. In the original analysis (PCNNG13), we described: the vanilla flattened shape model $B^{\rm flat} \propto S^{\rm flat}$ already given in Eq. (13); a more realistic flattened model $B^{\rm NBD}$ (Chen et al. 2007b) from power-law $k$-inflation, with excitations generated at time $\tau_c$, yielding an oscillation period (and cut-off) $k_c \approx (\tau_c c_s)^{-1}$; the two leading-order shapes for excited canonical single-field inflation labelled $B^{\rm NBD1\text{-}cos}$ and $B^{\rm NBD2\text{-}cos}$ (Agullo & Parker 2011); and a non-oscillatory sharply flattened model $B^{\rm NBD3}$, with large enhancements from a small sound speed $c_s$ (Chen 2010b). In the second (temperature plus polarization) analysis (PCNG15), we also studied additional NBD shapes including a sinusoidal version of the original NBD bispectrum $B^{\rm NBD\text{-}sin}$ (Chen 2010a), and similar extensions for the single-field excited models (Agullo & Parker 2011), labelled $B^{\rm NBD1\text{-}sin}$ and $B^{\rm NBD2\text{-}sin}$, with the former dominated by oscillatory squeezed configurations.

### 2.7. Directional-dependent NG

The standard local bispectrum in the squeezed limit ($k_1 \ll k_2 \approx k_3$) has an amplitude that is the same for all different angles between the large-scale mode with wavevector $k_1$ and the small-scale modes parameterized by wavevector $\boldsymbol{k}_s = (\boldsymbol{k}_2 - \boldsymbol{k}_3)/2$. More generally, we can consider "anisotropic" bispectra, in the sense of an angular dependence on the orientation of the large-scale and small-scale modes, where in the squeezed limit the bispectrum depends on all even powers of $\mu = \hat{\boldsymbol{k}}_1\cdot\hat{\boldsymbol{k}}_s$ (where $\hat{\boldsymbol{k}} = \boldsymbol{k}/k$, and all odd powers vanish by symmetry even out of the squeezed limit). Expanding the squeezed bispectrum into Legendre polynomials with even multipoles $L$, the $L > 0$ shapes then be used to cleanly isolate new physical effects. In the literature it is more common to expand in the angle $\mu_{12} = \hat{\boldsymbol{k}}_1\cdot\hat{\boldsymbol{k}}_2$, which makes some aspects of the analysis simpler while introducing non-zero odd $L$ moments (since they no longer vanish when defined using the non-symmetrized small-scale wavevector). We then parameterize variations of local NG using (Shiraishi et al. 2013a):

$$B_\Phi(k_1,k_2,k_3) = \sum_L c_L[P_L(\mu_{12})P_\Phi(k_1)P_\Phi(k_2) + 2\,\text{perms.}], \tag{20}$$

where $P_L(\mu)$ is the Legendre polynomial with $P_0 = 1$, $P_1 = \mu$, and $P_2 = \frac{1}{2}(3\mu^2 - 1)$. For instance, in the $L = 1$ case the shape is given by

$$B_\Phi^{L=1}(k_1,k_2,k_3) = \frac{2A^2 {f_{\rm NL}}^{L=1}}{(k_1k_2k_3)^2}\left[\frac{k_3^2}{k_1^2k_2^2}(k_1^2+k_2^2-k_3^2) + 2\,\text{perms.}\right]. \tag{21}$$

Bispectra of the directionally dependent class in general peak in the squeezed limit ($k_1 \ll k_2 \approx k_3$), but they feature a non-trivial dependence on the parameter $\mu_{12} = \hat{\boldsymbol{k}}_1\cdot\hat{\boldsymbol{k}}_2$. The local NG template corresponds to $c_i = 2f_{\rm NL}\delta_{i0}$. The non-linearity parameters $f_{\rm NL}^L$ are related to the $c_L$ coefficients by $c_0 = 2f_{\rm NL}^{L=0}$, $c_1 = -4f_{\rm NL}^{L=1}$, and $c_2 = -16f_{\rm NL}^{L=2}$. The $L = 1$ and 2 shapes are characterized by sharp variations in the flattened limit, for example, for $k_1 + k_2 \approx k_3$, while in the squeezed limit, $L = 1$ is suppressed, unlike $L = 2$, which grows like the local bispectrum shape (i.e. the $L = 0$ case).

Bispectra of the type in Eq. (20) can arise in different inflationary models, for example, models where anisotropic sources contribute to the curvature perturbation. Bispectra of this type are indeed a general and unavoidable outcome of models that sustain long-lived superhorizon gauge vector fields during inflation (Bartolo et al. 2013a). A typical example is the case of the inflaton field $\varphi$ coupled to the kinetic term $F^2$ of a $U(1)$ gauge field $A^\mu$, via the interaction term $I^2(\varphi)F^2$, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and the coupling $I^2(\varphi)F^2$ can allow for scale invariant vector fluctuations to be generated on superhorizon scales

(Barnaby et al. 2012a; Bartolo et al. 2013a)[4]. Primordial magnetic fields sourcing curvature perturbations can also generate a dependence on both $\mu$ and $\mu^2$ (Shiraishi 2012). The $I^2(\varphi)F^2$ models predict $c_2 = c_0/2$, while models where the primordial curvature perturbations are sourced by large-scale magnetic fields produce $c_0$, $c_1$, and $c_2$. The so-called "solid inflation" models (Endlich et al. 2013; see also Bartolo et al. 2013b, 2014; Endlich et al. 2014; Sitwell & Sigurdson 2014) also predict bispectra of the form Eq. (20). In this case $c_2 \gg c_0$ (Endlich et al. 2013, 2014). Inflationary models that break rotational invariance and parity also generate this kind of NG with the specific prediction $c_0 : c_1 : c_2 = 2 : -3 : 1$ (Bartolo et al. 2015). Therefore, measurements of the $c_i$ coefficients can test for the existence of primordial vector fields during inflation, fundamental symmetries, or non-trivial structure underlying the inflationary model (as in solid inflation).

Recently much attention has been focused on the possibility of testing the presence of higher-spin particles via their imprints on higher-order inflationary correlators. Measuring primordial NG can allow us to pin down masses and spins of the particle content present during inflation, making inflation a powerful cosmological collider (Baumann & Green 2012; Chen 2010b; Chen & Wang 2010; Noumi et al. 2013; Arkani-Hamed & Maldacena 2015; Baumann et al. 2018; Arkani-Hamed et al. 2018). In the case of long-lived superhorizon higher-spin (effectively massless or partially massless higher spin fields) bispectra like in Eq. (20) are generated, where even coefficients up to $c_{n=2s}$ are excited, $s$ being the spin of the field (Franciolini et al. 2018). A structure similar to Eq. (20) arises in the case of massive spin particles, where the coefficients $c_i$ have a specific non-trivial dependence on the mass and spin of the particles, as well as on wavenumber (Arkani-Hamed & Maldacena 2015; Baumann et al. 2018; Moradinezhad Dizgah et al. 2018).

### 2.8. Parity-violating tensor non-Gaussianity motivated by pseudo-scalars

Together with the scalar mode, the NG signature in the tensor-mode sector also provides a powerful probe of inflation. In single-field inflation with Einstein gravity the tensor NG is highly suppressed. However, the tensor NG is enhanced for several modifications of Einstein gravity (e.g., Maldacena & Pimentel 2011; Gao et al. 2011; Akita & Kobayashi 2016; Domènech et al. 2017; Bartolo & Orlando 2017; Naskar & Pal 2018; Anninos et al. 2019; Ozsoy et al. 2019) or the addition of some extra source fields (in this context see, e.g., Cook & Sorbo 2012, 2013; Barnaby et al. 2012b; Senatore et al. 2014; Mirbabayi et al. 2015; Namba et al. 2016; Agrawal et al. 2018).

In some inflationary scenarios involving the axion field, there are chances to realize the characteristic NG signal in the tensor-mode sector. In these cases a non-vanishing bispectrum of primordial gravitational waves, $B_{\rm h}^{s_1 s_2 s_3}$, arises via the non-linear interaction between the axion and the gauge field. Its magnitude varies depending on the shape of the axion-gauge coupling, and, in the best-case scenario, the tensor mode can be comparable in size to or dominate the scalar mode (Cook & Sorbo 2013; Namba et al. 2016; Agrawal et al. 2018).

The induced tensor bispectrum is polarized as $B_{\rm h}^{+++} \gg B_{\rm h}^{++-}, B_{\rm h}^{+--}, B_{\rm h}^{---}$ (because the source gauge field is maximally chiral), and peaked at around the equilateral limit (because the tensor-mode production is a subhorizon event). Its size is therefore quantified by the so-called tensor non-linearity parameter,

$$f_{\rm NL}^{\rm tens} \equiv \lim_{k_i \to k} \frac{B_{\rm h}^{+++}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)}{F_\zeta^{\rm equil}(k_1, k_2, k_3)}, \tag{22}$$

with $F_\zeta^{\rm equil} \equiv (5/3)^3 B_\Phi^{\rm equil} / f_{\rm NL}^{\rm equil}$.

In this paper we constrain $f_{\rm NL}^{\rm tens}$ by measuring the CMB temperature and $E$-mode bispectra computed from $B_{\rm h}^{+++}$ (for the exact shape of $B_{\rm h}^{+++}$ see PCNG15). By virtue of their parity-violating nature, the induced CMB bispectra have non-vanishing signal for not only the even but also the odd $\ell_1 + \ell_2 + \ell_3$ triplets (Shiraishi et al. 2013b). Both are investigated in our analysis, yielding more unbiased and accurate results. The *Planck* 2015 paper (PCNG15) found a best limit of $f_{\rm NL}^{\rm tens} = (0 \pm 13) \times 10^2$ (68% CL), from the foreground-cleaned temperature and high-pass filtered $E$-mode data, where the $E$-mode information for $\ell < 40$ was entirely discarded in order to avoid foreground contamination. This paper updates those limits with additional data, including large-scale $E$-mode information.

We stress that there are also other theoretical models that yield interesting NG signals in the tensor-tensor-tensor bispectrum, as well as mixed correlations such as tensor-tensor-scalar and tensor-scalar-scalar bispectra. Such NG signals would be testable by using the $B$-mode polarization, while the analysis of $T$-mode and $E$-mode bispectra provides quite limited information, even with the *Planck* resolution because of large contamination due to the scalar-scalar-scalar bispectrum (Meerburg et al. 2016a; Domènech et al. 2017; Bartolo et al. 2019). In contrast, for models involving the axion and the gauge field, the tensor contribution dominates the scalar one at large scales and hence such a scalar-mode bias is negligible. This is one reason why we are now focusing on the axion-model $f_{\rm NL}$ template (Eq. (22)) and measuring it with the *Planck* $T$- and $E$-mode data.

## 3. Estimators and data analysis procedures

### 3.1. Bispectrum estimators

We give here a short description of the data-analysis procedures used in this paper. For additional details, we refer the reader to the primordial NG analysis associated with previous *Planck* releases (PCNG13; PCNG15) and to references provided below.

For a rotationally invariant CMB sky and even parity bispectra (as is the case for combinations of $T$ and $E$), the angular bispectrum can be written as

$$\langle a_{\ell_1 m_1}^{X_1} a_{\ell_2 m_2}^{X_2} a_{\ell_3 m_3}^{X_3} \rangle = \mathcal{G}_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} \, b_{\ell_1 \ell_2 \ell_3}^{X_1 X_2 X_3}, \tag{23}$$

where $b_{\ell_1 \ell_2 \ell_3}^{X_1 X_2 X_3}$ defines the "reduced bispectrum," and $\mathcal{G}_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3}$ is the Gaunt integral, tahis is the integral over solid angle of the product of three spherical harmonics,

$$\mathcal{G}_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} \equiv \int Y_{\ell_1 m_1}(\hat{\boldsymbol{n}}) \, Y_{\ell_2 m_2}(\hat{\boldsymbol{n}}) \, Y_{\ell_3 m_3}(\hat{\boldsymbol{n}}) \, {\rm d}^2\hat{\boldsymbol{n}}. \tag{24}$$

The Gaunt integral (which can be expressed as a product of Wigner 3$j$-symbols) enforces rotational symmetry. It satisfies both a triangle inequality and a limit given by some maximum experimental resolution $\ell_{\rm max}$. This defines a tetrahedral domain of allowed bispectrum triplets, $\{\ell_1, \ell_2, \ell_3\}$.

In order to estimate the $f_{\rm NL}$ value for a given primordial shape, we need to compute a theoretical prediction of the corresponding CMB bispectrum ansatz $b_{\ell_1 \ell_2 \ell_3}^{\rm th}$ and fit it to the observed 3-point function (see e.g., Komatsu & Spergel 2001).

[4] Notice that indeed these models generate bispectra (and power spectra) that break statistical isotropy and, after an angle average, the bispectrum takes the above expression (Eq. (20)).

Optimal cubic bispectrum estimators were first discussed in Heavens (1998). It was then shown that, in the limit of small NG, the optimal polarized $f_{\rm NL}$ estimator is described by (Creminelli et al. 2006)

$$\hat{f}_{\rm NL} = \frac{1}{N}\sum_{X_i,X_i'}\sum_{\ell_i,m_i}\sum_{\ell_i',m_i'} \mathcal{G}^{\ell_1\,\ell_2\,\ell_3}_{m_1m_2m_3} b^{X_1X_2X_3,\,\rm th}_{\ell_1\ell_2\ell_3} \Bigg\{\Bigg[\left(\mathsf{C}^{-1}_{\ell_1m_1,\ell_1'm_1'}\right)^{X_1X_1'} a^{X_1'}_{\ell_1'm_1'} \times \left(\mathsf{C}^{-1}_{\ell_2m_2,\ell_2'm_2'}\right)^{X_2X_2'} a^{X_2'}_{\ell_2'm_2'} \left(\mathsf{C}^{-1}_{\ell_3m_3,\ell_3'm_3'}\right)^{X_3X_3'} a^{X_3'}_{\ell_3'm_3'}\Bigg] - \Bigg[\left(\mathsf{C}^{-1}_{\ell_1m_1,\ell_2m_2}\right)^{X_1X_2}\left(\mathsf{C}^{-1}_{\ell_3m_3,\ell_3'm_3'}\right)^{X_3X_3'} a^{X_3'}_{\ell_3'm_3'} + \text{cyclic}\Bigg]\Bigg\}, \quad (25)$$

where the normalization $N$ is fixed by requiring unit response to $b^{\rm th}_{\ell_1\ell_2\ell_3}$ when $f_{\rm NL} = 1$. $\mathsf{C}^{-1}$ is the inverse of the block matrix:

$$\mathsf{C} = \left(\begin{array}{c|c} \mathsf{C}^{TT} & \mathsf{C}^{TE} \\ \hline \mathsf{C}^{ET} & \mathsf{C}^{EE} \end{array}\right). \quad (26)$$

The blocks represent the full $TT$, $TE$, and $EE$ covariance matrices, with $\mathsf{C}^{ET}$ being the transpose of $\mathsf{C}^{TE}$. CMB $a_{\ell m}$ coefficients, bispectrum templates, and covariance matrices in the previous relation are assumed to include instrumental beam and noise.

As shown in the formula above, these estimators are always characterized by the presence of two distinct contributions. One is cubic in the observed multipoles, and computes the correlation between the observed bispectrum and the theoretical template $b^{\rm th}_{\ell_1\ell_2\ell_3}$. This is generally called the "cubic term" of the estimator. The other is instead linear in the observed multipoles. Its role is that of correcting for mean-field contributions to the uncertainties, generated by the breaking of rotational invariance, due to the presence of a mask or to anisotropic/correlated instrumental noise (Creminelli et al. 2006; Yadav et al. 2008).

Performing the inverse-covariance filtering operation implied by Eq. (25) is numerically very demanding (Smith et al. 2009; Elsner & Wandelt 2012). An alternative, simplified approach, is that of working in the "diagonal covariance approximation," yielding (Yadav et al. 2007)

$$\hat{f}_{\rm NL} = \frac{1}{N}\sum_{X_i,X_i'}\sum_{\ell_i,m_i} \mathcal{G}^{\ell_1\,\ell_2\,\ell_3}_{m_1m_2m_3} (\mathsf{C}^{-1})^{X_1X_1'}_{\ell_1} (\mathsf{C}^{-1})^{X_2X_2'}_{\ell_2} (\mathsf{C}^{-1})^{X_3X_3'}_{\ell_3} b^{X_1X_2X_3,\,\rm th}_{\ell_1\ell_2\ell_3} \times \Bigg[a^{X_1'}_{\ell_1m_1} a^{X_2'}_{\ell_2m_2} a^{X_3'}_{\ell_3m_3} - \mathsf{C}^{X_1'X_2'}_{\ell_1m_1,\ell_2m_2} a^{X_3'}_{\ell_3m_3} - \mathsf{C}^{X_1'X_3'}_{\ell_1m_1,\ell_3m_3} a^{X_2'}_{\ell_2m_2} - \mathsf{C}^{X_2'X_3'}_{\ell_2m_2,\ell_3m_3} a^{X_1'}_{\ell_1m_1}\Bigg]. \quad (27)$$

Here, $\mathsf{C}_\ell^{-1}$ represents the inverse of the following $2\times 2$ matrix:

$$\mathsf{C}_\ell = \begin{pmatrix} \mathsf{C}_\ell^{TT} & \mathsf{C}_\ell^{TE} \\ \mathsf{C}_\ell^{ET} & \mathsf{C}_\ell^{EE} \end{pmatrix}. \quad (28)$$

As already described in PCNG13, we find that this simplification, while avoiding the covariance-inversion operation, still leads to uncertainties that are very close to optimal, provided that the multipoles are pre-filtered using a simple diffusive inpainting method. As in previous analyses, we stick to this approach here.

A brute-force implementation of Eq. (27) would require the evaluation of all the possible bispectrum configurations in our data set. This is completely unfeasible, as it would scale as $\ell_{\rm max}^5$. The three different bispectrum estimation pipelines employed in this analysis are characterized by the different approaches used to address this issue.

Before describing these methods in more detail in the following sections, we would like to stress here, the importance of having these multiple approaches. The obvious advantage is that this redundancy enables a stringent cross-validation of our results. There is, however, much more than that, as different methods allow a broad range of applications, beyond $f_{\rm NL}$ estimation, such as, for example, model-independent reconstruction of the bispectrum in different decomposition domains, precise characterization of spurious bispectrum components, monitoring direction-dependent NG signals, and so on.

#### 3.1.1. KSW and skew-$C_\ell$ estimators

Komatsu-Spergel-Wandelt (KSW) and skew-$C_\ell$ estimators (Komatsu et al. 2005; Munshi & Heavens 2010) can be applied to bispectrum templates that can be factorized, they can be written or well approximated as a linear combination of separate products of functions. This is the case for the standard local, equilateral, and orthogonal shapes, which cover a large range of theoretically motivated scenarios. The idea is that factorization leads to a massive reduction in computational time, via reduction of the three-dimensional summation over $\ell_1$, $\ell_2$, $\ell_3$ into a product of three separate one-dimensional sums over each multipole.

The skew-$C_\ell$ pipeline differs from KSW essentially in that, before collapsing the estimate into the $f_{\rm NL}$ parameter, it initially determines the so called "bispectrum-related power spectrum" (in short, "skew-$C_\ell$") function (see Munshi & Heavens 2010) for details). The slope of this function is shape-dependent, which makes the skew-$C_\ell$ extension very useful to separate and monitor multiple and spurious NG components in the map.

#### 3.1.2. Running of primordial non-Gaussianity

In the previous 2015 analysis, the KSW pipeline was used only to constrain the separable local, equilateral, orthogonal, and lensing templates. In the current analysis we extend its scope by adding the capability to constrain running of non-Gaussianity, encoded in the spectral index of the non-linear amplitude $f_{\rm NL}$, denoted $n_{\rm NG}$.

In our analysis we consider both the two local running templates, described by Eqs. (8) and (10) in Sect. 2.4.1, and the general parametrizazion for equilateral running of Sect. 2.4.2, which reads:

$$f_{\rm NL} \to f^*_{\rm NL}\left(\frac{k_1+k_2+k_3}{3k_{\rm piv}}\right)^{n_{\rm NG}}, \quad (29)$$

where $n_{\rm NG}$ is the running parameter and $k_{\rm piv}$ is a pivot scale needed to constrain the amplitude. Contrary to the two local running shapes this expression is not explicitly separable. To make it suitable for the KSW estimator (e.g. to preserve the factorizability over $k_i$), we can use a Schwinger parametrization and rearrange it as

$$f_{\rm NL} \to \frac{f^*_{\rm NL}}{3k_{\rm piv}^{n_{\rm NG}}}\frac{k_{\rm sum}}{\Gamma(1-n_{\rm NG})}\int_0^\infty {\rm d}t\, t^{-n_{\rm NG}} e^{-tk_{\rm sum}}, \quad (30)$$

where $k_{\rm sum} = k_1 + k_2 + k_3$.

Alternatively, but not equivalently, factorizability can be preserved by replacing the arithmetic mean of the three wavenumbers with the geometric mean (Sefusatti et al. 2009):

$$f_{\rm NL} \to f^*_{\rm NL}\left(\frac{k_1k_2k_3}{k_{\rm piv}^3}\right)^{\frac{n_{\rm NG}}{3}}. \quad (31)$$

Making one of these substitutions immediately yields the scale-dependent version of any bispectrum shape. Analysis in

Oppizzi et al. (2018) has shown strong correlation between the two templates, where the former behaves better numerically and is the template of choice for the running in this analysis.

A generalization of the local model, taking into account the scale dependence of $f_{\rm NL}$, can be found in Byrnes et al. (2010b), as summarized in Sect. 2.4.1.

Unlike $f_{\rm NL}$, the running parameter $n_{\rm NG}$ cannot be estimated via direct template fitting. The optimal estimation procedure, developed in Becker & Huterer (2012) and extended to all the scale-dependent shapes treated here in Oppizzi et al. (2018), is based instead on the reconstruction of the likelihood function, with respect $n_{\rm NG}$. The method exploits the KSW estimator to obtain estimates of $f^*_{\rm NL}$ for different values of the running, using explicitly separable bispectrum templates. With these values in hand, the running parameter probability density function (PDF) is computed from its analytical expression.

The computation of the marginalized likelihood depends on the choice of the prior distributions; in Becker & Huterer (2012) and Oppizzi et al. (2018) a flat prior on $f^*_{\rm NL}$ was assumed. This prior depends on the choice of the arbitrary pivotal scale $\boldsymbol{k}_{\rm piv}$, since a flat prior on $f^*_{\rm NL}$ defined at a certain scale, corresponds to a non-flat prior for another scale. The common solution is to select the pivot scale that minimizes the correlation between the parameters. This is in general a good choice, and would work properly in the case of a significant detection of a bispectrum signal. In the absence of a clear detection, however, it is worth noting some caveats. Since the range of scales available is obviously finite, a fit performed at a certain pivot scale tend to favour particular values of $n_{\rm NG}$. Therefore, there is not a perfectly 'fair' scale for the fit. As a consequence, statistical artefacts can affect the estimated constraints in the case of low significance of the measured $f^*_{\rm NL}$ central value. To prevent this issue, we resort to two additional approaches that make the final $n_{\rm NG}$ PDF pivot independent: the implementation of a parametrization invariant Jeffreys prior; and frequentist likelihood profiling. Assuming that the bispectrum configurations follow a Gaussian distribution, the likelihood can be written as (see Becker & Huterer 2012, for a derivation)

$$\mathcal{L}(n_{\rm NG}, f^*_{\rm NL}) \propto \exp\left[-\frac{N(f^*_{\rm NL}-\hat{f}_{\rm NL})^2}{2}\right] \exp\left(\frac{\hat{f}^2_{\rm NL}N}{2}\right), \tag{32}$$

where $\hat{f}_{\rm NL}$ is the value of the NG amplitude recovered from the KSW estimator for a fixed $n_{\rm NG}$ value of the running, and $N$ is the KSW normalization factor. Integrating this expression with respect to $f^*_{\rm NL}$ we obtain the marginalized likelihood. Assuming a constant prior we obtain

$$\mathcal{L}(n_{\rm NG}) \propto \frac{1}{\sqrt{N}} \exp\left(\frac{\hat{f}^2_{\rm NL}N}{2}\right). \tag{33}$$

The Jeffreys prior is defined as the square root of the determinant of the Fisher information matrix $\mathcal{I}(f_{\rm NL}, n_{\rm NG})$. In the case of separable scale-dependent bispectra, the Fisher matrix is

$$\mathcal{I}_{\alpha,\beta} \equiv \sum_{\ell_1\le\ell_2\le\ell_3} \frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi} \begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ 0 & 0 & 0\end{pmatrix}^2 \times \frac{1}{\sigma^2_{\ell_1\ell_2\ell_3}} \frac{\partial b_{\ell_1\ell_2\ell_3}}{\partial\theta_\alpha} \frac{\partial b_{\ell_1\ell_2\ell_3}}{\partial\theta_\beta}, \tag{34}$$

where $\theta_\alpha$ and $\theta_\beta$ correspond to $f^*_{\rm NL}$ or $n_{\rm NG}$ (depending on the value of the index), $b_{\ell_1\ell_2\ell_3}$ is the reduced bispectrum, and the matrix is a Wigner-3j symbol.

We search an expression for the posterior distribution marginalized over $f^*_{\rm NL}$. Assuming the Jeffreys prior for both parameters and integrating over $f_{\rm NL}$, we obtain the marginalized posterior

$$\mathcal{P}(n_{\rm NG}) \propto \left[\hat{f}_{\rm NL}\sqrt{\frac{2\pi}{N}}\ \exp\left(\frac{\hat{f}^2_{\rm NL}N}{2}\right) \mathrm{erf}\left(\hat{f}_{\rm NL}\sqrt{\frac{N}{2}}\right) + \frac{2}{N}\right] \times \sqrt{\det(\mathcal{I}(f^*_{\rm NL}=1, n_{\rm NG}))}. \tag{35}$$

The implementation of this expression in the estimator is straightforward; the only additional step is the numerical computation of the Fisher matrix determinant for each value of $n_{\rm NG}$ considered. The derived expression is independent of the pivot scale.

Alternatively, in the frequentist approach, instead of marginalizing over $f^*_{\rm NG}$, the likelihood is sampled along its maximum for every $n_{\rm NG}$ value. For fixed $n_{\rm NG}$, the maximum likelihood $f^*_{\rm NG}$ is given exactly by the KSW estimator $\hat{f}_{\rm NL}$. From Eq. (32), we see that for this condition the first exponential is set to 1 (since $f^*_{\rm NG} = \hat{f}_{\rm NL}$ at the maximum), and the profile likelihood reduces to

$$\mathcal{L}(n_{\rm NG}) \propto \exp\left(\frac{\hat{f}^2_{\rm NL}N}{2}\right). \tag{36}$$

Notice that this expression also does not depend on the pivot scale. We additionally used this expression to perform a likelihood ratio test between our scale-dependent models and the standard local and equilateral shapes.

### 3.1.3. Modal estimators

Modal estimators (Fergusson et al. 2010, 2012) are based on constructing complete, orthogonal bases of separable bispectrum templates ("bispectrum modes") and finding their amplitudes by fitting them to the data. This procedure can be made fast, due to the separability of the modes, via a KSW type of approach. The vector of estimated mode amplitudes is referred to as the "mode spectrum". This mode spectrum is theory independent and it contains all the information that needs to be extracted from the data. It is also possible to obtain theoretical mode spectra, by expanding primordial shapes in the same modal basis used to analyse the data. This allows us to measure $f_{\rm NL}$ for any given primordial bispectrum template, by correlating the theoretical mode vectors, which can be quickly computed for any shape, with the data mode spectrum. This feature makes modal techniques ideal for analyses of a large number of competing models. Also important is that non-separable bispectra are expanded with arbitrary precision into separable basis modes. Therefore the treatment of non-separable shapes is always numerically efficient in the modal approach. Finally, the data mode spectrum can be used, in combination with measured mode amplitudes, to build linear combinations of basis templates, which provide a model-independent reconstruction of the full data bispectrum. This reconstruction is of course smoothed in practice, since we use a finite number of modes. The modal bispectrum presented here follows the same approach as in 2015. In particular we use two modal pipelines, "Modal 1" and "Modal 2", characterized both by a different approach to the decomposition of polarized bispectra and by a different choice of basis, as detailed in PCNG13, PCNG15, and at the end of Sect. 3.2.2.

#### 3.1.4. Binned bispectrum estimator

The "Binned" bispectrum estimator (Bucher et al. 2010, 2016) is based on the exact optimal $f_{\rm NL}$ estimator, in combination with the observation that many bispectra of interest are relatively smooth functions in $\ell$ space. This means that data and templates can be binned in $\ell$ space with minimal loss of information, but with large computational gains. As a consequence, no KSW-like approach is required, and the theoretical templates and observational bispectra are computed and stored completely independently, and only combined at the very last stage in a sum over the bins to obtain $f_{\rm NL}$. This has several advantages: the method is fast; it is easy to test additional shapes without having to rerun the maps; the bispectrum of a map can be studied on its own in a non-parametric approach (a binned reconstruction of the full data bispectrum is provided, which can additionally be smoothed); and the dependence of $f_{\rm NL}$ on $\ell$ can be investigated for free, simply by leaving out bins from the final sum. All of these advantages are used to good effect in this paper.

The Binned bispectrum estimator was described in more detail in the papers associated with the 2013 and 2015 *Planck* releases, and full details can be found in Bucher et al. (2016). The one major change made to the Binned estimator code compared to the 2015 release concerns the computation of the linear correction term, required to make the estimator optimal in the case that rotational invariance is broken, as it is in the *Planck* analysis because of the mask and anisotropic noise. The version of the code used in 2015, while fast to compute the linear correction for a single map, scaled poorly with the number of maps, as the product of the data map with all the Gaussian maps squared had to be recomputed for each data map. Hence computing real errors, which requires analysing a large set of realistic simulations, was slow. The new code can precompute the average of the Gaussian maps squared, and then quickly apply it to all the data maps. For the full *Planck* analysis, with errors based on 300 simulations, one gains an order of magnitude in computing time (see Bucher et al. 2016, for more details).

#### 3.1.5. High-frequency feature and resonant-model estimator

In addition to the modal expansion described in Sect. 3.1.3, we have considered an extended frequency range of the constant feature model and the constant resonance model with specialized high-frequency estimators, as in PCNG15. The constant feature bispectrum in Eq. (15) is separable and thus allows for the construction of a KSW estimator (Münchmeyer et al. 2014) for direct bispectrum estimation at any given frequency. For the constant resonance model in Eq. (12), we use the method of Münchmeyer et al. (2015), which expands the logarithmic oscillations in terms of separable linear oscillatory functions. We used 800 sine and cosine modes for this expansion.

### 3.2. Data set and analysis procedures

#### 3.2.1. Data set and simulations

For our temperature and polarization data analyses we use the *Planck* 2018 CMB maps, as constructed with the four component-separation methods, `SMICA`, `SEVEM`, `NILC`, and `Commander` (Planck Collaboration IV 2020). We also make much use of simulated maps, for several different purposes, from computing errors to evaluating the linear mean-field correction terms for our estimators, as well as for performing data-validation checks. Where not otherwise specified we used the FFP10 simulation data set described in Planck Collaboration II (2020), Planck Collaboration III (2020), and Planck Collaboration IV (2020), which are the most realistic *Planck* simulations currently available. The maps we consider have been processed through the same four component-separation pipelines. The same weights used by the different pipelines on actual data have been adopted to combine different simulated frequency channels.

Simulations and data are masked using the common masks of the *Planck* 2018 release in temperature and polarization; see Planck Collaboration IV (2020) for a description of how these masks have been produced. The sky coverage fractions are, $f_{\rm sky} = 0.779$ in temperature and $f_{\rm sky} = 0.781$ in polarization.

#### 3.2.2. Data analysis details

Now we describe the setup adopted for the analysis of *Planck* 2018 data by the four different $f_{\rm NL}$ estimators described earlier in this section.

In order to smooth mask edges and retain optimality, as explained earlier, we inpaint the mask via a simple diffusive inpainting method (Bucher et al. 2016). First, we fill the masked regions with the average value of the non-masked part of the map. Then we replace each masked pixel with the average value of its neighbours and iterate this 2000 times. This is exactly the same procedure as adopted in 2013 and 2015.

Linear correction terms and $f_{\rm NL}$ errors are obtained from the FFP10 simulations, processed through the four component-separation pipelines. To this end, all pipelines use all the available 300 FFP10 noise realizations, except both modal estimators, which use only 160 maps (in order to speed up the computation). The good convergence of the modal pipelines with 160 maps was thoroughly tested in previous releases. There, we showed with a large number of tests on realistic simulations that the level of agreement between all our bispectrum estimators was perfectly consistent with theoretical expectations. Accurate tests have found some mismatch between the noise levels in the data and that of the FFP10 simulations (Planck Collaboration III 2020; Planck Collaboration IV 2020). This is roughly at the 3% level in the noise power spectrum at $\ell \approx 2000$, in temperature, and at the percent level, in polarization. We find (see Sect. 6.2 for details) that this mismatch does not play a significant role in our analysis, and can safely be ignored.

All theoretical quantities (e.g. bispectrum templates and lensing bias) are computed assuming the *Planck* 2018 best-fit cosmology and making use of the CAMB computer code[5] (Lewis et al. 2000) to compute radiation transfer functions and theoretical power spectra. The HEALPix computer code[6] (Górski et al. 2005) is used to perform spherical harmonic transforms.

As far as temperature is concerned, we maintain the same multipole ranges as in the 2013 and 2015 analyses, which is $2 \leq \ell \leq 2000$ for the KSW and modal estimators and $2 \leq \ell \leq 2500$ for the Binned estimator. The different choice of $\ell_{\rm max}$ does not produce any significant effect on the results, since the $2000 < \ell \leq 2500$ range is noise dominated and the measured value of $f_{\rm NL}$ remains very stable in that range, as confirmed by validation tests discussed in Sect. 6. The angular resolution (beam FWHM) of both the cleaned temperature and polarization maps is 5 arcmin.

The main novelty of the current analysis is the use of the low-$\ell$ polarization multipoles that were not exploited in 2015 ($\ell < 40$ polarization multipoles were removed by means of a high-pass filter). More precisely, KSW and modal estimators work in the

[5] `http://camb.info/`
[6] `http://healpix.sourceforge.net/`

polarization multipole range $4 \leq \ell \leq 1500$, while the Binned estimator considers $4 \leq \ell \leq 2000$. For the same reasons as above, this different choice of $\ell_{\max}$ does not have any impact on the results. The impact of these extra polarization modes on the final results is discussed at the end of Sect. 5.1.

The choice of using $\ell_{\min} = 4$ for polarization, thus removing the first two polarization multipoles, is instead dictated by the presence of some anomalous results in tests on simulations. When $\ell = 2$ and $\ell = 3$ are included, we observe some small bias arising in the local $f_{\rm NL}$ measurement extracted from FFP10 maps, together with a spurious increase of the uncertainties. We also notice larger discrepancies between the different estimation pipelines than expected from either theoretical arguments or previous validation tests on simulations. This can be ascribed to the presence of some small level of non-Gaussianity in the polarization noise at very low $\ell$. We stress that the choice of cutting the first two polarization multipoles does not present any particular issue, since it is performed a priori, before looking at the data (as opposed to the simulations), and generates an essentially negligible loss of information.

In addition, the Binned bispectrum estimator removes from the analysis all bispectrum $TEE$ configurations (i.e. those involving one temperature mode and two polarization modes) with the temperature mode in the bin [2, 3]. This is again motivated by optimality considerations: with these modes included the computed errors are much larger than the optimal Fisher errors, while after removing them the errors are effectively optimal. As errors are computed from simulations, this is again an a priori choice, made before looking at the data. It is not clear why only the Binned bispectrum estimator requires this additional removal, but of course the estimators are all quite different, with different sensitivities, which is exactly one of the strengths of having multiple estimators for our analyses.

The Binned bispectrum estimator uses a binning that is identical to the one in 2015, with 57 bins. The boundary values of the bins are 2, 4, 10, 18, 30, 40, 53, 71, 99, 126, 154, 211, 243, 281, 309, 343, 378, 420, 445, 476, 518, 549, 591, 619, 659, 700, 742, 771, 800, 849, 899, 931, 966, 1001, 1035, 1092, 1150, 1184, 1230, 1257, 1291, 1346, 1400, 1460, 1501, 1520, 1540, 1575, 1610, 1665, 1725, 1795, 1846, 1897, 2001, 2091, 2240, and 2500 (i.e. the first bin is [2, 3], the second [4, 9], etc., while the last one is [2240, 2500]). This binning was determined in 2015 by minimizing the increase in the theoretical variance for the primordial shapes due to the binning.

As in our 2015 analysis, we use two different polarized modal estimators. The Modal 1 pipeline expands separately the $TTT$, $EEE$, $TTE$, and $EET$ bispectra (Shiraishi et al. 2019). It then writes the estimator normalization in separable expanded form and estimates $f_{\rm NL}$ via a direct implementation of Eq. (27). The Modal 2 pipeline uses a different approach (see Fergusson 2014, for details). It first orthogonalizes $T$ and $E$ multipoles to produce new, uncorrelated, $\hat{a}^T_{\ell m}$ and $\hat{a}^E_{\ell m}$ coefficients. It then builds uncorrelated bispectra out of these coefficients, which are constrained independently, simplifying the form and reducing the number of terms in the estimator. However, the rotation procedure does not allow a direct estimation of the $EEE$ bispectrum. Direct $EEE$ reconstruction is generally useful for validation purposes and can be performed with the Modal 1 estimator.

As in 2015, Modal 1 is used to study in detail the local, equilateral, and orthogonal shapes, as well as to perform a large number of validation and robustness tests. Modal 2 is mostly dedicated to a thorough study of non-standard shapes having a large parameter space (like oscillatory bispectra). The two pipelines are equipped with modal bases optimized for their respective purposes. Modal 1 uses 600 polynomial modes, augmented with radial modes extracted from the KSW expansion of the local, equilateral, and orthogonal templates, in order to speed up convergence for these shapes. The Modal 2 expansion uses a higher-resolution basis, including 2000 polynomial modes and a Sachs-Wolfe local template, to improve efficiency in the squeezed limit.

For oscillating non-Gaussianities the Modal 2 estimator loses resolution for models with high frequencies. In order to constrain this region of parameter space we use two specialized estimators (Münchmeyer et al. 2014, 2015) that specifically target the high-frequency range of shapes with the low frequency range used for cross-valiation with Modal 2. Both of these estimators are equivalent to those used in PCNG15.

## 4. Non-primordial contributions to the CMB bispectrum

In this section we investigate those non-primordial contributions to the CMB bispectrum that we can detect in the cleaned maps, namely lensing and extragalactic point sources. These then potentially have to be taken into account when determining the constraints on the various primordial NG shapes in Sect. 5. On the other hand, the study of other non-primordial contaminants (that we do not detect in the cleaned maps) is part of the validation work in Sect. 6.

### 4.1. Non-Gaussianity from the lensing bispectrum

CMB lensing generates a significant CMB bispectrum (Goldberg & Spergel 1999; Hanson & Lewis 2009; Mangilli & Verde 2009; Lewis et al. 2011; Mangilli et al. 2013). In temperature, this is due to correlations between the lensing potential and the integrated Sachs-Wolfe (ISW; Sachs & Wolfe 1967) contribution to the CMB anisotropies. In polarization, the dominant contribution instead comes from correlations between the lensing potential and $E$ modes generated by scattering at reionization. Both the ISW and reionization contributions affect large scales, while lensing is a small-scale effect, so that the resulting bispectra peak on squeezed configurations. Therefore, they can significantly contaminate primordial NG measurements, especially for the local shape. It has been known for a while that for a high-precision experiment like *Planck* the effect is large enough that it must be taken into account. The temperature-only 2013 *Planck* results (PCNG13, Planck Collaboration XIX 2014, Planck Collaboration XVII 2014) showed the first detection of the lensing CMB temperature bispectrum and the associated bias. This was later confirmed in the 2015 *Planck* results (PCNG15, Planck Collaboration XXI 2016) both for $T$-only and for the full $T$+$E$ results. The template of the lensing bispectrum[7] is given by (Hu 2000; Lewis et al. 2011),

$$\begin{aligned} b^{X_1X_2X_3,\,\rm lens}_{\ell_1\ell_2\ell_3} = f^{\rm lens}_{\rm NL}\Big( & C^{X_2\phi}_{\ell_2}\tilde{C}^{X_1X_3}_{\ell_3} f^{X_1}_{\ell_1\ell_2\ell_3} + C^{X_3\phi}_{\ell_3}\tilde{C}^{X_1X_2}_{\ell_2} f^{X_1}_{\ell_1\ell_3\ell_2} \\ & + C^{X_1\phi}_{\ell_1}\tilde{C}^{X_2X_3}_{\ell_3} f^{X_2}_{\ell_2\ell_1\ell_3} + C^{X_3\phi}_{\ell_3}\tilde{C}^{X_1X_2}_{\ell_1} f^{X_2}_{\ell_2\ell_3\ell_1} \\ & + C^{X_1\phi}_{\ell_1}\tilde{C}^{X_2X_3}_{\ell_2} f^{X_3}_{\ell_3\ell_1\ell_2} + C^{X_2\phi}_{\ell_2}\tilde{C}^{X_1X_3}_{\ell_1} f^{X_3}_{\ell_3\ell_2\ell_1}\Big) \end{aligned} \tag{37}$$

[7] As a reminder, let us stress that the expression 'lensing bispectrum' in this paper always refers to the 3-point function generated by correlations between the lensing potential and ISW or reionization contributions, as explained in the main text. We are therefore not referring here to NG lensing signatures arising from the deflection potential alone, such as those considered in the context of CMB lensing reconstruction (and producing a leading trispectrum contribution).

**Table 1.** Results for the amplitude of the lensing bispectrum $f_{\rm NL}^{\rm lens}$ from the SMICA, SEVEM, NILC, and Commander foreground-cleaned CMB maps, for different bispectrum estimators.

| | Lensing amplitude | | | |
|---|---|---|---|---|
| Estimator | SMICA | SEVEM | NILC | Commander |
| *T* | | | | |
| Binned . . . | 0.64 ± 0.33 | 0.42 ± 0.33 | 0.65 ± 0.33 | 0.45 ± 0.33 |
| Modal 1 . . | 0.74 ± 0.33 | 0.59 ± 0.32 | 0.72 ± 0.32 | 0.54 ± 0.33 |
| Modal 2 . . | 0.73 ± 0.27 | 0.61 ± 0.27 | 0.73 ± 0.27 | 0.62 ± 0.27 |
| *T*+*E* | | | | |
| Binned . . . | 0.81 ± 0.27 | 0.62 ± 0.27 | 0.77 ± 0.27 | 0.67 ± 0.27 |
| Modal 1 . . | 0.90 ± 0.26 | 0.82 ± 0.25 | 0.83 ± 0.25 | 0.73 ± 0.26 |

**Notes.** Uncertainties are 68% CL.

where the $X_i$ are either $T$ or $E$. The tilde on $\tilde{C}_\ell^{X_iX_j}$ indicates that it is the lensed power spectrum, while $C_\ell^{T\phi}$ and $C_\ell^{E\phi}$ are the temperature/polarization-lensing potential cross-power spectra. The functions $f^{T,E}_{\ell_1\ell_2\ell_3}$ are defined by

$$f^T_{\ell_1\ell_2\ell_3} = \frac{1}{2}[\ell_2(\ell_2+1)+\ell_3(\ell_3+1)-\ell_1(\ell_1+1)],$$
$$f^E_{\ell_1\ell_2\ell_3} = \frac{1}{2}[\ell_2(\ell_2+1)+\ell_3(\ell_3+1)-\ell_1(\ell_1+1)] \times \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^{-1}, \quad (38)$$

if the sum $\ell_1+\ell_2+\ell_3$ is even and $\ell_1, \ell_2, \ell_3$ satisfies the triangle inequality, and zero otherwise. Unlike for all other templates, the amplitude parameter $f_{\rm NL}^{\rm lens}$ is not unknown, but should be exactly equal to 1 in the context of the assumed ΛCDM cosmology.

The results for $f_{\rm NL}^{\rm lens}$ can be found in Table 1. Error bars have been determined based on FFP10 simulations. As we have seen in the previous releases, the results for $T$-only are on the low side. SMICA and NILC have remained stable compared to 2015 and are marginally consistent with the expected value at the $1\sigma$ level. SEVEM and Commander, on the other hand, have both decreased compared to 2015 and are now further than $1\sigma$ away from unity. However, when polarization is added all results increase and become mostly consistent with unity at the $1\sigma$ level. Using the SMICA map (which we often focus on in the rest of the paper, see Sect. 6 for discussion) and the Modal 1 estimator (because it is one of the two estimators for which the lensing template has been implemented in both $T$ and $E$ and the Binned estimator is slightly less well-suited for this particular shape, since it is a difficult template to bin), we conclude that we have a significant detection of the lensing bispectrum; the hypothesis of having no lensing bispectrum is excluded at $3.5\sigma$ using the full temperature and polarization data.

It was pointed out in Hill (2018) that the coupling between the ISW effect and the thermal Sunyaev-Zeldovich (tSZ) effect can produce a significant ISW-tSZ-tSZ temperature bispectrum, which peaks for squeezed modes, and can therefore contaminate especially our local and lensing results. The semi-analytic approach in Hill (2018) makes predictions for single frequency channels, and cannot be directly applied to the multi-frequency component-separated data we are using. However, it is interesting that such an approach shows an anti-correlation at all frequencies between the ISW-tSZ-tSZ contamination and the lensing bispectrum shape. This could be a possible explanation for the slightly low $T$-only value of $f_{\rm NL}^{\rm lens}$ observed in the final data, across all component-separated maps. We therefore investigate this issue further, by measuring the amplitude of the lensing bispectrum using SMICA maps in which the tSZ signal has been subtracted in addition to the usual components. The results for this "no-SZ" map are given in Table 2. It is clear that the results have increased and are now closer to unity. Using these values, the hypothesis of having no lensing bispectrum is excluded at $3.8\sigma$ using the full temperature and polarization data, and the Modal 1 estimator.

**Table 2.** Results for the amplitude $f_{\rm NL}^{\rm lens}$ of the lensing bispectrum from the SMICA no-SZ temperature map, combined with the standard SMICA polarization map, for the Binned and Modal 1 bispectrum estimators.

| Estimator | Lensing amplitude |
|---|---|
| *T* | |
| Binned . . . . . | 0.83 ± 0.35 |
| Modal 1 . . . . . | 0.90 ± 0.34 |
| *T*+*E* | |
| Binned . . . . . | 0.90 ± 0.28 |
| Modal 1 . . . . . | 1.03 ± 0.27 |

**Notes.** Uncertainties are 68% CL.

The hypothesis that ISW-tSZ-tSZ residuals are contributing to the temperature-only lensing bispectrum amplitude result is reinforced by the fact that our $f_{\rm NL}^{\rm lens}$ measurements from $T$+$E$ are systematically closer to 1 (polarization does not correlate with ISW and helps in debiasing the result). The SZ-removed (hereafter "no-SZ") SMICA measurements of local $f_{\rm NL}$, however, do not support this hypothesis, since they do not display any large shift, whereas a residual ISW-tSZ-tSZ bispectrum should correlate with all shapes that peak in the squeezed limit. Still, one cannot exclude the possibility that multi-frequency component-separation affects the two cases differently. Both our local no-SZ results and a further discussion of this effect are presented in Sect. 6.3.2, where the impact on other primordial shapes is also evaluated and comparisons with simulations are carried out. The final conclusion is that the evidence for ISW-tSZ-tSZ contamination on the temperature-only lensing bispectrum measurements is not very strong, but the possibility cannot be ruled out.

We also note here that another potentially important source of contamination for the local and lensing shapes is given by the coupling between lensing and the CIB. However, the frequency-by-frequency analysis in Hill (2018) shows that in this case the expected bias is positive at all frequencies. The systematically low values of $f_{\rm NL}^{\rm lens}$ observed in temperature seem therefore to indicate that the CIB bispectrum contamination does not leak into the final component-separated maps, at least not at a level that is significant for our analysis. This is further reinforced by the fact that we do not detect any CIB signal directly in the cleaned maps (see Sect. 4.2).

In this paper our main concern with the lensing bispectrum is its influence on the primordial shapes. The bias due to the lensing bispectrum on the estimation of the $f_{\rm NL}$ parameter of another shape $S$ is given by the inner product of the lensing bispectrum (Eq. (37)) with the bispectrum of that shape $S$, divided by the inner product of the bispectrum $S$ with itself (see PCNG15, for more details, or e.g. Bucher et al. 2016 for a derivation). The values for the bias, as computed from theory, are given in Table 3. Note that the bias values that can be read off from, for example, Table 5 in Sect. 5 can differ slightly from these, because each estimator uses values computed using the approximations appropriate to the estimator. However, those differences are insignificant compared to the uncertainties. As seen already in the previous releases, for $T$-only data and for $T$+$E$ the bias

**Table 3.** Bias in the three primordial $f_{\rm NL}$ parameters due to the lensing signal for the four component-separation methods.

| Shape | Lensing $f_{\rm NL}$ bias | | | |
|---|---|---|---|---|
| | SMICA | SEVEM | NILC | Commander |
| $T$ Local . . . . . | 7.3 | 7.2 | 7.3 | 7.3 |
| $T$ Equilateral . . . . . | −0.7 | −0.6 | −0.7 | −0.7 |
| $T$ Orthogonal . . . . . | −23 | −23 | −23 | −23 |
| $E$ Local . . . . . | 0.5 | 0.5 | 0.5 | 0.5 |
| $E$ Equilateral . . . . . | 0.5 | 0.5 | 0.5 | 0.5 |
| $E$ Orthogonal . . . . . | −0.7 | −0.7 | −0.7 | −0.7 |
| $T$+$E$ Local . . . . . | 5.0 | 5.0 | 5.0 | 5.0 |
| $T$+$E$ Equilateral . . . | 1.0 | 1.1 | 1.1 | 1.0 |
| $T$+$E$ Orthogonal . . . | −9.1 | −9.4 | −9.2 | −9.4 |

is very significant for local and to a lesser extent for orthogonal NG. For $T$-only local NG the bias is even larger than the error bars on $f_{\rm NL}$. Hence it is quite important to take this bias into account. On the other hand, for $E$-only the effect is completely negligible.

Lastly, we would like to point out that lensing can also contribute to the covariance (Babich & Zaldarriaga 2004). To lowest order, as was done in this analysis, the spectra in Eq. (27) are replaced with the lensed spectra. However, it was shown by Babich & Zaldarriaga (2004), and later confirmed by Lewis et al. (2011), that the contribution from the connected four-point function induced by lensing can become an important contribution to the covariance. Although the analysis here has not shown the effect of lensing to be large (since tests were performed and analysis was done on lensed simulations with no obvious degradation over the Gaussian Fisher errors), it is expected that this will become an important challenge in CMB non-Gaussianity constraints beyond *Planck*. Furthermore, unlike the signal contamination discussed above, this will likely be equally important for polarization. Because of the shape of the lensing-induced covariance, which tends to be largest for squeezed configurations, the local shape will be most affected. The estimates performed by Babich & Zaldarriaga (2004) were done in the flat-sky limit and did not include polarization; in addition, computations were truncated at linear order in the potential. Although the latter seems justified, it was later shown that for the lensing power spectrum covariance the linear contribution is actually subdominant to the quadratic contribution (Peloton et al. 2017). For future CMB analysis it will be important to address these open questions. One obvious solution to the extra covariance would be to delens the maps (Green et al. 2017) before applying the estimators. Interestingly, this would automatically remove some of the signal-induced lensing contributions discussed above.

### 4.2. Non-Gaussianity from extragalactic point sources

As seen in the previous releases, extragalactic point sources are a contaminant present in the bispectrum as measured by *Planck*. They are divided into populations of unclustered and clustered sources. The former are radio and late-type infrared galaxies (see e.g. Toffolatti et al. 1998; González-Nuevo et al. 2005), while the latter are primarily dusty star-forming galaxies constituting the cosmic infrared background (CIB; Lagache et al. 2005). For both types of point sources analytic (heuristic) bispectrum templates have been determined, which can be fitted jointly with the primordial NG templates to deal with the contamination.

**Table 4.** Joint estimates of the bispectrum amplitudes of unclustered and clustered point sources in the cleaned *Planck* temperature maps, determined with the Binned bispectrum estimator.

| Map | $b_{\rm PS}/(10^{-29})$ | $A_{\rm CIB}/(10^{-27})$ |
|---|---|---|
| SMICA . . . . . | 4.7 ± 2.6 | 0.8 ± 1.3 |
| SEVEM . . . . . | 7.0 ± 2.8 | 1.4 ± 1.4 |
| NILC . . . . . | 5.2 ± 2.7 | 0.3 ± 1.3 |
| Commander . . . . | 3.4 ± 2.6 | 1.1 ± 1.3 |

**Notes.** Uncertainties are 68 % CL.

The templates used here are the same as those used in PCNG15 (see that paper for more information and references). The reduced angular bispectrum template of the unclustered sources is (Komatsu & Spergel 2001)

$$b^{\rm unclust}_{\ell_1\ell_2\ell_3} = b_{\rm PS} = \text{constant}. \tag{39}$$

This template is valid in polarization as well as temperature. However, we do not detect all the same point sources in polarization as we detected in temperature, so that a full $T$+$E$ analysis does not make sense for this template. In fact, there is no detection of unclustered point sources in the cleaned *Planck* polarization map, so that we do not include the $E$-only values in the table either. The reduced angular bispectrum template for the clustered sources, for example the CIB, is (Lacasa et al. 2014; Pénin et al. 2014)

$$b^{\rm CIB}_{\ell_1\ell_2\ell_3} = A_{\rm CIB}\left[\frac{(1+\ell_1/\ell_{\rm break})(1+\ell_2/\ell_{\rm break})(1+\ell_3/\ell_{\rm break})}{(1+\ell_0/\ell_{\rm break})^3}\right]^q, \tag{40}$$

where the index is $q = 0.85$, the break is located at $\ell_{\rm break} = 70$, and $\ell_0 = 320$ is the pivot scale for normalization. This template is valid only for temperature; the CIB is negligibly polarized.

The results for both extragalactic point source templates, as determined by the Binned bispectrum estimator applied to the *Planck* temperature map cleaned with the four component-separation methods, can be found in Table 4. Because the two templates are highly correlated (93%), the results have been determined through a joint analysis. Contamination from unclustered sources is detected in all component-separated maps, although at different levels, with SEVEM having the largest contamination. The CIB bispectrum, on the other hand, is not detected in a joint analysis. Both point-source templates are negligibly correlated with the primordial NG templates and the lensing template (all well below 1% for the unclustered point sources). For this reason, and despite the detection of unclustered point sources in the cleaned maps, it makes no difference for the primordial results in the next sections if point sources are included in a joint analysis or completely neglected.

## 5. Results

### 5.1. Constraints on local, equilateral, and orthogonal $f_{\rm NL}$

We now describe our analysis of the standard local, equilateral, and orthogonal shapes. We employ the four bispectrum estimators described in Sects. 3.1.1, 3.1.3, and 3.1.4 on the temperature and polarization maps generated by the SMICA, SEVEM, NILC, and Commander component-separation pipelines. Further details about our data analysis setup were provided in Sect. 3.2.2. As explained there, the main novelty, compared to the 2015 release, is the use of polarization multipoles in the range $4 \le \ell < 40$, which were previously excluded.

**Table 5.** Results for the $f_{\rm NL}$ parameters of the primordial local, equilateral, and orthogonal shapes, determined by the KSW, Binned and Modal estimators from the SMICA, SEVEM, NILC, and Commander foreground-cleaned maps.

| | Independent | | | Lensing subtracted | | | |
|---|---|---|---|---|---|---|---|
| Shape | KSW | Binned | Modal 1 | KSW | Binned | Modal 1 | Modal 2 |
| SMICA $T$ | | | | | | | |
| Local | 6.7 ± 5.6 | 6.8 ± 5.6 | 6.3 ± 5.8 | −0.5 ± 5.6 | −0.1 ± 5.6 | −0.6 ± 5.8 | −0.6 ± 6.4 |
| Equilateral | 6 ± 66 | 27 ± 69 | 28 ± 64 | 7 ± 66 | 26 ± 69 | 24 ± 64 | 34 ± 67 |
| Orthogonal | −38 ± 36 | −37 ± 39 | −29 ± 39 | −15 ± 36 | −11 ± 39 | −4 ± 39 | −26 ± 43 |
| SMICA $E$ | | | | | | | |
| Local | 48 ± 28 | 49 ± 26 | 44 ± 25 | 47 ± 28 | 48 ± 26 | 44 ± 25 | |
| Equilateral | 170 ± 161 | 170 ± 140 | 190 ± 160 | 169 ± 161 | 170 ± 140 | 200 ± 160 | |
| Orthogonal | −209 ± 86 | −180 ± 83 | −210 ± 85 | −208 ± 86 | −180 ± 83 | −220 ± 85 | |
| SMICA $T$+$E$ | | | | | | | |
| Local | 4.1 ± 5.1 | 2.2 ± 5.0 | 4.6 ± 4.7 | −0.9 ± 5.1 | −2.5 ± 5.0 | −0.1 ± 4.7 | −2.0 ± 5.0 |
| Equilateral | −17 ± 47 | −16 ± 48 | −6 ± 48 | −18 ± 47 | −19 ± 48 | −8 ± 48 | −4 ± 43 |
| Orthogonal | −46 ± 23 | −45 ± 24 | −37 ± 24 | −37 ± 23 | −34 ± 24 | −28 ± 24 | −40 ± 24 |
| SEVEM $T$ | | | | | | | |
| Local | 5.0 ± 5.6 | 6.8 ± 5.7 | 6.2 ± 6.0 | −2.3 ± 5.6 | 0.0 ± 5.7 | −1.6 ± 6.0 | 0.0 ± 6.5 |
| Equilateral | 16 ± 66 | 45 ± 70 | 39 ± 64 | 17 ± 66 | 43 ± 70 | 33 ± 64 | 34 ± 68 |
| Orthogonal | 2 ± 37 | −17 ± 39 | −5 ± 40 | 24 ± 37 | 9 ± 39 | 25 ± 40 | −14 ± 43 |
| SEVEM $E$ | | | | | | | |
| Local | 38 ± 29 | 56 ± 29 | 55 ± 22 | 38 ± 29 | 55 ± 29 | 37 ± 22 | |
| Equilateral | 174 ± 166 | 230 ± 160 | 250 ± 160 | 173 ± 166 | 230 ± 160 | 260 ± 160 | |
| Orthogonal | −182 ± 88 | −140 ± 88 | −160 ± 87 | −181 ± 88 | −140 ± 88 | −170 ± 87 | |
| SEVEM $T$+$E$ | | | | | | | |
| Local | 3.3 ± 5.1 | 3.6 ± 5.2 | 7.8 ± 4.7 | −1.7 ± 5.1 | −1.2 ± 5.2 | 1.8 ± 4.7 | 1.5 ± 5.1 |
| Equilateral | −8 ± 47 | 2 ± 48 | 5 ± 51 | −9 ± 47 | −1 ± 48 | 4 ± 51 | 23 ± 45 |
| Orthogonal | −24 ± 23 | −28 ± 24 | −19 ± 24 | −15 ± 23 | −18 ± 24 | −6 ± 24 | −29 ± 25 |
| NILC $T$ | | | | | | | |
| Local | 6.8 ± 5.7 | 6.8 ± 5.6 | 6.2 ± 5.9 | −0.4 ± 5.7 | 0.0 ± 5.6 | −1.7 ± 5.9 | −0.5 ± 6.6 |
| Equilateral | −10 ± 66 | 6 ± 69 | 12 ± 61 | −10 ± 66 | 5 ± 69 | 6 ± 61 | 20 ± 67 |
| Orthogonal | −23 ± 36 | −21 ± 39 | −12 ± 40 | −1 ± 36 | 4 ± 39 | 17 ± 40 | −10 ± 43 |
| NILC $E$ | | | | | | | |
| Local | 9 ± 30 | 22 ± 28 | 20 ± 22 | 9 ± 30 | 21 ± 28 | 15 ± 22 | |
| Equilateral | 39 ± 163 | 60 ± 150 | 65 ± 150 | 38 ± 163 | 59 ± 150 | 66 ± 150 | |
| Orthogonal | −133 ± 88 | −120 ± 85 | −150 ± 83 | −133 ± 88 | −120 ± 85 | −160 ± 83 | |
| NILC $T$+$E$ | | | | | | | |
| Local | 3.9 ± 5.1 | 2.1 ± 5.1 | 3.9 ± 4.7 | −1.1 ± 5.1 | −2.7 ± 5.1 | −1.5 ± 4.7 | −2.4 ± 5.1 |
| Equilateral | −30 ± 46 | −29 ± 47 | −22 ± 49 | −31 ± 46 | −32 ± 47 | −24 ± 49 | −15 ± 43 |
| Orthogonal | −33 ± 23 | −31 ± 23 | −26 ± 24 | −24 ± 23 | −21 ± 23 | −14 ± 24 | −24 ± 24 |
| Commander $T$ | | | | | | | |
| Local | 4.9 ± 5.6 | 5.6 ± 5.6 | 4.5 ± 5.9 | −2.3 ± 5.6 | −1.3 ± 5.6 | −3.0 ± 5.9 | −1.9 ± 6.6 |
| Equilateral | 14 ± 66 | 33 ± 69 | 33 ± 62 | 15 ± 66 | 32 ± 69 | 25 ± 62 | 36 ± 68 |
| Orthogonal | 3 ± 37 | 3 ± 39 | 14 ± 40 | 25 ± 37 | 29 ± 39 | 42 ± 40 | 5 ± 43 |
| Commander $E$ | | | | | | | |
| Local | 31 ± 29 | 43 ± 27 | 31 ± 21 | 31 ± 29 | 42 ± 27 | 27 ± 21 | |
| Equilateral | 163 ± 167 | 170 ± 150 | 190 ± 160 | 162 ± 167 | 170 ± 150 | 180 ± 160 | |
| Orthogonal | −179 ± 88 | −160 ± 85 | −180 ± 85 | −178 ± 88 | −160 ± 85 | −190 ± 85 | |
| Commander $T$+$E$ | | | | | | | |
| Local | 3.0 ± 5.1 | 2.3 ± 5.1 | 3.4 ± 4.6 | −2.0 ± 5.1 | −2.5 ± 5.1 | −1.4 ± 4.6 | −1.7 ± 5.1 |
| Equilateral | −9 ± 47 | −6 ± 48 | 5 ± 50 | −10 ± 47 | −9 ± 48 | 2 ± 50 | 35 ± 44 |
| Orthogonal | −23 ± 23 | −23 ± 24 | −14 ± 24 | −13 ± 23 | −12 ± 24 | −2 ± 24 | −21 ± 25 |

**Notes.** Results have been determined using an independent single-shape analysis and are reported both without (first set of columns) and with (second set of columns) subtraction of the lensing bias. Uncertainties are 68 % CL.

Our final results are summarized in Table 5, while data validation tests will be presented in Sect. 6. As in 2015, we show final $f_{\rm NL}$ estimates for $T$-only, $E$-only, and for the full $T$+$E$ data set, with and without subtraction of the lensing bias. When we subtract the lensing bias, we assume a theoretical prediction for the lensing bispectrum amplitude based on the *Planck* best-fit ΛCDM cosmological parameters (see Sect. 4.1). We note that propagating uncertainties in the parameters has a negligible effect on the predicted lensing bispectrum, and the validity of the ΛCDM assumption is of course consistent with

**Table 6.** Results for the $f_{\rm NL}$ parameters of the primordial local, equilateral, and orthogonal shapes, determined by the KSW estimator from the SMICA foreground-cleaned map.

| Shape | Independent | Lensing subtracted |
|---|---|---|
| | SMICA $T$ | |
| Local | $6.7 \pm 5.6$ | $-0.5 \pm 5.6$ |
| Equilateral | $4 \pm 67$ | $5 \pm 67$ |
| Orthogonal | $-38 \pm 37$ | $-15 \pm 37$ |
| | SMICA $T$+$E$ | |
| Local | $4.1 \pm 5.1$ | $-0.9 \pm 5.1$ |
| Equilateral | $-25 \pm 47$ | $-26 \pm 47$ |
| Orthogonal | $-47 \pm 24$ | $-38 \pm 24$ |

**Notes.** Both independent single-shape results and results with the lensing bias subtracted are reported; uncertainties are 68 % CL. The difference between this table and the corresponding values in the previous table is that here the equilateral and orthogonal shapes have been analysed jointly.

all *Planck* measurements. An alternative to the direct subtraction of the predicted lensing bias of the primordial shapes would be performing a full joint bispectrum analysis (accounting for the measured lensing bispectrum amplitude and propagating the uncertainty into the final primordial NG error budget). While the latter approach is in principle more conservative, we opt for the former both for simplicity and for consistency with previous analyses, as it turns out that the difference between the two methods has no significant impact on our results. If, as an example, we perform a $T$-only joint bispectrum analysis using the SMICA map and the Binned estimator, we obtain ${f_{\rm NL}}^{\rm local} = 2.7 \pm 5.7$ and ${f_{\rm NL}}^{\rm lens} = 0.60 \pm 0.34$. The joint $T$+$E$ analysis produces instead ${f_{\rm NL}}^{\rm local} = -1.7 \pm 5.2$ and ${f_{\rm NL}}^{\rm lens} = 0.82 \pm 0.27$. Hence the uncertainties obtained from the joint analysis are practically stable with respect to those shown in Table 5, while the small shift in the local central value (due to the low value of the measured ${f_{\rm NL}}^{\rm lens}$) would not change our conclusions in any way. More details about the lensing contribution are provided in Sect. 4, where we also discuss the negligible impact of point source contamination on primordial bispectra.

Table 5 constitutes the most important result of this section. As done in 2013 and 2015, we select the KSW estimator and the SMICA map to provide the final *Planck* results for the local, equilateral, and orthogonal bispectra; these results are summarized in Table 6. The motivations for this choice are as in the past: SMICA performs well in all validation tests and shows excellent stability across different data releases; the KSW estimator, while not able to deal with non-separable shapes or reconstruct the full bispectrum, can treat exactly the local, equilateral, and orthogonal templates that are analysed here.

As for the two previous data releases, we note that the agreement between the different estimators – for all the maps and all the shapes considered, both in temperature and polarization, as well as in the full $T$+$E$ results – is well in line with both our theoretical expectations and Monte Carlo studies, where we found an *average* measured $f_{\rm NL}$ scatter at the level of $\lesssim \sigma_{f_{\rm NL}}/3$ between different pipelines (this was discussed at length in PCNG13 and PCNG15). The observed scatter between two estimators (for a given realization) can be larger than what the very small reported differences in their variance might suggest. In an ideal, noiseless, full-sky experiment, the observed scatter is only due to differences in the estimator weights, coming from the use of different bispectrum expansions or binning schemes. In such an ideal case, if two different bispectrum templates have a correlation coefficient $r$, we showed in Appendix B of PCNG13 that the standard deviation of the expected scatter $\delta f_{\rm NL}$ is given by $\sigma_{\delta f_{\rm NL}} = \sigma_{f_{\rm NL}} \sqrt{(1-r^2)}/r$. This leads to differences between $f_{\rm NL}$ results that are a sizeable fraction of the estimator standard deviation, even for highly correlated weights. To be more explicit, a 95% correlation between different input templates leads, using the formula above, to a $\sigma_{f_{\rm NL}}/3$ average scatter in $f_{\rm NL}$ estimates; on the other hand, it would produce only a 5% difference in the final uncertainties.

Considering a more concrete example, let us focus on the current SMICA KSW $T$+$E$ results, which we quote as our final, recommended bounds. Let us consider, for example, the difference between the KSW and Binned estimator for the local shape. This is relatively large, at approximately $0.3\sigma_{f_{\rm NL}}$ after lensing-bias subtraction, compared to an error bar difference of 2%. Let us now assume a correlation $r \approx 0.98$ between the weights, consistent both with what we see in simulations and with such a small difference in the errors. If we substitute this into our formula, we obtain an expected scatter of $\sigma_{\delta f_{\rm NL}} = 0.2\sigma_{f_{\rm NL}}$. The observed scatter is then $1.5\sigma$ away from this average and therefore fairly consistent with it. The chosen example displays a relatively large difference, compared to all the other combinations that can be built in Table 5; note also that the formula we are using represents an ideal case, and our validation tests on simulations in realistic conditions (e.g. masking and anisotropic noise) show, as expected, that the actual scatter between two pipelines is generally a bit larger than this ideal expectation (PCNG13; PCNG15).

If, instead of comparing central values, we look at uncertainties, we see as expected that all pipelines produce nearly optimal constraints. For the local shape, we see that the Modal 1 pipeline produces 6% smaller errors than for example KSW; however, this is within the expected Monte Carlo error and it seems to be just an effect of the selected simulation sample used to compute $\sigma_{f_{\rm NL}}$. One should also consider that Modal 1 uses 160 FFP10 maps to extract the standard deviation, versus 300 maps for the other pipelines. This explanation is confirmed by our many validation tests on different sets of simulations.

A high level of internal consistency is also displayed between $f_{\rm NL}$ estimates obtained from different component-separated maps, as well as in the comparison of current results with those from previous releases. One small exception is provided by the orthogonal $f_{\rm NL}$ estimate obtained from Commander $T$-only data, for which we notice both a larger fluctuation with respect to 2013 and 2015 results and a larger discrepancy with other foreground cleaned maps, in particular with the SMICA one. However, we do not find this worrisome for a number of reasons. First of all, the SMICA – Commander difference is still at the level of the $1\sigma$ orthogonal $f_{\rm NL}$ uncertainty and all methods, including Commander, show full consistency with $f_{\rm NL}^{\rm ortho} = 0$. Therefore, this discrepancy does not pose any problem for the theoretical interpretation of the result. Moreover, this fluctuation completely goes away when accounting for polarization data, the reliability of which has become significantly higher with respect to our previous analysis (see Sect. 6 for details). Finally, the discrepancy is for a very specific shape and it is entirely driven by the already-noted fluctuation in the Commander orthogonal result with respect to 2013 and 2015. The other methods remain stable, in particular SMICA, which we take as the map of choice for our final results. The observed fluctuation in orthogonal $f_{\rm NL}$ from Commander can likely be explained by the unavailability of "detset" (i.e. detector-subset) maps for this release, which constituted in the past a useful input for improving the accuracy

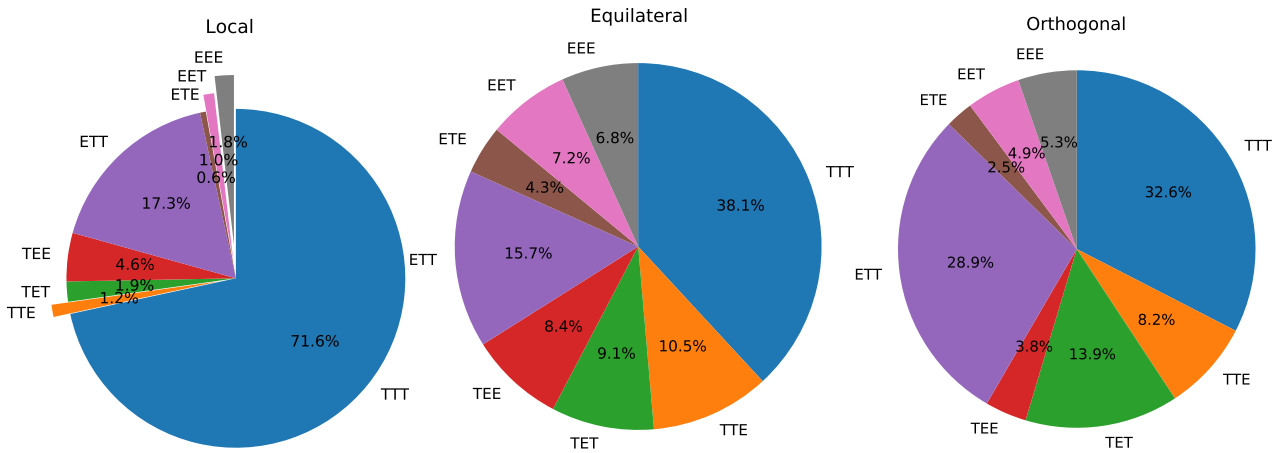


**Fig. 1.** Weights of each polarization configuration going into the total value of $f_{\rm NL}$ for, *from left to right*, local, equilateral, and orthogonal shapes. Since we impose $\ell_1 \leq \ell_2 \leq \ell_3$, there is a difference between, *TEE* for example (smallest $\ell$ is temperature) and *EET* (largest $\ell$ is temperature).

of the `Commander` map. It is, however, important to stress that `Commander` itself shows excellent agreement with other methods, when measuring $f_{\rm NL}$ for all other shapes and also when correlating the bispectrum modes and bins in a model-independent fashion, both in temperature and polarization (again see Sect. 6 for a complete discussion of these tests).

Comparing the uncertainties in Table 5 to those in the corresponding table in the 2015 analysis paper (PCNG15), and focussing on the ones for the local shape, since those are most sensitive to low-$\ell$ modes, we see the following: for $T$-only data the errors are approximately equal on average, slightly better in some cases, and slightly worse in others. A possible explanation for the slightly larger errors could be the fact that the realism of the simulations (both in temperature and polarization) has improved from FFP8 used in 2015 to FFP10 used here. For our default `SMICA` component-separation method this leads to a slightly higher noise power spectrum of a few percent in the regions where noise is relevant. So the errors in 2015 might actually have been slightly underestimated. However, the differences are small enough that they could just be random fluctuations, especially given that not all estimators show the same effect. For $E$-only data, on the other hand, we see a clear improvement of the errors for all estimators. That is as expected, since we are now including all the additional polarization modes with $4 \leq \ell < 40$ in the analysis, and the local shape is quite sensitive to these low-$\ell$ modes. Finally, for the full $T$+$E$ analysis, we see similar results as for $T$-only: all errors have remained the same, to within fluctuations of around a few percent, at most. So one might wonder why the improvement in the $E$-only analysis has not translated into a corresponding improvement in the $T$+$E$ analysis. The answer is relatively simple: the *EEE*-bispectra only have a very small contribution to the final $T$+$E$ analysis, as shown in Fig. 1. This figure also shows that the equilateral and orthogonal shapes are more sensitive to polarization modes than the local shape, which explains why the errors for the former improve more when going from $T$-only to the full $T$+$E$ analysis than the errors for the latter.

In conclusion, our current results show no evidence for non-Gaussianity of the local, equilateral, or orthogonal type and are in very good agreement with the previous 2013 and 2015 analyses. We also show in Sect. 6 that the overall robustness and internal consistency of the polarization data set has significantly improved, as far as primordial non-Gaussian measurements are concerned.

### 5.2. Further bispectrum shapes

#### 5.2.1. Isocurvature non-Gaussianity

In this section we present a study of the isocurvature NG in the *Planck* 2018 `SMICA` map using the Binned bispectrum estimator. This analysis is complementary to the one based on the power spectrum presented in Planck Collaboration X (2020). The underlying modelling approach was discussed in Sect. 2.3, and as explained there, we only investigate isocurvature NG of the local type, and in addition always consider the adiabatic mode together with only one isocurvature mode, there we consider separately CDM-density, neutrino-density, and neutrino-velocity isocurvature. In that case there are in general six different $f_{\rm NL}$ parameters: the purely adiabatic one (a,aa) from Sect. 5.1; a purely isocurvature one (i,ii); and four mixed ones.

The results can be found in Table 7, both for an independent analysis of the six parameters (i.e. assuming that only one of the six parameters is present) and for a fully joint analysis (i.e. assuming that all six parameters are present, which is clearly the correct thing to do in the correlated framework described above). The reason for the very differently sized errors is a combination of two effects. The first is a simple normalization issue, due to switching from the more natural $\zeta$ and $S$ variables (commonly used in the inflation literature) to $\Phi_{\rm adi} = -3\zeta/5$ and $\Phi_{\rm iso} = -S/5$, which are more commonly used in the CMB literature[8]. The second is that certain parameters depend more on the high-$\ell$ adiabatic modes (which are well-determined), while others are dominated by the much suppressed, and hence unconstrained, high-$\ell$ isocurvature modes. For example, for the joint CDM $T$+$E$ case, when compensating for the normalization factor, one would find the error for a,aa and a,ai to be around 10, and the other four errors to be around 200 (see Langlois & van Tent 2012, for further discussion of these effects).

As in our analysis of the 2015 *Planck* data (PCNG15), we see no clear sign of any isocurvature NG. There are a few values that deviate from zero by up to about $2.5\sigma$, but such a small deviation cannot be considered a detection, given the large number of tests and the fact that the deviations are not consistent between $T$-only and $T$+$E$ data. For example, looking at the 300 Gaussian simulations that were used to determine the linear correction and the uncertainties, we find that 84 of them have at least one

[8] Conversion factors to obtain results based on $\zeta$ and $S$ are $-6/5$, $-2/5$, $-2/15$, $-18/5$, $-6/5$, and $-2/5$, for the six modes, respectively.

**Table 7.** Results for $f_{\rm NL}$ for local isocurvature NG determined from the SMICA *Planck* 2018 map with the Binned bispectrum estimator.

| | Independent | | | Joint | | |
|---|---|---|---|---|---|---|
| Shape | Cold dark matter | Neutrino density | Neutrino velocity | Cold dark matter | Neutrino density | Neutrino velocity |
| $T$ a,aa . . . . . | −0.1 ± 5.6 | −0.1 ± 5.6 | −0.1 ± 5.6 | 15 ± 16 | −15 ± 54 | −29 ± 49 |
| $T$ a,ai . . . . . | −3 ± 11 | −6 ± 16 | 27 ± 29 | −30 ± 31 | 87 ± 220 | 310 ± 360 |
| $T$ a,ii . . . . . | 440 ± 950 | −150 ± 290 | 480 ± 360 | 11000 ± 8500 | −1800 ± 4500 | −1100 ± 3800 |
| $T$ i,aa . . . . . | 33 ± 53 | 0.3 ± 9.2 | −0.1 ± 4.9 | 95 ± 120 | 46 ± 110 | −27 ± 49 |
| $T$ i,ai . . . . . | 28 ± 67 | −9 ± 23 | 14 ± 21 | −1300 ± 1000 | −32 ± 670 | 98 ± 170 |
| $T$ i,ii . . . . . | 110 ± 280 | −130 ± 250 | 270 ± 230 | 2500 ± 2000 | −140 ± 2600 | −1000 ± 1400 |
| $E$ a,aa . . . . . | 48 ± 26 | 48 ± 26 | 48 ± 26 | 25 ± 42 | −60 ± 120 | 43 ± 130 |
| $E$ a,ai . . . . . | 140 ± 91 | 170 ± 98 | 72 ± 51 | 28 ± 150 | 330 ± 590 | 280 ± 370 |
| $E$ a,ii . . . . . | 100 ± 2500 | 2000 ± 1500 | 310 ± 460 | −8100 ± 4400 | −9400 ± 5500 | −6100 ± 2800 |
| $E$ i,aa . . . . . | 110 ± 77 | 75 ± 34 | 31 ± 21 | 43 ± 140 | 90 ± 170 | −94 ± 120 |
| $E$ i,ai . . . . . | 330 ± 180 | 210 ± 100 | 43 ± 35 | 830 ± 580 | 500 ± 890 | 460 ± 360 |
| $E$ i,ii . . . . . | 1800 ± 1200 | 1900 ± 1100 | 170 ± 230 | −1400 ± 2300 | −1300 ± 4800 | −380 ± 1000 |
| $T$+$E$ a,aa . . . . | −2.5 ± 5.0 | −2.5 ± 5.0 | −2.5 ± 5.0 | 4 ± 10 | −53 ± 28 | 2 ± 26 |
| $T$+$E$ a,ai . . . . | −10 ± 10 | −14 ± 14 | −1 ± 21 | −14 ± 21 | 160 ± 110 | 250 ± 110 |
| $T$+$E$ a,ii . . . . . | −450 ± 520 | −350 ± 260 | −56 ± 210 | −3100 ± 1500 | −4100 ± 1600 | −2100 ± 920 |
| $T$+$E$ i,aa . . . . | 20 ± 28 | −0.5 ± 8.0 | −3.5 ± 4.2 | 96 ± 52 | 44 ± 49 | −37 ± 26 |
| $T$+$E$ i,ai . . . . . | −32 ± 46 | −18 ± 20 | −7 ± 14 | 190 ± 180 | 350 ± 240 | 23 ± 77 |
| $T$+$E$ i,ii . . . . . | −290 ± 210 | −340 ± 210 | −51 ± 110 | −640 ± 400 | −2000 ± 990 | 38 ± 300 |

**Notes.** In each case the adiabatic mode is considered together with one isocurvature mode (either cold-dark-matter, neutrino-density, or neutrino-velocity isocurvature). As explained in the text this gives six different $f_{\rm NL}$ parameters, indicated by the different combinations of the adiabatic (a) and isocurvature (i) modes. Results are shown for both an independent and a fully joint analysis, for $T$-only, $E$-only, and full $T$+$E$ data. In all cases the lensing bias has been subtracted.

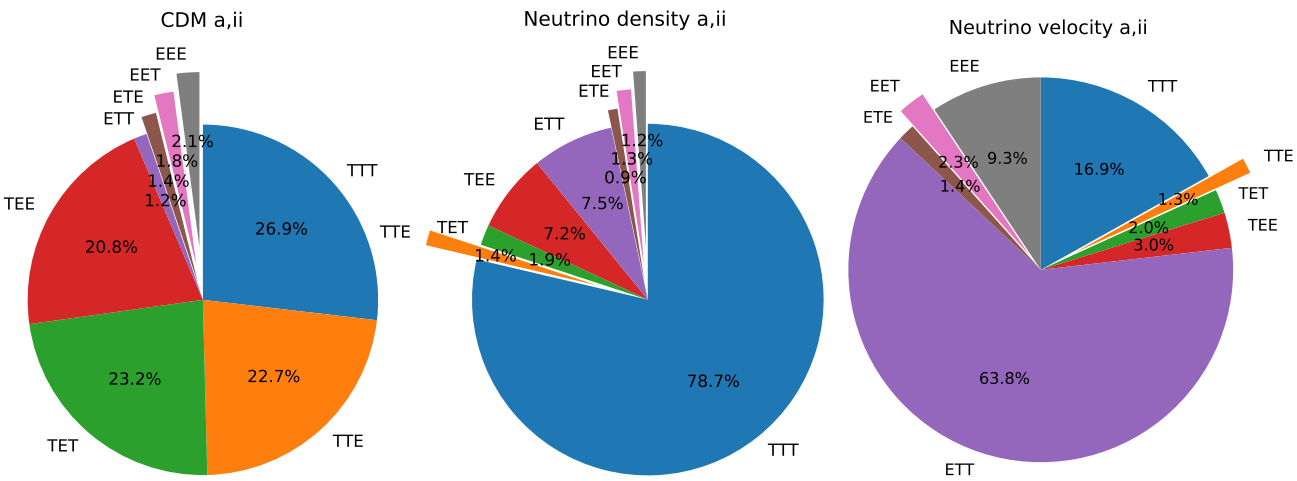


**Fig. 2.** Weights of each polarization configuration going into the total value of the a,ii mixed $f_{\rm NL}$ parameter for, *from left to right*, CDM, neutrino-density, and neutrino-velocity isocurvature, in addition to the adiabatic mode. Since we impose $\ell_1 \leq \ell_2 \leq \ell_3$, there is a difference between, *TEE* for example (where the smallest $\ell$ is temperature) and *EET* (where the largest $\ell$ is temperature).

>2.5$\sigma$ result in the two CDM columns of Table 7, while for neutrino density and neutrino velocity the numbers are 62 and 80, respectively.

We see that many constraints are tightened considerably when including polarization, by up to the predicted factor of about 5–6 for the CDM a,ii, i,ai, and i,ii modes in the joint analysis (e.g. −1300±1000 for $T$ only, decreasing to 190±180 for $T$+$E$ in the CDM i,ai case). Focussing now on the independent results, where it is easier to understand their behaviour (as things are not mixed together), we see that the uncertainties of some of the CDM and neutrino-velocity modes improve by a factor of about 2 when going from the $T$-only to the full $T$+$E$ analysis (e.g. neutrino velocity i,ii changes from 270±230 to −51±110), while the improvements for the neutrino-density modes are much smaller, of the order of what we see for the pure adiabatic mode. This can be explained if we look at the contribution of the various polarization modes to the total $f_{\rm NL}$ parameters, as indicated in Fig. 2 for the a,ii modes (to give an example). While the diagram for the neutrino-density mode very much resembles the one for the pure adiabatic local case in Fig. 1, those for CDM and neutrino velocity are quite different and have a much larger contribution from polarization modes (but quite different between the two cases). Hence isocurvature NG also provides a good test of the quality of the polarization maps, and a clear motivation for extending our analyses to include polarization in addition to temperature.

Comparing the results in Table 7 to those of the corresponding table in our 2015 analysis (PCNG15) we see in the first place that the $T$-only results are mostly very stable, generally

**Table 8.** Similar to Table 7, except that we now assume that the adiabatic and isocurvature modes are completely uncorrelated.

| | Independent | | | Joint | | |
|---|---|---|---|---|---|---|
| Shape | Cold dark matter | Neutrino density | Neutrino velocity | Cold dark matter | Neutrino density | Neutrino velocity |
| $T$ a,aa . . . . . | $-0.1 \pm 5.6$ | $-0.1 \pm 5.6$ | $-0.1 \pm 5.6$ | $-0.6 \pm 5.7$ | $15 \pm 13$ | $-1.3 \pm 5.8$ |
| $T$ i,ii . . . . . | $110 \pm 280$ | $-130 \pm 250$ | $270 \pm 230$ | $120 \pm 290$ | $-710 \pm 590$ | $280 \pm 240$ |
| $E$ a,aa . . . . . | $48 \pm 26$ | $48 \pm 26$ | $48 \pm 26$ | $41 \pm 27$ | $33 \pm 35$ | $62 \pm 35$ |
| $E$ i,ii . . . . . | $1800 \pm 1200$ | $1900 \pm 1100$ | $170 \pm 230$ | $1400 \pm 1200$ | $1000 \pm 1400$ | $-200 \pm 300$ |
| $T$+$E$ a,aa . . . . | $-2.5 \pm 5.0$ | $-2.5 \pm 5.0$ | $-2.5 \pm 5.0$ | $-1.3 \pm 5.1$ | $11.4 \pm 8.4$ | $-1.9 \pm 5.2$ |
| $T$+$E$ i,ii . . . . . | $-290 \pm 210$ | $-340 \pm 210$ | $-51 \pm 110$ | $-280 \pm 210$ | $-710 \pm 360$ | $-41 \pm 120$ |

**Notes.** Hence there are only two $f_{\rm NL}$ parameters in this case, a purely adiabatic one and a purely isocurvature one.

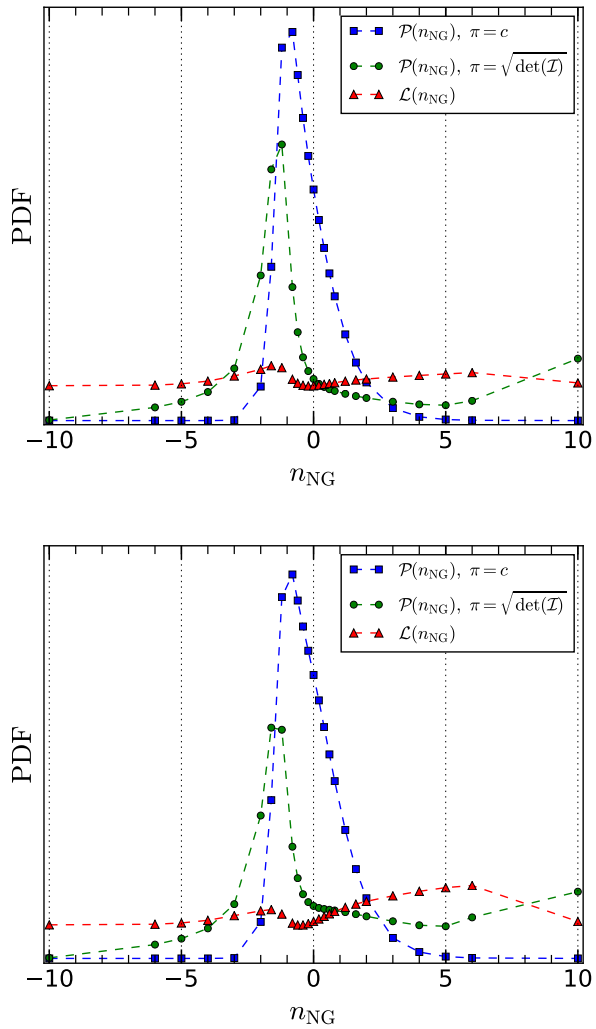


**Fig. 3.** PDF of the running parameter $n_{\rm NG}$ for the one-field local model. *Top*: SMICA map. *Bottom*: Commander map. Blue squares give the marginalized posterior assuming a constant prior, green circles are the posterior assuming a Jeffreys prior, and red triangles are the profiled Likelihood.

shifting by much less than $1\sigma$ (e.g. joint neutrino velocity i,ii is $-970 \pm 1400$ in 2015 versus $-1000 \pm 1400$ now). Secondly, for the $E$-only results we see that all the errors have significantly decreased (e.g. taking again joint neutrino velocity i,ii, we had $2200 \pm 1600$ in 2015 versus $-380 \pm 1000$ now), in line with the fact that we are now using the additional $4 \leq \ell < 40$ $E$ modes. We also generally see larger changes in the central values, with several shifts of $1\sigma$ or larger; nevertheless, the results are consistent with zero. Because we have added new polarization data to the analysis compared to 2015, and the quality of the polarization data has improved overall, these larger shifts are not unexpected. Finally, looking at the full $T$+$E$ analysis, we see that the central values shift a bit more than for $T$ only, but remain consistent with zero. For CDM and neutrino density the errors are marginally larger than in 2015, similarly to what we see with the Binned estimator for the local adiabatic shape (see Sect. 5.1). As shown in Fig. 2, both of these depend very little on low-$\ell$ polarization, and so are likely mostly driven by the marginal increase in the $T$-only errors. For the neutrino-velocity mode, on the other hand, some of the errors do decrease (e.g. $480 \pm 430$ in 2015 versus $38 \pm 300$ now for joint neutrino velocity i,ii). As seen in Fig. 2, this mode depends more strongly on the $ETT$ combination, which does involve low-$\ell$ polarization.

In the results so far we looked at the most general case, having a possible correlation between the isocurvature and adiabatic modes. However, if we assume that the adiabatic and the isocurvature modes have a cross-power spectrum of zero and are completely uncorrelated, then there are only two free $f_{\rm NL}$ parameters, the a,aa and the i,ii ones. In Table 8 we present the results for this uncorrelated case. The independent results are the same as in the previous table and have been repeated for convenience. The significant increase in the $T$-only and $T$+$E$ errors for the neutrino density case when going from the independent to the joint analysis clearly illustrates the fact that its bispectrum template has a large overlap with the adiabatic one, something that also explains the similarity of Figs. 1 and 2 for that case. The CDM and neutrino-velocity modes, on the other hand, have templates that are very different from the adiabatic one and their errors hardly increase (except for neutrino velocity for $E$-only data). Again there is no evidence for any isocurvature NG: we do not consider the almost $2\sigma$ result for the neutrino-density isocurvature mode in the $T$+$E$ joint analysis to be significant, although it will be interesting to keep an eye on this in future CMB experiments with even better polarization measurements.

### 5.2.2. Running non-Gaussianity

In this section we present our analysis of the scale-dependent bispectrum shapes described in Sect. 2.4; we obtain these results following the pipeline described in Sect. 3.1.2. In Figs. 3–5 we show the results, respectively, for the one-field local model (Eq. (8)), the two-field local model (Eq. (10)), and the geometric-mean equilateral model, where $f^*_{\rm NL}$ is parametrized as in Eq. (31). Here we present results for both the SMICA and Commander temperature maps. Since the $E$-mode polarization maps are not expected to significantly improve the constraints

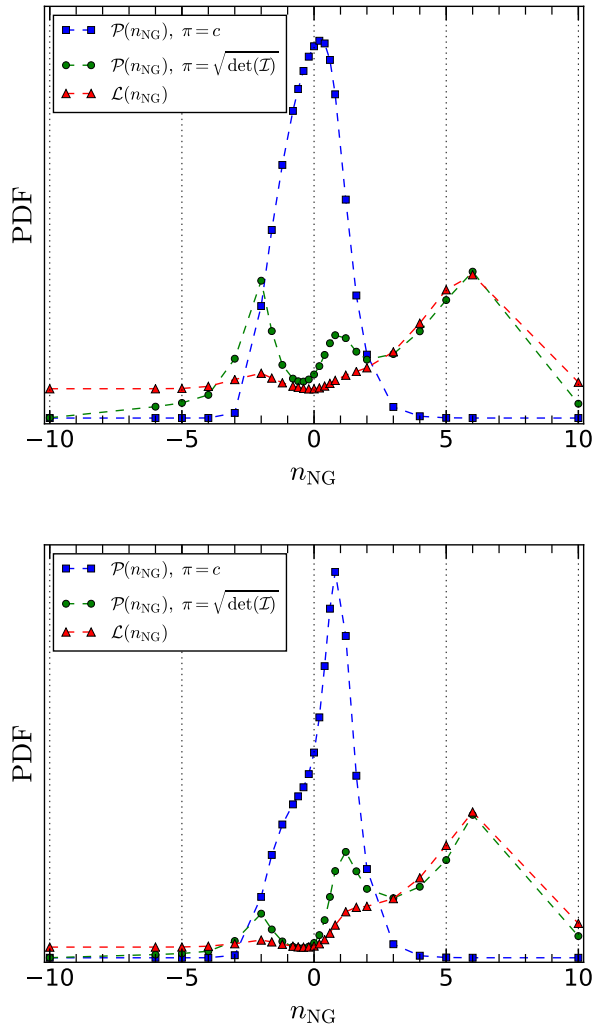


**Fig. 4.** PDF of the running parameter $n_{\rm NG}$ for the two-field local model. *Top*: SMICA map. *Bottom*: Commander map. Blue squares give the marginalized posterior assuming a constant prior, green circles are the posterior assuming a Jeffreys prior, and red triangles are the profiled Likelihood.

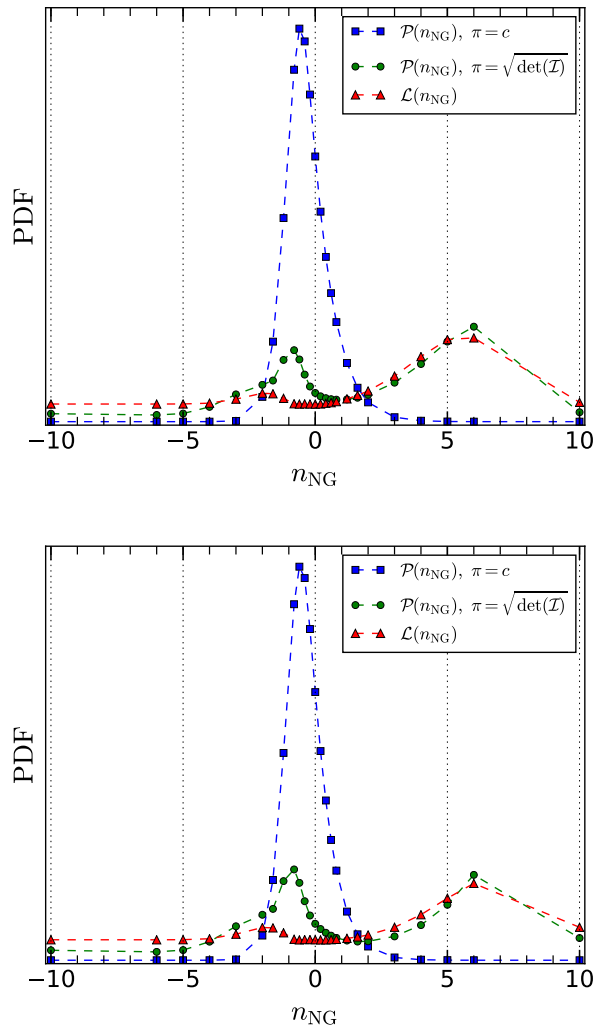


**Fig. 5.** PDF of the running parameter $n_{\rm NG}$ for the geometric mean equilateral parametrization. *Top*: SMICA map. *Bottom*: Commander map. Blue squares give the marginalized posterior assuming a constant prior, green circles are the posterior assuming a Jeffreys prior, and red triangles are the profiled Likelihood.

on $n_{\rm NG}$, while requiring a significant growth in the computational cost of the estimation pipeline, we do not include polarization in this particular analysis. We show the PDF inferred from all three of the methodologies described. Each point is derived from a KSW estimate of the amplitude for the corresponding scale-dependent template with the given running. The KSW pipeline and the map processing steps are the same as applied to the estimation of the local, equilateral, and orthogonal shapes. We include in this analysis the multipole range from 2 to 2000 and the results are corrected for the lensing bias. All curves in Figs. 3–5 are normalized to integrate to one. We consider possible values of the running in the interval $n_{\rm NG} = [-10, 10]$. This interval is two orders of magnitude wider than the theoretical expectation of the models, which are valid in the regime of mild scale dependence, having $n_{\rm NG} \approx 0.1$.

The effects of the choice of prior are very obvious: if we adopt a constant prior (blue squares) we can always identify a peak in the distribution and define proper constraints; however, this is not the case for the other priors. If we implement an uninformative prior (green circles), the shape of the distribution becomes complex, showing multiple peaks or even diverging on the boundaries, making it impossible to define constraints. A similar behaviour appears in the profiled likelihood approach (red triangles; see Eq. (36)). We used the likelihood to also perform a likelihood ratio test between its maximum value and the value for $n_{\rm NG} = 0$. Notice that in the case of zero running, these models reduce to the usual local or equilateral shapes. We assume an acceptance threshold of $\alpha = 0.01$ As expected, we do not find any evidence in favour of scale dependent models.

### 5.2.3. Resonance and axion monodromy

Now we present results from the Modal 2 and from an adapted KSW-type estimator for the broad class of resonance-type models. These models can be characterized by a bispectrum template having logarithmic oscillations with a scale, as introduced in Sect. 2.5.1. For the Modal 2 analysis we examine templates with frequency in the range $0 < \omega \leq 50$ and with a range of possible cross-sectional behaviour covering constant, flat, local, and equilateral type templates to span a broader range of possible models. The raw results have been maximised over phase and are presented in Fig. 6 for the four component-separation methods. Since the errors both grow with frequency, due to increasing suppression by the both transfer functions and projection effects,

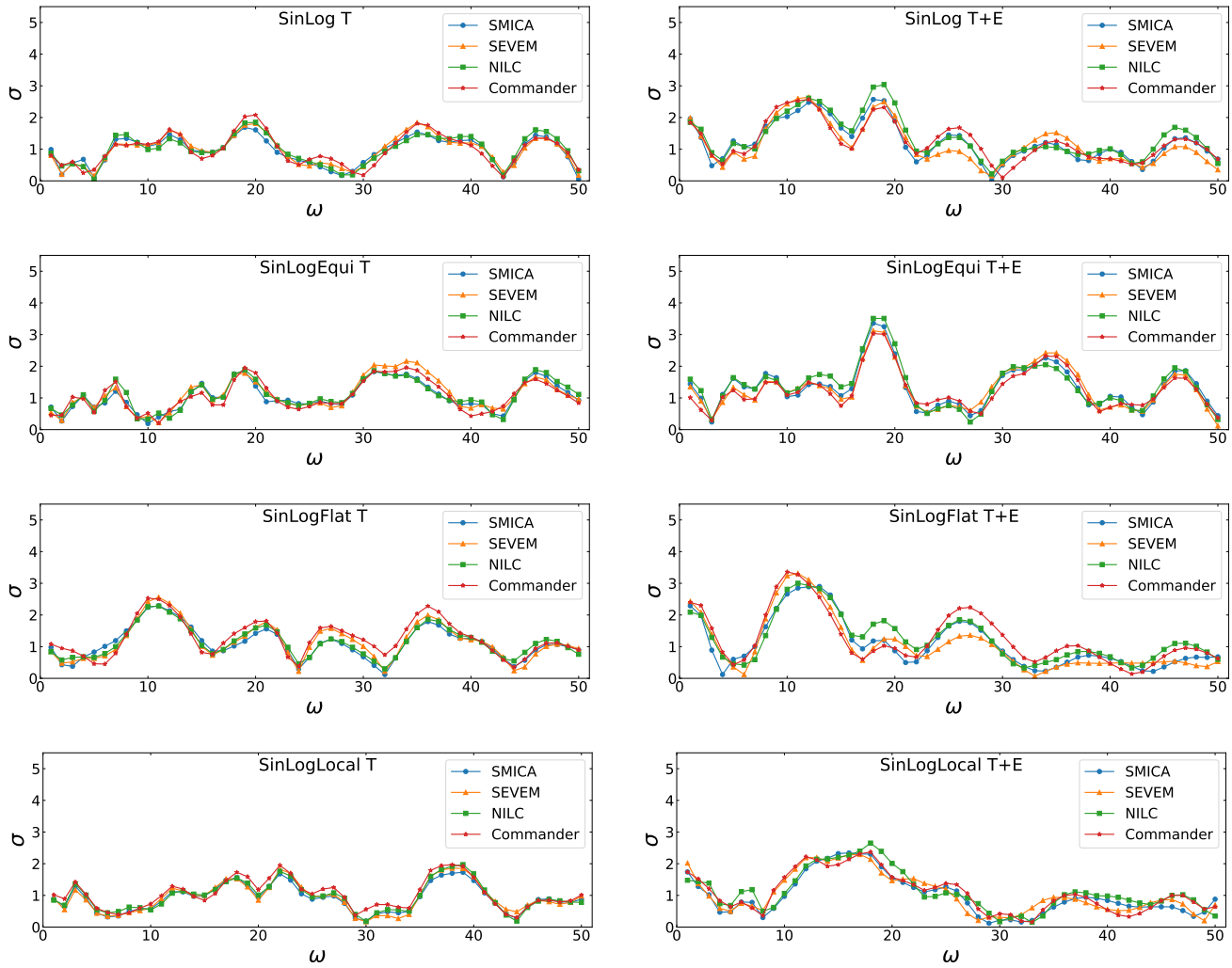


**Fig. 6.** Generalized resonance-model significance surveyed over the Modal 2 frequency range with the uppermost panels for the constant resonance model (Eq. 12), showing $T$-only (*left*) and $T$+$E$ (*right*) results, then below this the equilateral resonance model (Eq. (13)), followed by the flattened model (Eq. (14)) with the bottom panels, representing the local resonance model with a squeezed envelope. These models have been investigated using *Planck* temperature data to $\ell_{\max} = 2000$ and polarization data to $\ell_{\max} = 1500$, with the *Planck* component-separation methods `SMICA` (blue), `SEVEM` (orange), `NILC` (green), and `Commander` (red). These results are broadly consistent with those found previously (PCNG15), with some broad peaks of moderate significance emerging in both the equilateral and flattened resonance models, somewhat enhanced by including polarization data (*upper middle* and *lower middle right panels*).

and at each frequency vary by up to 15% with phase, we only plot the raw significance. Approximate 1 sigma error bars for the constant resonance model can be obtained from the formula $\sigma^T \approx 3.5\omega + 51$ and $\sigma^{T+E} \approx 1.7\omega + 32$ with errors for the equilateral and flattened cross-sections approximately 1.7 times larger and those for the local cross section 0.5 times smaller. The results are consistent across component-separation methods and are comparable with those previously presented in PCNG15, but with slight reductions in significance. Since we are surveying a large range of frequencies, we must correct our results for the "look-elsewhere" effect to asses their true significance. This is done using the optimal methods first proposed in Fergusson et al. (2015), which assess the true significance of both single peaks and also clusters of multiple peaks (which may indicate a model in this class but with a waveform that only partially correlates with the templates used). The results are presented in Table 9, with the largest look-elsewhere-corrected result being for the equilateral cross-section with around $2\sigma$ for a single peak. This is not on its own significant, but may warrant further investigation with more realistic exact templates near this frequency.

#### 5.2.4. Scale-dependent oscillatory features

In this section we show the broad range of oscillatory models sourced through features in the inflationary potential or sound speed, described in Sect. 2.5.2, which can be described by a sinusoid multiplied by either a cross-sectional template or a scaling function. The results are plotted after maximization over phase in Figs. 7 and over envelope widths in 8. Since the errors both grow with frequency, due to increasing suppression by the both transfer functions and projection effects, and at each frequency vary by up to 20% with phase, we only plot the raw significance. Approximate 1 sigma error bars for the feature model can be obtained from the formula $\sigma^T \approx 0.41\omega + 60$ and $\sigma^{T+E} \approx 0.22\omega + 33$ with errors for the equilateral and flattened cross-sections approximately 1.7 times larger. However, when we vary the envelope the error bars change by several orders of magnitude so it is not possible to provide approximate formula in this case. Look-elsewhere-adjusted results are presented in Table 10. Here all results are consistent with Gaussianity for all models in this class.

**Table 9.** Peak statistics, as defined in Fergusson et al. (2015), for the resonance models, showing the maximum "Raw" peak significance, the "Single" peak significance after accounting for the parameter survey look-elsewhere effect, and the "Multi"-peak statistic integrating across all peaks (also accounting for the look-elsewhere correction).

| | SMICA | | | SEVEM | | | NILC | | | Commander | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Raw | Single | Multi | Raw | Single | Multi | Raw | Single | Multi | Raw | Single | Multi |
| Sin(log) constant $T$ only . . . . . | 1.7 | 0.0 | 0.0 | 1.8 | 0.1 | 0.1 | 1.8 | 0.1 | 0.1 | 2.1 | 0.3 | 0.2 |
| Sin(log) constant $T$+$E$ . . . . . | 2.6 | 0.9 | 0.9 | 2.6 | 1.0 | 1.0 | 3.0 | 1.6 | 1.4 | 2.6 | 0.9 | 0.9 |
| Sin(log) equilateral $T$ only . . . . . | 1.8 | 0.1 | 0.1 | 2.2 | 0.4 | 0.3 | 1.9 | 0.1 | 0.1 | 2.0 | 0.2 | 0.2 |
| Sin(log) equilateral $T$+$E$ . . . . . | 3.4 | 2.0 | 1.5 | 3.1 | 1.7 | 1.4 | 3.5 | 2.2 | 1.8 | 3.0 | 1.6 | 1.2 |
| Sin(log) flattened $T$ only . . . . . | 2.3 | 0.5 | 0.4 | 2.6 | 0.9 | 0.7 | 2.3 | 0.5 | 0.4 | 2.5 | 0.9 | 0.7 |
| Sin(log) flattened $T$+$E$ . . . . . | 2.9 | 1.4 | 1.4 | 3.3 | 2.0 | 1.8 | 3.0 | 1.5 | 1.5 | 3.4 | 2.0 | 1.9 |
| Sin(log) local $T$ only . . . . . | 1.7 | 0.1 | 0.0 | 1.9 | 0.1 | 0.1 | 2.0 | 0.2 | 0.1 | 2.0 | 0.2 | 0.2 |
| Sin(log) local $T$+$E$ . . . . . | 2.3 | 0.6 | 0.6 | 2.3 | 0.5 | 0.6 | 2.7 | 1.0 | 0.8 | 2.4 | 0.6 | 0.6 |

**Notes.** This table does not include the results of the high-frequency resonance-model estimator, whose significance was assessed independently and presented in Sect. 5.2.5.

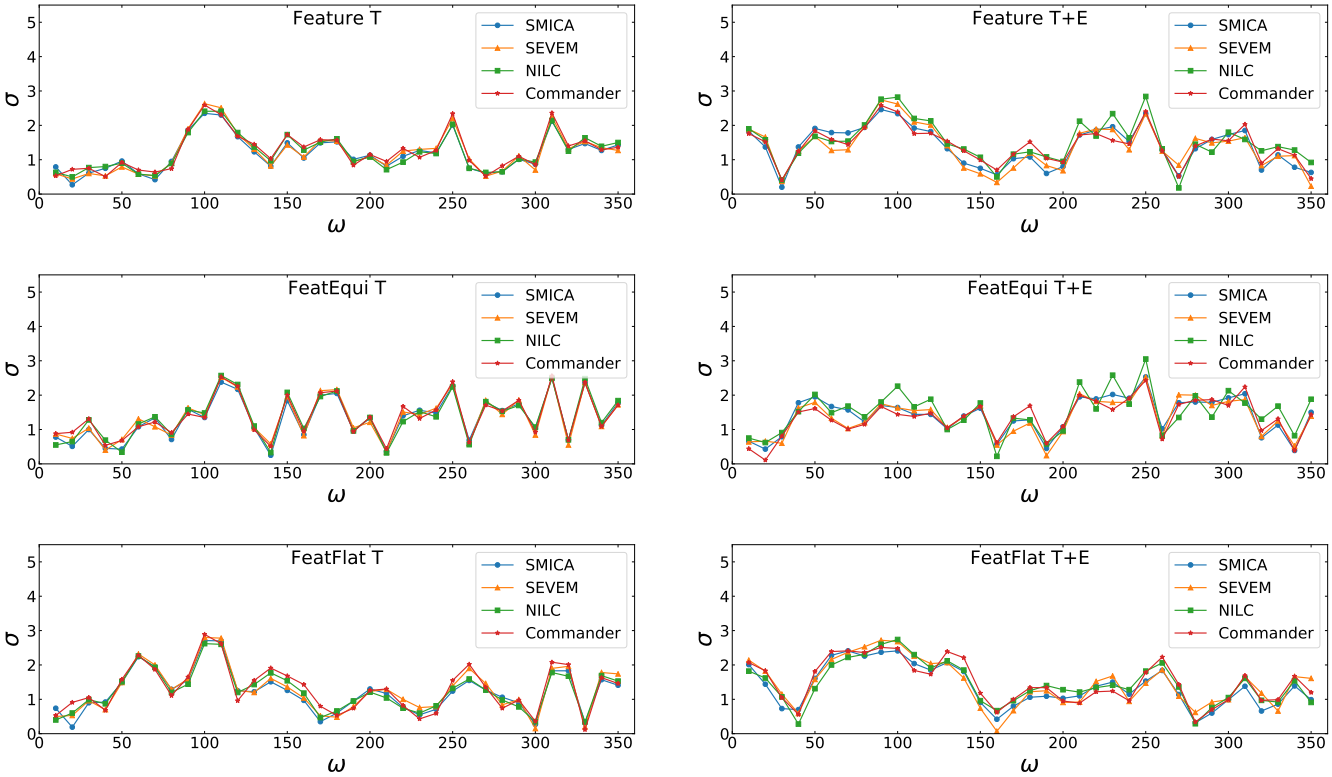


**Fig. 7.** Generalized feature model significance surveyed over the Modal 2 frequency range $0 < \omega < 350$, after marginalizing over the phase $\phi$. *Top panels*: results for the constant feature model (Eq. (15)) with $T$ only (*left*) and $T$+$E$ (*right*), *middle panels*: for the equilateral feature model (Eq. (18)), and *bottom panels*: for the flattened feature model (Eq. (19)). The same conventions apply as in Fig. 6. These feature model results generally have lower significance than obtained previously (PCNG15), with polarization data not tending to reinforce apparent peaks found using temperature data only.

### 5.2.5. High-frequency feature and resonant-model estimator

We used the high-frequency linear oscillation estimator described in Sect. 3.1.5 to probe the constant feature model with frequencies up to $\omega = 3000$. In the overlapping frequency range we also confirmed that the results are compatible with those of the Modal pipeline. The results are shown in the upper two panels of Fig. 9. Due to the computational demands of this estimator we have only analysed SMICA maps, since it was already shown in PCNG15 that all component-separations methods agree well for this estimator. No statistically significant peak is found, and the distribution of peaks is consistent with the 2015 analysis. The highest peak is $2.9\sigma$ in $T$ only and $3.5\sigma$ in $T + E$, compared to the Gaussian expectation of $3.1(\pm 0.3)\sigma$[9].

For the constant resonance model, we used the high frequency estimator described in Sect. 3.1.5 to cover the frequency range $0 < \omega < 1000$. The results are shown in Fig. 7 (lower two panels) for SMICA data. The highest peak is $3.1\sigma$ for $T$ only and $3.0\sigma$ for $T + E$, compared to the Gaussian expectation

[9] The Gaussian expectation is not zero because it is the average value of the most significant fluctuation, maximized over the frequency and phase parameters of the shape. The expectation value of a single amplitude with a fixed frequency and phase is zero for a Gaussian map.

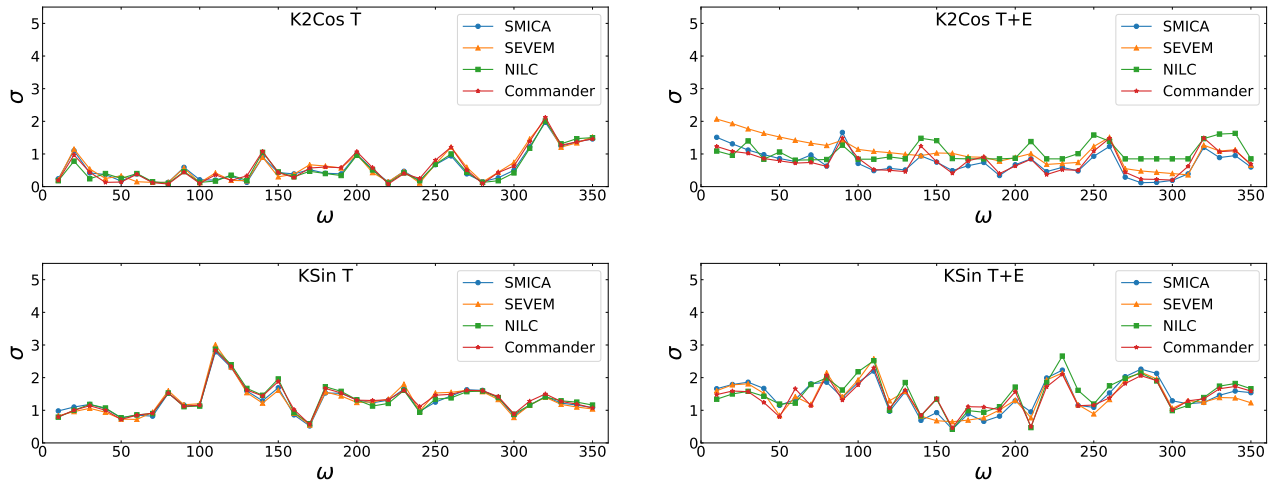


**Fig. 8.** *Top*: significance of single-field models for the potential feature case with a $K^2 \cos \omega K$ scaling dependence (Eq. (16)). *Bottom*: significance for rapidly varying sound speed with a $K \sin \omega K$ scaling (Eq. (17)). *Left panels*: for $T$ only and *right panels*: for $T + E$. These results have been marginalized over the envelope parameter $\alpha$ (determined by feature width and height) from $\alpha = 0$ to $\alpha\omega = 90$. There appears to be no evidence for these very specific signatures in this frequency range when polarization data are included.

**Table 10.** Peak statistics, as defined in Fergusson et al. (2015), for the different feature models, showing the Raw peak maximum significance (for the given Modal 2 survey domain), the corrected significance of this Single maximum peak after accounting for the parameter survey size (the look-elsewhere effect) and the Multi-peak statistic which integrates across the adjusted significance of all peaks to determine consistency with Gaussianity.

| | SMICA | | | SEVEM | | | NILC | | | Commander | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Raw | Single | Multi | Raw | Single | Multi | Raw | Single | Multi | Raw | Single | Multi |
| Features constant $T$ only | 2.4 | 0.2 | 0.2 | 2.6 | 0.5 | 0.5 | 2.4 | 0.2 | 0.3 | 2.6 | 0.4 | 0.5 |
| Features constant $T$+$E$ | 2.5 | 0.3 | 0.3 | 2.7 | 0.6 | 0.7 | 2.8 | 0.8 | 1.2 | 2.6 | 0.4 | 0.5 |
| Features equilateral $T$ only | 2.5 | 0.3 | 0.4 | 2.5 | 0.3 | 0.5 | 2.6 | 0.4 | 0.6 | 2.6 | 0.4 | 0.6 |
| Features equilateral $T$+$E$ | 2.5 | 0.3 | 0.3 | 2.5 | 0.3 | 0.3 | 3.0 | 1.1 | 1.1 | 2.4 | 0.2 | 0.2 |
| Features flattened $T$ only | 2.7 | 0.6 | 0.8 | 2.8 | 0.7 | 1.0 | 2.6 | 0.5 | 0.6 | 2.9 | 0.9 | 0.9 |
| Features flattened $T$+$E$ | 2.4 | 0.2 | 0.4 | 2.7 | 0.6 | 0.9 | 2.7 | 0.6 | 0.7 | 2.5 | 0.3 | 0.5 |
| $K^2$ cos features $T$ only | 2.0 | 0.0 | 0.0 | 2.0 | 0.0 | 0.0 | 2.0 | 0.0 | 0.0 | 2.1 | 0.0 | 0.0 |
| $K^2$ cos features $T$+$E$ | 1.7 | 0.0 | 0.0 | 2.1 | 0.0 | 0.0 | 1.6 | 0.0 | 0.0 | 1.5 | 0.0 | 0.0 |
| $K$ sin features $T$ only | 2.8 | 0.7 | 0.7 | 3.0 | 1.0 | 1.0 | 2.9 | 0.8 | 0.8 | 2.8 | 0.8 | 0.8 |
| $K$ sin features $T$+$E$ | 2.3 | 0.1 | 0.1 | 2.6 | 0.4 | 0.4 | 2.7 | 0.5 | 0.6 | 2.3 | 0.1 | 0.1 |

**Notes.** This table does not include the results of the high-frequency resonance-model estimator, whose significance was assessed independently and presented in Sect. 5.2.5.

of $3.4\sigma \pm 0.4\sigma$. The Gaussian expectation was obtained from the covariance matrix of the estimators using the method of Meerburg et al. (2016b). In summary we do not find evidence for non-Gaussianity in the high-frequency feature and resonance-model analysis.

We used the results obtained here to perform a joint power and bispectrum analysis, which was presented in the accompanying paper Planck Collaboration X (2020) in a search for correlated features. No evidence was found for such correlated features in either the power spectrum or the bispectrum.

### 5.2.6. Equilateral-type models and the effective field theory of inflation

Many physically well-motivated inflationary models produce non-Gaussianity of the equilateral type. While the equilateral template accurately approximates models in this class, it is nevertheless interesting to constrain the exact templates for two reasons. Firstly, while the equilateral template correlates well with the true models, they are only normalized in the equilateral limit, which neglects the rest of configuration space and so is only approximate. This is demonstrated by the fact that the uncertainties for the true templates differ by up to a factor of 2 compared with the equilateral limit. Secondly, even small deviations in correlation can result in much larger shifts in measurements, for example models that are 95% correlated should produce measurements that are on average $1/3\sigma$ apart. Here we constrain the exact templates for all inflationary bispectra that fit into the broad equilateral class; for details on these exact templates we refer the reader to the paper PCNG15. The results for all models in this class are presented in Table 11 and are all below $1\sigma$, as expected due to their correlation with the equilateral template, but with significant spread due to the small differences in correlation between templates.

### 5.2.7. Models with excited initial states (non-Bunch-Davies vacua)

Inflationary models that modify the initial vacuum for the inflaton generally produce shapes in the flattened class, meaning that

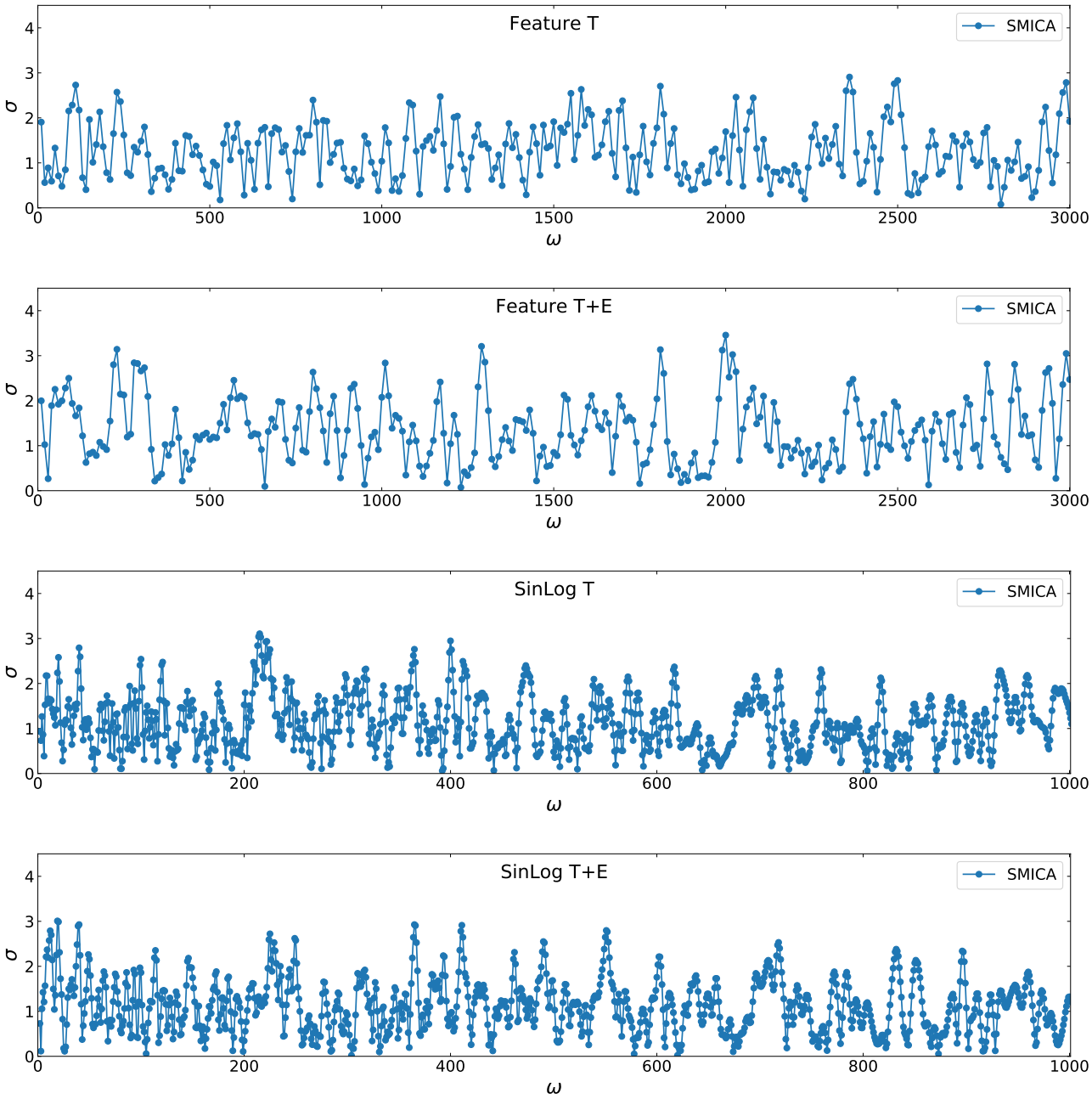


**Fig. 9.** High-frequency estimator results for feature and resonance models. *Upper two panels*: significance for the constant feature model (Eq. (15)) surveyed over the frequency range $0 < \omega < 3000$ after marginalizing over the phase $\phi$, for $T$ and $T+E$ SMICA maps. *Bottom two panels*: significance for the constant resonance model (Eq. (12)) surveyed over the frequency range $0 < \omega < 1000$ after marginalizing over the phase $\phi$, for $T$ and $T+E$ SMICA maps. As for the Modal expansion case, the high-frequency results generally have lower significance than obtained previously (PCNG15), with polarization data not tending to reinforce apparent peaks found using temperature data only.

they peak in triangle configurations with zero area. There is a wide variety of models of this type proposed in the literature, which are described in detail in Sect. 2.6. For simplicity we also include here the results for the warm inflation template, described in Sect. 2.2, since it exhibits similar behaviour once projected, despite the mechanism behind it being quite different. The results for this class are presented in Table 12. The most significant measurement is for the NBD sin template, which produces results around $2\sigma$. However, it is important to note that this result involves a marginalization over a frequency type parameter so there is a look-elsewhere effect that has not yet been taken into account. Because of this we expect the true significance will be lower, so we can safely claim that our results are consistent with Gaussianity, while noting that it may be interesting to revisit oscillatory NBD templates, like the NBD sin model, using future data sets.

#### 5.2.8. Direction-dependent primordial non-Gaussianity

Here we present results for inflationary models where gauge fields induce a direction dependance to a local bispectrum, as described in Sect. 2.7. Results for the $L = 1$ and $L = 2$ templates are presented in Table 13 from the Modal 2 pipeline. Due to the complicated behaviour in the squeezed limit, the convergence for

**Table 11.** Constraints on models with equilateral-type NG covering the shapes predicted by the effective field theory of inflation, together with constraints on specific non-canonical inflation models, such as DBI inflation.

| | SMICA | | SEVEM | | NILC | | Commander | |
|---|---|---|---|---|---|---|---|---|
| Equilateral-type model | $A \pm \sigma_A$ | $S/N$ | $A \pm \sigma_A$ | $S/N$ | $A \pm \sigma_A$ | $S/N$ | $A \pm \sigma_A$ | $S/N$ |
| Constant $T$ only | $31 \pm 38$ | 0.8 | $24 \pm 38$ | 0.6 | $15 \pm 38$ | 0.4 | $15 \pm 38$ | 0.4 |
| Constant $T$+$E$ | $21 \pm 24$ | 0.9 | $18 \pm 25$ | 0.7 | $7 \pm 24$ | 0.3 | $14 \pm 25$ | 0.6 |
| Equilateral $T$ only | $34 \pm 67$ | 0.5 | $34 \pm 68$ | 0.5 | $20 \pm 67$ | 0.3 | $36 \pm 68$ | 0.5 |
| Equilateral $T$+$E$ | $-5 \pm 43$ | −0.1 | $2 \pm 45$ | 0.1 | $-16 \pm 43$ | −0.4 | $4 \pm 44$ | 0.1 |
| EFT shape 1 $T$ only | $43 \pm 61$ | 0.7 | $41 \pm 63$ | 0.6 | $26 \pm 61$ | 0.4 | $39 \pm 62$ | 0.6 |
| EFT shape 1 $T$+$E$ | $8 \pm 40$ | 0.2 | $10 \pm 41$ | 0.2 | $-7 \pm 40$ | −0.2 | $10 \pm 41$ | 0.3 |
| EFT shape 2 $T$ only | $51 \pm 47$ | 1.1 | $45 \pm 48$ | 1.0 | $33 \pm 46$ | 0.7 | $37 \pm 47$ | 0.8 |
| EFT shape 2 $T$+$E$ | $28 \pm 30$ | 0.9 | $23 \pm 31$ | 0.8 | $11 \pm 30$ | 0.4 | $21 \pm 31$ | 0.7 |
| DBI inflation $T$ only | $46 \pm 58$ | 0.8 | $43 \pm 60$ | 0.7 | $29 \pm 58$ | 0.5 | $39 \pm 59$ | 0.7 |
| DBI inflation $T$+$E$ | $14 \pm 38$ | 0.4 | $14 \pm 39$ | 0.4 | $-2 \pm 38$ | −0.1 | $14 \pm 39$ | 0.4 |
| Ghost inflation $T$ only | $6 \pm 81$ | 0.1 | $13 \pm 83$ | 0.2 | $-0 \pm 81$ | −0.0 | $25 \pm 82$ | 0.3 |
| Ghost inflation $T$+$E$ | $-48 \pm 52$ | −0.9 | $-30 \pm 54$ | −0.6 | $-50 \pm 52$ | −1.0 | $-25 \pm 53$ | −0.5 |
| Inverse decay $T$ only | $38 \pm 41$ | 0.9 | $32 \pm 42$ | 0.8 | $21 \pm 41$ | 0.5 | $23 \pm 42$ | 0.5 |
| Inverse decay $T$+$E$ | $24 \pm 27$ | 0.9 | $19 \pm 27$ | 0.7 | $8 \pm 26$ | 0.3 | $17 \pm 27$ | 0.6 |

**Notes.** See Sect. 2 of PCNG13 for detailed explanation of these specific models, with further implications discussed in Sect. 8.

**Table 12.** Constraints on models with excited initial states (non-Bunch-Davies models), as well as warm inflation.

| | SMICA | | SEVEM | | NILC | | Commander | |
|---|---|---|---|---|---|---|---|---|
| Flattened-type model | $A \pm \sigma_A$ | $S/N$ | $A \pm \sigma_A$ | $S/N$ | $A \pm \sigma_A$ | $S/N$ | $A \pm \sigma_A$ | $S/N$ |
| Flat model $T$ only | $56 \pm 72$ | 0.8 | $37 \pm 72$ | 0.5 | $23 \pm 71$ | 0.3 | $7 \pm 72$ | 0.1 |
| Flat model $T$+$E$ | $60 \pm 43$ | 1.4 | $44 \pm 44$ | 1.0 | $29 \pm 43$ | 0.7 | $33 \pm 44$ | 0.8 |
| Non-Bunch-Davies $T$ only | $13 \pm 89$ | 0.1 | $6 \pm 90$ | 0.1 | $-25 \pm 88$ | −0.3 | $-17 \pm 90$ | −0.2 |
| Non-Bunch-Davies $T$+$E$ | $60 \pm 54$ | 1.1 | $38 \pm 56$ | 0.7 | $32 \pm 53$ | 0.6 | $44 \pm 55$ | 0.8 |
| NBD sin $T$ only | $-630 \pm 445$ | −1.4 | $-673 \pm 450$ | −1.5 | $-629 \pm 444$ | −1.4 | $-747 \pm 445$ | −1.7 |
| NBD sin $T$+$E$ | $-496 \pm 247$ | −2.0 | $-483 \pm 255$ | −1.9 | $-547 \pm 249$ | −2.2 | $-505 \pm 253$ | −2.0 |
| NBD1 cos flattened $T$ only | $6 \pm 21$ | 0.3 | $8 \pm 21$ | 0.4 | $3 \pm 21$ | 0.1 | $5 \pm 22$ | 0.2 |
| NBD1 cos flattened $T$+$E$ | $2 \pm 15$ | 0.1 | $6 \pm 16$ | 0.4 | $-7 \pm 15$ | −0.5 | $1 \pm 16$ | 0.0 |
| NBD2 cos squeezed $T$ only | $48 \pm 167$ | 0.3 | $41 \pm 170$ | 0.2 | $-14 \pm 166$ | −0.1 | $11 \pm 170$ | 0.1 |
| NBD2 cos squeezed $T$+$E$ | $-5.0 \pm 6.0$ | −0.8 | $-6.5 \pm 6.2$ | −1.1 | $-6.7 \pm 6.0$ | −1.1 | $-5.6 \pm 6.0$ | −0.9 |
| NBD1 sin flattened $T$ only | $-22 \pm 26$ | −0.8 | $-33 \pm 26$ | −1.3 | $-27 \pm 27$ | −1.0 | $-38 \pm 27$ | −1.4 |
| NBD1 sin flattened $T$+$E$ | $-14 \pm 20$ | −0.7 | $-22 \pm 20$ | −1.1 | $-19 \pm 20$ | −1.0 | $-24 \pm 20$ | −1.2 |
| NBD2 sin squeezed $T$ only | $-0.5 \pm 0.6$ | −0.7 | $-0.3 \pm 0.5$ | −0.6 | $-2.9 \pm 2.6$ | −1.1 | $-0.8 \pm 0.7$ | −1.1 |
| NBD2 sin squeezed $T$+$E$ | $-0.3 \pm 0.4$ | −0.7 | $-0.2 \pm 0.4$ | −0.5 | $-0.4 \pm 0.4$ | −0.9 | $-0.4 \pm 0.4$ | −1.0 |
| NBD3 non-canonical (×1000) $T$ only | $-4.4 \pm 7.8$ | −0.6 | $-5.6 \pm 7.9$ | −0.7 | $-4.6 \pm 8.0$ | −0.6 | $-5.5 \pm 7.9$ | −0.7 |
| NBD3 non-canonical (×1000) $T$+$E$ | $-6.4 \pm 5.8$ | −1.1 | $-7.5 \pm 5.9$ | −1.3 | $-5.8 \pm 5.9$ | −1.0 | $-6.2 \pm 5.9$ | −1.0 |
| WarmS inflation $T$ only | $-39 \pm 44$ | −0.9 | $-35 \pm 44$ | −0.8 | $-34 \pm 44$ | −0.8 | $-18 \pm 44$ | −0.4 |
| WarmS inflation $T$+$E$ | $-48 \pm 27$ | −1.8 | $-37 \pm 28$ | −1.3 | $-41 \pm 28$ | −1.5 | $-27 \pm 28$ | −1.0 |

**Notes.** See Sect. 2 for further explanation and the labelling of these classes of NBD models. NBD, NBD1, and NBD2 models contain free parameters, so here we quote the maximum significance found over the entire parameter range. The location of the maximum for $T$ and $T$+$E$ can occur for different parameter values for the model, and so the results with and without polarization cannot be directly compared; however, the model parameters are held fixed across different component-separation methods.

these models is lower than for some others, but the Modal correlation remains above 93% so a perfect estimator may see shifts of order $0.5\sigma$. The largest result is $1.9\sigma$ for $L = 1$ and SMICA, but this is not seen consistently across all component-separation methods so cannot be considered robust. We can then conclude that our results are consistent with Gaussianity.

#### 5.2.9. Parity-violating tensor non-Gaussianity motivated by pseudo-scalars

In this section, we report constraints on the tensor non-linearity parameter $f_{\rm NL}^{\rm tens}$ (Eq. (22)) obtained from temperature and $E$-mode polarization maps. As in our 2015 analysis, we examine even and odd $\ell_1+\ell_2+\ell_3$ multipole domains, employing the original parity-even Modal estimator, $\hat{f}_{\rm NL}^{\rm even}$ (Fergusson et al. 2010, 2012; Shiraishi et al. 2019), and its parity-odd version, $\hat{f}_{\rm NL}^{\rm odd}$ (Shiraishi et al. 2014, 2015, 2019), respectively. The constraints obtained from both domains are also combined by computing

$$\hat{f}_{\rm NL}^{\rm all} = \frac{F^{\rm even}\hat{f}_{\rm NL}^{\rm even} + F^{\rm odd}\hat{f}_{\rm NL}^{\rm odd}}{F^{\rm even} + F^{\rm odd}}, \tag{41}$$

where $F^{\rm even/odd}$ is the Fisher matrix from $\ell_1+\ell_2+\ell_3 = $ even/odd. This analysis is performed for under the multipole ranges $2 \leq \ell \leq 500$ in temperature and $4 \leq \ell \leq 500$ in polarization. Here, the use of the first 40 multipoles of the $E$-mode polarization data, which were disregarded in the 2015 analysis, boosts contributions of $TTE$, $TEE$, and $EEE$ to constraining $f_{\rm NL}^{\rm tens}$. Although the data, simulations (used for the computation of the linear terms and error bars) and analysis details (e.g. masks, beams, and noise distributions) are somewhat different, other settings for the

**Table 13.** Direction-dependent NG results for both the $L = 1$ and $L = 2$ models from the Modal 2 pipeline.

| | SMICA | | SEVEM | | NILC | | Commander | |
|---|---|---|---|---|---|---|---|---|
| Equilateral-type model | $A \pm \sigma_A$ | $S/N$ | $A \pm \sigma_A$ | $S/N$ | $A \pm \sigma_A$ | $S/N$ | $A \pm \sigma_A$ | $S/N$ |
| $L = 1$ | | | | | | | | |
| Modal 2 $T$ only . . . . . | $-51 \pm 51$ | $-1.0$ | $-40 \pm 51$ | $-0.8$ | $-41 \pm 52$ | $-0.8$ | $-18 \pm 52$ | $-0.3$ |
| Modal 2 $T$+$E$ . . . . . | $-57 \pm 30$ | $-1.9$ | $-42 \pm 31$ | $-1.4$ | $-42 \pm 30$ | $-1.4$ | $-31 \pm 31$ | $-1.0$ |
| $L = 2$ | | | | | | | | |
| Modal 2 $T$ only . . . . . | $1.2 \pm 3.0$ | 0.4 | $1.7 \pm 3.1$ | 0.5 | $1.3 \pm 3.1$ | 0.4 | $0.5 \pm 3.1$ | 0.1 |
| Modal 2 $T$+$E$ . . . . . | $0.9 \pm 2.4$ | 0.4 | $1.8 \pm 2.4$ | 0.7 | $-0.2 \pm 2.4$ | $-0.1$ | $-0.2 \pm 2.4$ | $-0.1$ |

$f_{\rm NL}^{\rm tens}$ estimation are basically the same as in the 2015 analysis (PCNG15).

The results from four different component-separated maps are summarized in Table 14. We confirm there that the sizes of errors in the parity-odd $T$+$E$ analysis reduce to almost the same level as the parity-even counterparts. This is because the low-$\ell$ signal of $TTE$, which dominates the signal-to-noise ratio from $\ell_1 + \ell_2 + \ell_3 =$ odd (Shiraishi et al. 2013b), is now taken into account. Although the errors for the parity-even case are as large as the 2015 ones, owing to the sensitivity improvement of the parity-odd part, the whole-domain constraints become more stringent.

Regardless of some updates, we find no $>2\sigma$ signal, which is consistent with the conclusion of the 2015 analysis. This indicates no parity violation in the primordial Universe and accordingly gives constraints on some axion inflationary models (see Sect. 8.3).

**Table 14.** Results for the tensor non-linearity parameter $f_{\rm NL}^{\rm tens}/10^2$ obtained from the SMICA, SEVEM, NILC, and Commander temperature and polarization maps.

| | Even | Odd | All |
|---|---|---|---|
| SMICA | | | |
| $T$ . . . . . | $4 \pm 17$ | $100 \pm 100$ | $6 \pm 16$ |
| $E$ . . . . . | $33 \pm 67$ | $-570 \pm 720$ | $29 \pm 67$ |
| $T$+$E$ . . . . . | $11 \pm 14$ | $1 \pm 18$ | $8 \pm 11$ |
| SEVEM | | | |
| $T$ . . . . . | $4 \pm 17$ | $90 \pm 100$ | $6 \pm 16$ |
| $E$ . . . . . | $75 \pm 75$ | $-790 \pm 830$ | $70 \pm 75$ |
| $T$+$E$ . . . . . | $16 \pm 14$ | $2 \pm 20$ | $13 \pm 12$ |
| NILC | | | |
| $T$ . . . . . | $4 \pm 17$ | $90 \pm 100$ | $6 \pm 16$ |
| $E$ . . . . . | $-16 \pm 81$ | $-540 \pm 820$ | $-19 \pm 80$ |
| $T$+$E$ . . . . . | $6 \pm 14$ | $3 \pm 21$ | $5 \pm 11$ |
| Commander | | | |
| $T$ . . . . . | $5 \pm 17$ | $90 \pm 100$ | $6 \pm 16$ |
| $E$ . . . . . | $21 \pm 69$ | $-1200 \pm 700$ | $13 \pm 69$ |
| $T$+$E$ . . . . . | $10 \pm 14$ | $-2 \pm 19$ | $7 \pm 11$ |

**Notes.** The central values and the errors (68 % CL) extracted from $\ell_1 + \ell_2 + \ell_3 =$ even ("Even"), $\ell_1 + \ell_2 + \ell_3 =$ odd ("Odd"), and their whole domain ("All") are separately described. One can see that all $T$-only results are in good agreement with both the *Planck* 2015 ones (PCNG15) and WMAP ones (Shiraishi et al. 2015).

### 5.3. Bispectrum reconstruction

#### 5.3.1. Modal bispectrum reconstruction

The Modal bispectrum estimator filters the *Planck* foreground-removed CMB maps, for example, SMICA, SEVEM, NILC, and Commander, using $n_{\rm max} = 2001$ polynomial modes to obtain model coefficients $\beta_n$. This procedure is undertaken to obtain all auto- and cross-correlations between the temperature and polarization components *TTT, TTE, TEE*, and *EEE*. We can then use the $\beta_n$ coefficients with the polynomial modes to obtain a full 3D reconstruction of the *Planck* temperature and polarization bispectra and this is shown in Fig. 10. These bispectra are in close agreement with those published in the paper analysing the *Planck* 2015 results (PCNG15) when comparison is made in the signal-dominated regime.

#### 5.3.2. Binned bispectrum reconstruction

The Binned bispectrum estimator can also be used to study the bispectrum itself, in addition to determining the $f_{\rm NL}$ amplitudes of specific templates, as in the previous sections. In particular we can investigate if any non-Gaussianity beyond that of the explicit models tested can be found in the cleaned CMB maps. Since we want to detect specific bin-triplets standing out from the noise, we work with the linear-term-corrected signal-to-noise-ratio bispectrum. Except for the high-$\ell$ bins, where a point-source signal is present, the bin-triplets are noise dominated. Instead of rebinning, we can use a Gaussian kernel to smooth the bispectrum, so that structure localized in harmonic space stands out from the noise. In this process, we mask out a few bin-triplets that have non-Gaussian noise (due to the fact that they contain very few valid $\ell$-triplets). We also have to take into account edge effects from the non-trivial domain of definition of the bispectrum. The method has been described in PCNG15 and more extensively in Bucher et al. (2016).

Slices of the smoothed binned signal-to-noise bispectrum $\mathcal{B}_{i_1 i_2 i_3}$, with a Gaussian smoothing of $\sigma_{\rm bin} = 2$, are shown in Figs. 11 and 12. These are slices for the 20th and 40th $\ell_3$-bin as a function of $\ell_1$ and $\ell_2$. For the cross-bispectra mixing $T$ and $E$ modes, we defined $B^{T2E}_{i_1 i_2 i_3} \equiv B^{TTE}_{i_1 i_2 i_3} + B^{TET}_{i_1 i_2 i_3} + B^{ETT}_{i_1 i_2 i_3}$; and $B^{TE2}_{i_1 i_2 i_3} \equiv B^{TEE}_{i_1 i_2 i_3} + B^{ETE}_{i_1 i_2 i_3} + B^{EET}_{i_1 i_2 i_3}$, with corresponding variances $\mathrm{Var}(B^{T2E}) = \mathrm{Var}(TTE) + \mathrm{Var}(TET) + \mathrm{Var}(ETT) + 2\,\mathrm{Cov}(TTE, TET) + 2\,\mathrm{Cov}(TTE, ETT) + 2\,\mathrm{Cov}(TET, ETT)$, and similarly for $\mathrm{Var}(B^{TE2})$, where we have omitted the bin indices for clarity. The red and blue regions correspond to a significant NG, whereas grey areas in Figs. 11 and 12 show regions where the bispectrum is not defined. Results are shown for the four component-separation methods SMICA, SEVEM, NILC, and Commander, and for *TTT, T2E, TE2*, and *EEE*. We also show the *TTT* bispectra after we remove the best joint-fit unclustered and clustered point-source contribution (see Table 4). In the top row of Fig. 12, we can clearly see this significant point-source signal at high $\ell$. The *T2E*

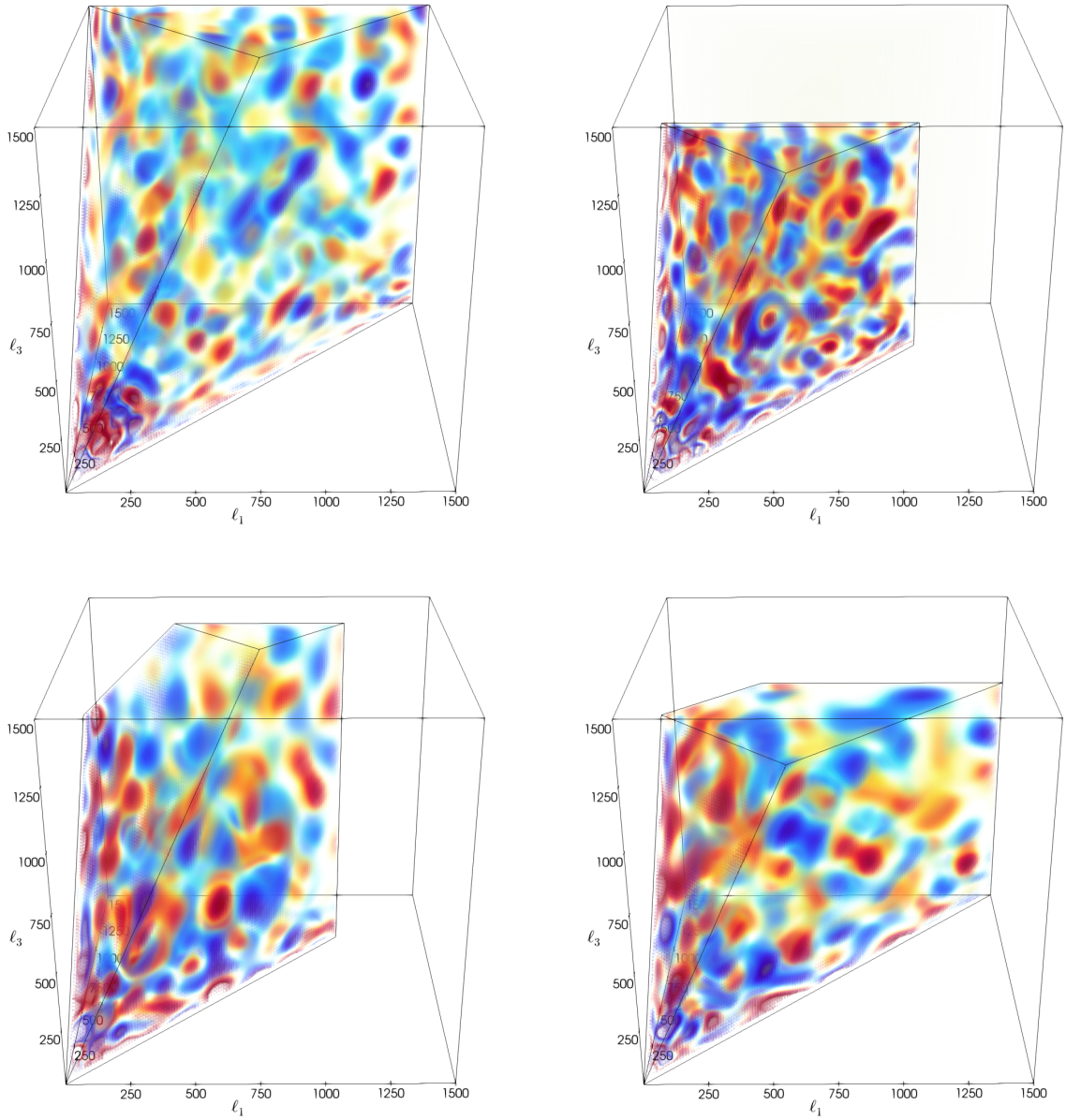


**Fig. 10.** Signal-to-noise-weighted temperature and polarization bispectra obtained from *Planck* SMICA maps using the Modal reconstruction with $n_{\max} = 2001$ polynomial modes. The bispecra are (clockwise from top left) the temperature *TTT* for $\ell \leq 1500$, the *E*-mode polarization *EEE* for $\ell \leq 1100$, the mixed temperature and polarization *TEE* (with *T* multipoles in the *z*-direction), and lastly *TTE* (with *E* multipoles in the *z*-direction). The S/N thresholds are the same between all plots.

and *TE2* bispectra do not have any obvious signals standing out, but we see some stronger NG in the *EEE* combination. Removing the best joint-fit contribution from all shapes (Tables 4 and 5) does not reduce this region (a case we do not show). The main difference with the previous release is the better qualitative agreement between the four component-separation methods in polarization, as was already the case for temperature. Commander and SEVEM now show similar structures as the NILC and SMICA bispectra, which have remained quite stable.

In order to quantify the possible residual non-Gaussianity in these smoothed bispectra, we focus on the minimum and maximum of our bin-triplets. In the case of statistically independent Gaussian numbers, we can calculate the probability distribution of the extreme value statistics. However, once we introduce correlations due to the smoothing with non-trivial boundary conditions, we do not have an analytical formula. We can instead generate Monte Carlo simulations of Gaussian random numbers with the same boundary conditions, apply the smoothing as for the data signal-to-noise bispectrum, and compute the *p*-values of the observed extremum statistics as the fraction of simulations having a more extreme extremum than our data. This requires many simulations to study the very unlikely events. In the current analysis, $10^6$ Monte Carlo simulations turn out to be sufficient and so we do not need to use the semi-analytical Ansatz introduced in Bucher et al. (2016).

In Table 15, we report for the smoothing lengths $\sigma_{\rm bin} = 1$, 2, and 3, the two-tailed *p*-value[10] of the maximum and of the minimum, defined as $p = 2\min[\mathrm{Prob}(X_{\rm MC} \leq X_{\rm data}), \mathrm{Prob}(X_{\rm MC} \geq$

[10] While a one-tailed *p*-value quantifies how likely it is for the maxima (a similar description holds for the minima) of the Monte Carlo simulations to be higher than the maximum of the data, the two-tailed *p*-value also considers how likely it is for it to be lower than that of the data. This means that a maximum that is too low yields a low *p*-value too.

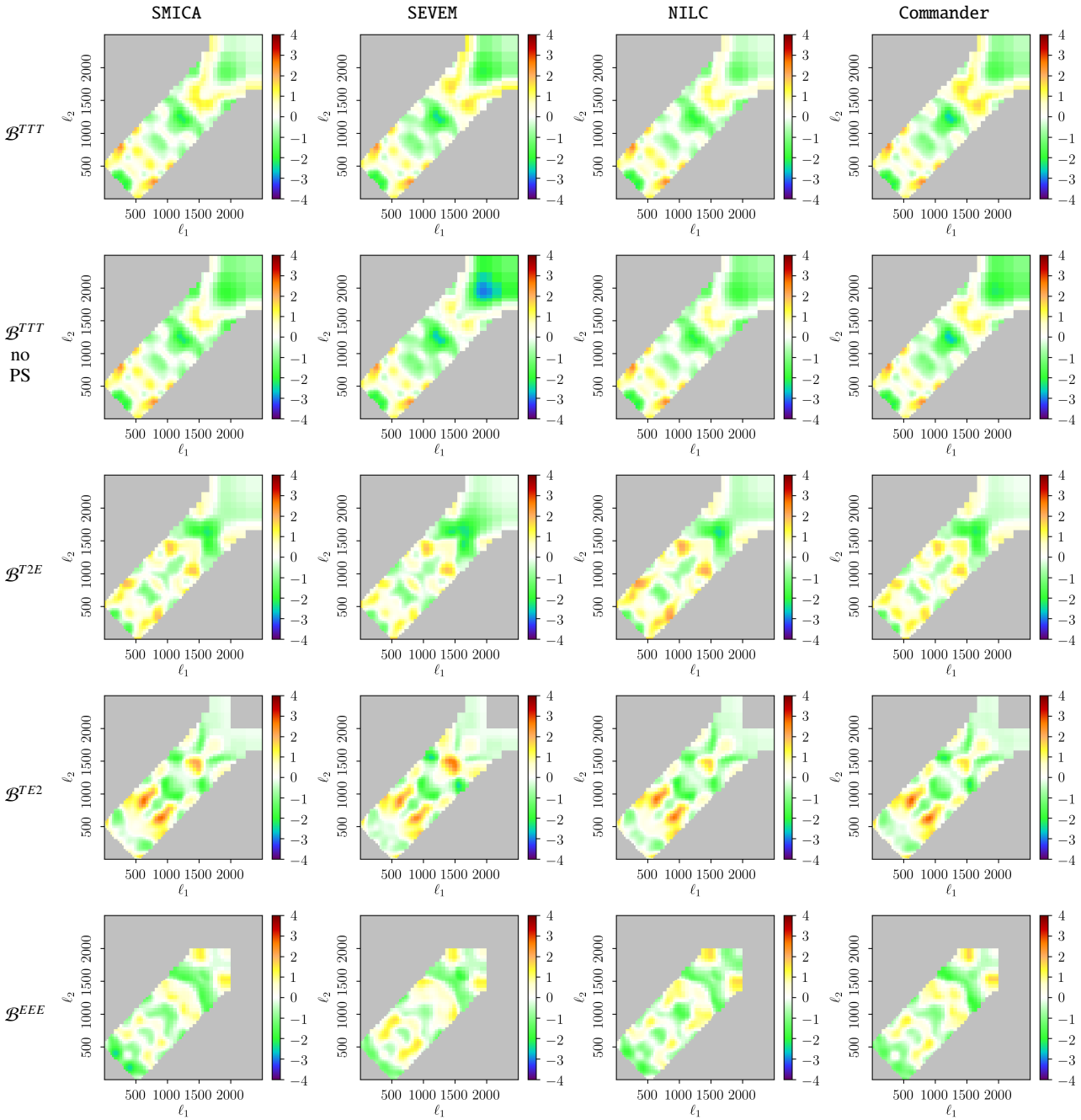


**Fig. 11.** Smoothed binned signal-to-noise bispectra $\mathcal{B}$ for the *Planck* 2018 cleaned sky maps. We show slices as a function of $\ell_1$ and $\ell_2$ for a fixed $\ell_3$-bin [518, 548]. *From left to right* results are shown for the four component-separation methods SMICA, SEVEM, NILC, and Commander. *From top to bottom* we show: *TTT; TTT* cleaned of clustered and unclustered point sources; *T2E; TE2*; and *EEE*. The colour range shows signal-to-noise from −4 to +4. The light grey regions are where the bispectrum is not defined, either because it lies outside the triangle inequality or because of the cut $\ell^E_{\rm max} = 2000$.

$X_{\rm data}$)], where $X$ is either the minimum or the maximum of the distribution. As expected, we detect a highly significant departure from Gaussian statistics in the maxima for *TTT* when we do not correct for the contribution from point sources, and the signal stands out more when increasing the smoothing kernel size. When looking at the SEVEM data for $\sigma_{\rm bin} = 2$ and 3, we find no simulation with a higher maximum, but for all the other cases our analysis should be robust. Most bispectra seem to be compatible with a simple Gaussian distribution, except for the *EEE* bispectrum in the region shown in Fig. 12, for multipole triplets around [900, 1300, 1800]. We also see some significantly high maxima for the *TTT* bispectrum of SEVEM and Commander (even after correcting for point sources), located around [800, 1100, 2000] (shown in Fig. 13). The origin of these signals clustered in multipole space is not understood.

## 6. Validation of *Planck* results

Two important potential sources of systematic effects in $f_{\rm NL}$ estimation are foreground residuals (which can also be related to the choice of mask) and an imperfect modelling of instrumental noise, either of which can lead to miscalibration of the linear correction term in the estimator. In this section, we perform a battery of tests aimed at testing the impact of these

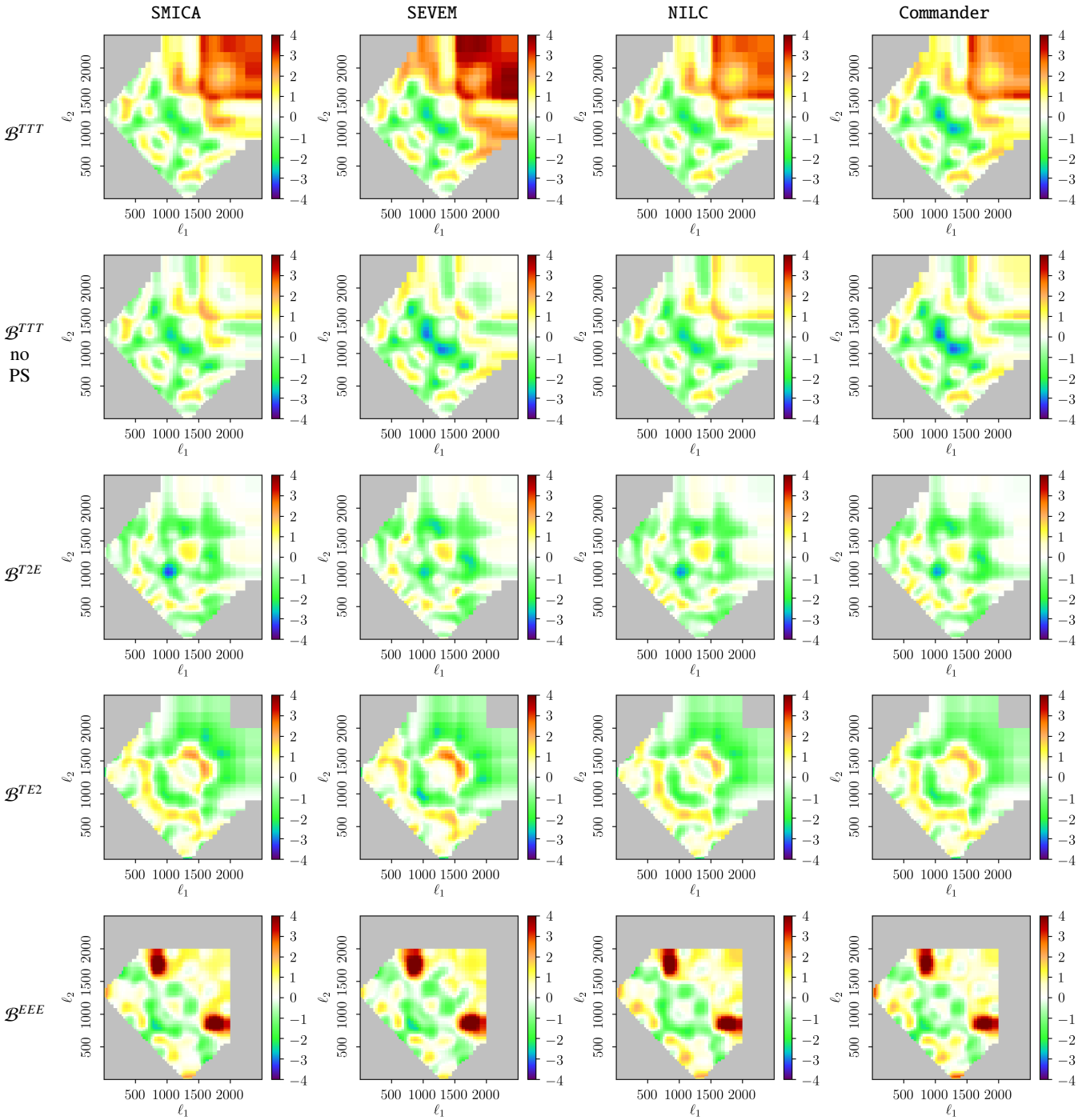


**Fig. 12.** Similar to Fig. 11, but for the $\ell_3$-bin [1291, 1345].

systematics. We typically chose only one of the KSW, Binned, or Modal pipelines for each of the tests described below. This is possible because of their excellent and well-verified agreement on both data and simulations.

### 6.1. Dependence on foreground-cleaning method

#### 6.1.1. Comparison between $f_{\rm NL}$ measurements

For this test, we consider 160 FFP10-based simulations with realistic beams and noise levels. A starting set of single-frequency Gaussian maps is processed through the different component-separation pipelines. Each pipeline combines the various frequency maps by adopting exactly the same coadding and filtering approach as done for the data. At the end, we have four sets of 160 frequency coadded realizations (the SMICA set, the SEVEM set, the NILC set and the Commander set). For each realization, in any of the four sets, we measure $f_{\rm NL}$ for the local, equilateral, and orthogonal shapes, using the Modal 1 pipeline. Then, for each shape and pair of methods, we build the difference between the two results and call it $\Delta f_{\rm NL}$. To give an example, let us consider SMICA, SEVEM, and the local shape. For each of the 160 realizations, we measure $f_{\rm NL}^{\rm local}$ from the SMICA map ($\equiv f_{\rm NL}^{\rm SMICA}$), repeat the operation with the SEVEM map ($\equiv f_{\rm NL}^{\rm SEVEM}$), and build the difference $\Delta f_{\rm NL} \equiv f_{\rm NL}^{\rm SMICA} - f_{\rm NL}^{\rm SEVEM}$. By repeating

**Table 15.** Two-tailed $p$-values of the maxima and the minima of the smoothed bispectra.

| | $\mathcal{B}^{\rm TTT}$ | $\mathcal{B}^{\rm TTT}$ (no PS) | $\mathcal{B}^{\rm T2E}$ | $\mathcal{B}^{\rm TE2}$ | $\mathcal{B}^{\rm EEE}$ |
|---|---|---|---|---|---|
| | | | Maximum $p$-values | | |
| SMICA . . . . . | (0.08, 0.04, 1.8 $\times 10^{-3}$) | (0.84, 0.80, 0.95) | (0.21, 0.11, 0.72) | (0.82, 0.36, 0.20) | (72, 0.34, 0.22) $\times 10^{-2}$ |
| SEVEM . . . . . | (1.2 $\times 10^{-4}$ , $< 10^{-6}$, $< 10^{-6}$) | (0.27, 0.03, 0.03) | (0.32, 0.18, 0.82) | (0.37, 0.50, 0.89) | (44, 0.32, 0.14) $\times 10^{-2}$ |
| NILC . . . . . | (0.18, 0.03, 3.9 $\times 10^{-3}$) | (0.59, 0.38, 0.68) | (0.74, 0.33, 0.82) | (0.94, 0.74, 0.49) | (31, 0.21, 0.25) $\times 10^{-2}$ |
| Commander . | (70, 1.4, 0.3) $\times 10^{-3}$ | (12, 0.66, 0.61) $\times 10^{-2}$ | (0.12, 0.19, 0.71) | (0.70, 0.40, 0.19) | (88, 2.0, 0.92) $\times 10^{-2}$ |
| | | | Minimum $p$-values | | |
| SMICA . . . . . | (0.02, 0.14, 0.20) | (0.25, 0.23, 0.62) | (0.47, 0.21, 0.54) | (0.59, 0.72, 0.86) | (0.27, 0.94, 0.68) |
| SEVEM . . . . . | (0.32, 0.43, 0.98) | (0.96, 0.84, 0.45) | (0.08, 0.75, 0.81) | (0.18, 0.70, 0.64) | (0.62, 0.68, 0.53) |
| NILC . . . . . | (8.1 $\times 10^{-3}$, 0.05, 0.27) | (0.30, 0.21, 0.67) | (0.32, 0.54, 0.93) | (0.73, 0.73, 0.95) | (0.89, 0.74, 0.72) |
| Commander . | (0.21, 0.42, 0.76) | (0.74, 0.99, 0.37) | (0.25, 0.71, 0.63) | (0.73, 0.69, 0.77) | (4.6 $\times 10^{-3}$, 0.25, 0.19) |

**Notes.** We report three smoothing scales ($\sigma_{\rm bin} = 1, \sigma_{\rm bin} = 2, \sigma_{\rm bin} = 3$), for five cases: $TTT$; $TTT$ cleaned of clustered and unclustered point sources; $T2E$; $TE2$; and $EEE$.

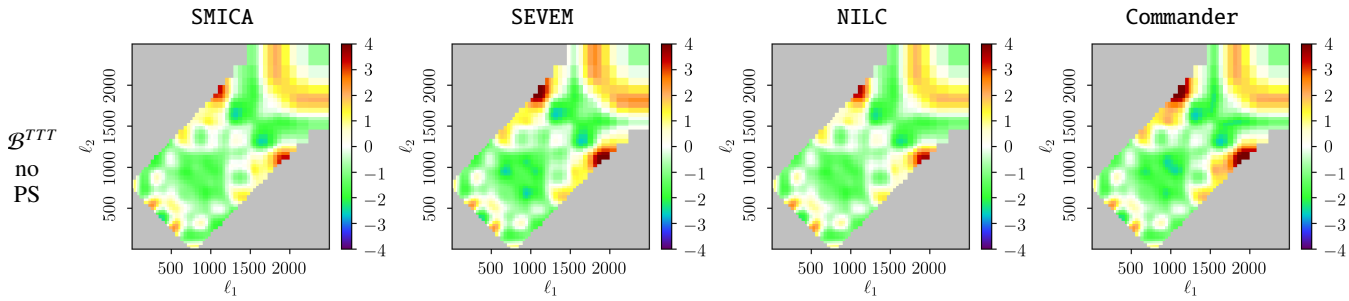


**Fig. 13.** Similar to Fig. 11, but for the $\ell_3$-bin [771,799], and only for $TTT$ cleaned of clustered and unclustered point sources.

this for each possible pair of methods, we obtain six sets of 160 differences $\Delta f_{\rm NL}$ for a given shape.

We then extract the standard deviation of each set, which we call $\sigma_{\Delta f_{\rm NL}}$ (not to be confused with $\sigma_{f_{\rm NL}}$; $\sigma_{\Delta f_{\rm NL}}$ is the standard deviation of the scatter between two cleaning methods, whereas $\sigma_{f_{\rm NL}}$ is simply the $f_{\rm NL}$ uncertainty for a given method). The simulations we use do not include any foreground component; therefore, the extracted $\sigma_{\Delta f_{\rm NL}}$ is only due to different frequency weighting and filtering schemes, and to further operations during NG estimation, such as inpainting. We can use these quantities to verify the consistency between the $f_{\rm NL}$ differences obtained from data and from simulations. A large $\Delta f_{\rm NL}$ scatter observed in the data, for a given pair of cleaned maps, compared to the corresponding expectation $\sigma_{\Delta f_{\rm NL}}$, would signal potential residual foreground contamination in at least one of the two maps.

Our results for this test are summarized in Table 16. For each pair of component-separation methods, and the three standard shapes, we show the measured scatter $\Delta f_{\rm NL}$ and the ratio $\Delta f_{\rm NL}/\sigma_\Delta$. In order to assess the statistical significance of the result, we need of course to take into account that a multiplicity correction is necessary, since we are considering six pairs of methods. In Table 17 we report the fraction of starting simulations for which, after the component-separation processing, $\Delta f_{\rm NL}/\sigma_\Delta$ is larger than 1, 2 or 3, for *any* pair of methods. We also report the largest value of $\Delta f_{\rm NL}/\sigma_\Delta$, measured for all combinations and simulations. We see from this table that measured values of $\Delta f_{\rm NL}/\sigma_\Delta$ from the data (Table 16) are not particularly unusual up to $\Delta f_{\rm NL}/\sigma_\Delta \approx 3$.

This leads us to a first interesting observation, namely that polarization-only $f_{\rm NL}$ results show no significant discrepancy among different cleaned maps. This is a large improvement with respect to our 2015 analysis, where in a similar test we found large differences in $EEE$ bispectra for several combinations. Such discrepancies, together with other anomalies, led us in the previous release to warn the reader that all polarization-based $f_{\rm NL}$ measurements were less robust than $TTT$ estimates, and had to be considered as preliminary. This is no longer the case, since both this and other tests (see next section) show that the polarization data are now fully reliable for primordial NG studies. Achieving such reliability was indeed one of the main goals for our analysis in this data release.

Despite consistency in polarization, somewhat surprisingly we now find some relevant discrepancies in $T$-only results, even though limited to specific methods and the orthogonal shape only. The most striking example is a large difference between the orthogonal $T$-only $f_{\rm NL}$ measurements obtained with SMICA and Commander ($\Delta f_{\rm NL}/\sigma_\Delta \approx 10$). Smaller but non-negligible are also the orthogonal $T$-only discrepancies for the NILC – Commander and SMICA – SEVEM pairs (both with $\Delta f_{\rm NL}/\sigma_\Delta \approx 4$). These results have been cross-checked using the Binned pipeline, finding agreement between estimators: the Binned estimator finds $\Delta f_{\rm NL}/\sigma_\Delta$ of approximately 8, 4, and 4, for these three cases, respectively.

As anticipated in Sect. 5.1, a closer inspection shows that these differences are less worrisome that they might appear at first glance. First of all, a comparison with our 2013 and 2015 results shows that SMICA and NILC orthogonal $TTT$ measurements have remained very stable, while SEVEM and especially Commander display significant changes in this data release. Considering that *all* pipelines agreed very well and displayed robustness to a large number of validation tests in temperature in both previous data releases, we conclude that the latter two methods can be identified as the sources of the current orthogonal $T$-only discrepancy. This is good news, since the main component-separation method that we have focused on for $f_{\rm NL}$ analysis (including in PCNG13 and PCNG15) is SMICA.

**Table 16.** Comparison between local, equilateral, and orthogonal $f_{\rm NL}$ results, obtained using the four different component-separation pipelines and the Modal 1 bispectrum estimator.

| Methods | Local | | Equilateral | | Orthogonal | |
|---|---|---|---|---|---|---|
| | $\Delta$ | $\Delta/\sigma$ | $\Delta$ | $\Delta/\sigma$ | $\Delta$ | $\Delta/\sigma$ |
| SMICA – SEVEM | | | | | | |
| $T$ | −0.1 | 0.0 | −11 | −1.2 | −24 | −4.1 |
| $E$ | −5.4 | −0.4 | −59 | −1.1 | −5.0 | −1.8 |
| $T$+$E$ | −1.1 | −0.6 | −10 | −1.0 | −19 | −3.8 |
| SMICA – NILC | | | | | | |
| $T$ | 0.1 | 0.1 | 16 | 2.4 | −17 | −3.0 |
| $E$ | 25 | 1.8 | 123 | 2.1 | −66 | −2.4 |
| $T$+$E$ | 0.3 | 0.2 | 17 | 1.8 | −12 | −2.4 |
| SMICA – Commander | | | | | | |
| $T$ | 1.7 | 2.1 | −4.5 | −0.5 | −43 | −10.0 |
| $E$ | −0.2 | 0.0 | 4.9 | 0.1 | −32 | −1.4 |
| $T$+$E$ | 0.9 | 0.9 | −10 | −1.1 | −23 | −5.2 |
| SEVEM – NILC | | | | | | |
| $T$ | 0.2 | 0.2 | 27 | 2.6 | 7.6 | 1.2 |
| $E$ | 30 | 2.2 | 183 | 2.7 | −6.9 | −0.2 |
| $T$+$E$ | 1.4 | 1.2 | 27 | 2.1 | 7.7 | 1.3 |
| SEVEM – Commander | | | | | | |
| $T$ | 1.8 | 0.8 | 6.5 | 0.6 | −19 | −2.6 |
| $E$ | 5.2 | 0.4 | 64 | 1.1 | 26 | 0.8 |
| $T$+$E$ | 1.9 | 0.9 | −0.4 | 0.0 | −4.4 | −0.7 |
| NILC – Commander | | | | | | |
| $T$ | 1.6 | 0.7 | −20 | −2.0 | −26 | −3.8 |
| $E$ | −25 | −1.8 | −119 | −1.6 | 33 | 1.0 |
| $T$+$E$ | 0.6 | 0.3 | −27 | −2.1 | −12 | −2.0 |

**Notes.** We calculate the difference $\Delta$ in the sense $f_{\rm NL}$(method 1) − $f_{\rm NL}$(method 2). For each pair of cleaned maps, we start by considering actual data and we compute the scatter $\Delta$ in our estimates of local, equilateral, and orthogonal $f_{\rm NL}$. We then compute the ratio of $\Delta$ over its standard deviation, obtained using 160 FFP10 simulations. These ratios have to be compared with the benchmarks provided in Table 17.

**Table 17.** Fraction of simulations for which the differences between pairs of component-separation methods are above various levels of $\sigma$.

| Data set | Local | | | | Equilateral | | | | Orthogonal | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\Delta f_{\rm NL}/\sigma_\Delta{}^{\rm max}$ | $N_{\rm sims}^{>1\sigma}$ | $N_{\rm sims}^{>2\sigma}$ | $N_{\rm sims}^{>3\sigma}$ | $\Delta f_{\rm NL}/\sigma_\Delta{}^{\rm max}$ | $N_{\rm sims}^{>1\sigma}$ | $N_{\rm sims}^{>2\sigma}$ | $N_{\rm sims}^{>3\sigma}$ | $\Delta f_{\rm NL}/\sigma_\Delta{}^{\rm max}$ | $N_{\rm sims}^{>1\sigma}$ | $N_{\rm sims}^{>2\sigma}$ | $N_{\rm sims}^{>3\sigma}$ |
| $T$ | 3.7 | 0.66 | 0.18 | 0.006 | 3.3 | 0.74 | 0.18 | 0.006 | 3.5 | 0.72 | 0.16 | 0.019 |
| $E$ | 3.6 | 0.72 | 0.16 | 0.025 | 3.2 | 0.71 | 0.16 | 0.013 | 3.3 | 0.76 | 0.19 | 0.019 |
| $T$+$E$ | 2.9 | 0.66 | 0.13 | 0 | 3.3 | 0.64 | 0.15 | 0.013 | 3.7 | 0.65 | 0.16 | 0.019 |

**Notes.** We use 160 FFP10 simulations (without foreground residuals) to compute the standard deviation $\sigma_{\Delta f_{\rm NL}}$ of the measured $f_{\rm NL}$ scatter, obtained by processing a given realization through the different component-separation pipelines. We report the largest measured significance $\Delta f_{\rm NL}/\sigma_{\Delta f_{\rm NL}}$, across all method pairs and simulations. We then report the fraction of the total number of simulations for which at least one method pair returns a value $\Delta_{f_{\rm NL}}$ larger than $1\,\sigma_{\Delta f_{\rm NL}}$, $2\,\sigma_{\Delta f_{\rm NL}}$, or $3\,\sigma_{\Delta f_{\rm NL}}$. These numbers provide a benchmark against which to assess the significance of the scatter measured on the data (Table 16), taking into account the multiple comparisons.

It is also important to stress that all significant discrepancies are specifically confined to the orthogonal $TTT$ case. The other shapes and also a mode-by-mode or bin-by-bin correlation analysis over the full bispectrum domain (see next section) show no other signs of anomalies for any component-separation method. Even considering the largest discrepancy, arising from the SMICA – Commander pair, this still amounts to just a $1\sigma$ deviation in $f_{\rm NL}{}^{\rm ortho}$. This is due to the fact that the level of agreement displayed by different cleaning methods on simulations typically amounts to a small fraction of the $f_{\rm NL}$ errors. The implications for inflationary constraints of shifting $f_{\rm NL}{}^{\rm ortho}$ by $1\sigma$ are essentially negligible. As mentioned earlier, the changes in orthogonal $f_{\rm NL}$ coming from Commander can be explained through the unavailability of detector-set maps for the current release.

The main conclusion to be drawn from the validation tests described in this section is that the reliability of our final $T$+$E$ $f_{\rm NL}$ results has significantly increased with respect to the previous release, thanks to a clear improvement in the robustness of the polarization data. The issues we find in the temperature data are confined to the orthogonal shape and to specific component-separation pipelines, not affecting the final SMICA measurements used for inflationary constraints.

Of course, any considerations in this section that might lead us to a preference for specific component-separation methods, apply only to primordial NG analysis, and not to other cosmological or astrophysical analyses. There is no generally preferred cleaning method for all applications, and separate assessments should be conducted case by case.

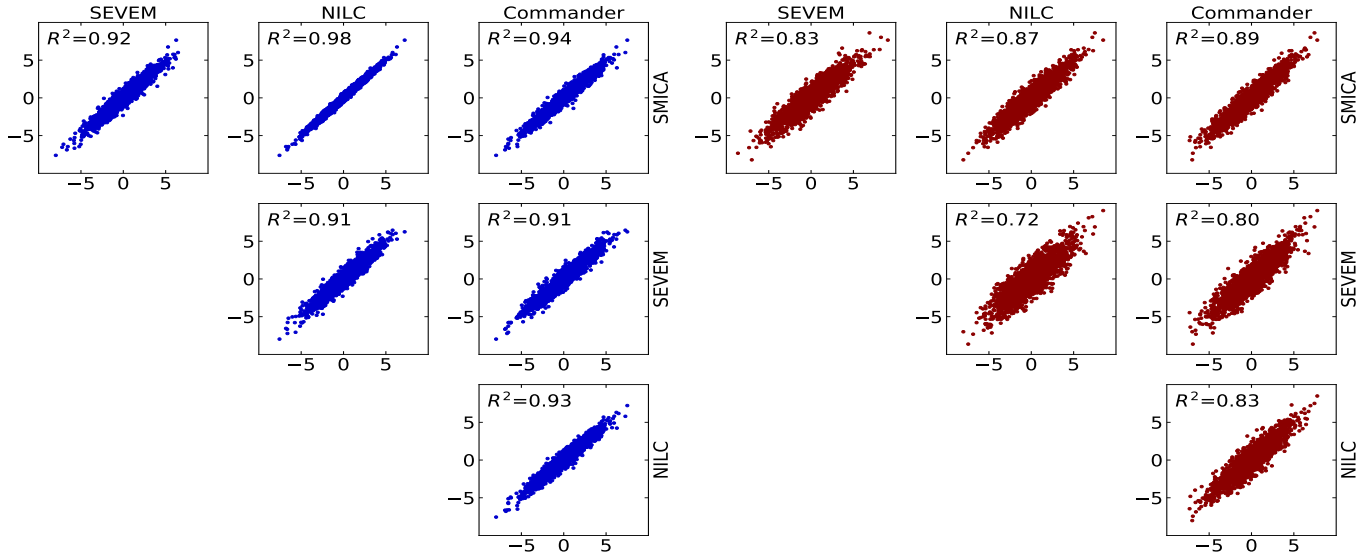


**Fig. 14.** Scatter plots of the 2001 Modal 2 coefficients for each combination of component-separation methods, labelled with the $R^2$ coefficient of determination. The figures on the *left* are for the temperature modes and those on the *right* for pure polarization modes (the Modal 2 pipeline reconstructs the component of the *EEE* bispectrum that is orthogonal to *TTT*, so this is not exactly the same as the *EEE* bispectrum of the other estimators).

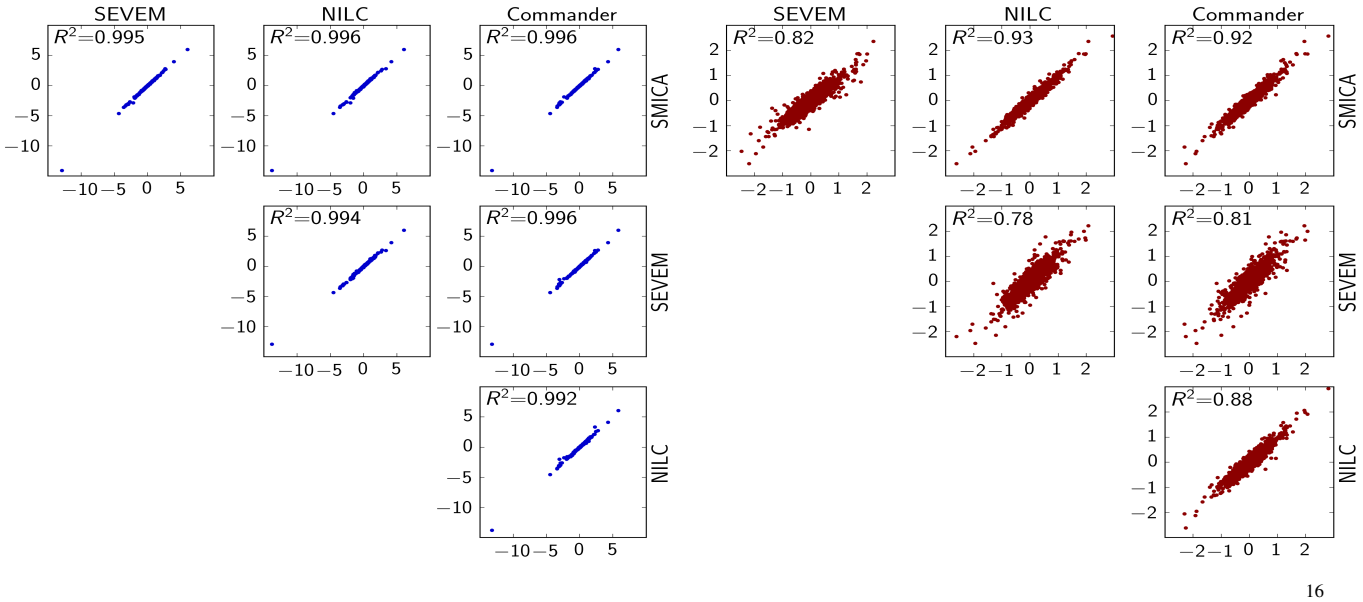


**Fig. 15.** Scatter plots of the bispectrum values in each bin-triplet (13 020 for *TTT*, all of which have been multiplied by $10^{16}$, and 11 221 for *EEE*, all of which have been multiplied by $10^{20}$) for all combinations of component-separation methods, with the $R^2$ coefficient of determination. The figures on the *left* are for *TTT* and those on the *right* for *EEE*.

### 6.1.2. Comparison between reconstructed bispectra

The local, equilateral, and orthogonal directions already cover a significant part of the entire bispectrum domain, and deserve special attention, since they are crucial for inflationary constraints. It is nevertheless useful to also check the agreement between component-separation pipelines in a more general, model-independent fashion. For this purpose we calculate the coefficient of determination (see Allen 1997), denoted by $R^2$, from the mode or bin amplitudes extracted from different foreground-cleaned maps. The coefficient of determination is a standard statistical measure of what proportion of the variance of one variable can be predicted from another. For example in our case a score of 0.9 for `SMICA` – `SEVEM` would mean that 90% of the mode or bin amplitudes measured in `SMICA` could be explained by the mode or bin amplitudes measured in `SEVEM` and 10% of the amplitudes would be unexplained.

Scatter plots of modes are shown in Fig. 14, while those for bins are shown in Fig. 15. For *TTT* the lowest Modal coefficient of determination is 0.91 (for `SEVEM` – `NILC`), while for the Binned bispectrum values all *TTT* coefficients of determination are larger than 0.99. For *EEE* the lowest Modal value is 0.71, again for `SEVEM` – `NILC`, while the corresponding Binned value is 0.78. The lowest Binned value overall is 0.71, for `SMICA` – `SEVEM` *ETT* (not shown in the figure). It is also possible to see

**Table 18.** Results of the noise mismatch test described in Sect. 6.2.

| Method | $T$ only | | $E$ only | | $T$+$E$ | |
|---|---|---|---|---|---|---|
| | $\sigma_{f_{\rm NL}}$ | $\sigma_{\Delta f_{\rm NL}}$ | $\sigma_{f_{\rm NL}}$ | $\sigma_{\Delta f_{\rm NL}}$ | $\sigma_{f_{\rm NL}}$ | $\sigma_{\Delta f_{\rm NL}}$ |
| $\ell$ space . . . . . | 5.8 | 0.1 | 26 | 3.5 | 4.7 | 0.2 |
| $A_{\rm all} = 0.2$ . . . . . | 5.8 | 0.6 | 26 | 5.1 | 4.8 | 0.6 |
| $A = 0.1$ . . . . . | | | 26 | 3.3 | 4.7 | 0.2 |
| $A = 0.2$ . . . . . | | | 28 | 5.8 | 4.7 | 0.3 |
| $A = 0.3$ . . . . . | | | 29 | 11.0 | 4.7 | 0.6 |

**Notes.** We consider FFP10 simulations and add extra noise to them. We then measure $f_{\rm NL}^{\rm local}$ using the Modal 1 pipeline, normalized and calibrated without the extra-noise component. We verify how this affects the local $f_{\rm NL}$ errors ($\sigma_{f_{\rm NL}}$ columns) and measure the standard deviation of the scatter $f_{\rm NL}^{1,i} - f_{\rm NL}^{2,i}$, where $f_{\rm NL}^{2,i}$ and $f_{\rm NL}^{1,i}$ are the measurements extracted with and without noise mismatch, respectively, for the $i$th simulation ($\sigma_{\Delta f_{\rm NL}}$ columns). Results in the first row are obtained by generating extra noise in harmonic space, while in the second row we generate extra noise in pixel space, both in temperature and polarization. In the following rows, we still generate noise in pixel space, but leave temperature noise maps unchanged, while scaling the polarization rms by the given factors.

the impact of excluding the lowest [2, 3] bin from the analysis for $E$; if it were included that value would drop from 0.71 to 0.59.

These results show a very good level of agreement between all methods. Again, we see a large improvement in the polarization maps, by comparing with a similar test performed in PCNG15. Also interesting is the fact that no anomalies show up in temperature data for any specific mode or bin. The issues discussed in the previous section, which affected only specific component-separation pipelines, seem also completely confined to the combination of modes and bins that selects the orthogonal shape region of the bispectrum domain, in such a way as to be invisible in this more general analysis.

### 6.2. Testing noise mismatch

Accurate tests of the FFP10 maps have found some level of mismatch between the noise model adopted in the simulations and the actual noise levels in the data (Planck Collaboration III 2020; Planck Collaboration IV 2020). In temperature this is roughly at the 3% level in the noise power spectrum at $\ell \approx 2000$. Percent level differences are also seen in polarization. This issue raises some concern for our estimators, since we use simulations to calibrate them. We thus decided to perform some simulation-based tests to check to what extent this noise mismatch may affect our Monte Carlo errors (as long as the noise is Gaussian, noise mismatches of this kind cannot bias the estimators). For all the work in this section we consider SMICA maps only.

Uncertainties in non-Gaussianity parameters might be affected by two effects. One is the suboptimality of the estimator weights in the cubic term, if the power spectra extracted from simulations do not match the data. The other is an imperfect Monte Carlo calibration of the linear correction term, leading to the inability to fully correct for anisotropic and correlated noise features. Given that we are considering percent-level corrections, we do not expect the former of the two effects to be of particular significance; this is also confirmed by past analyses, in which the cosmological parameters have been updated several times, leading to percent changes in the fiducial power spectrum, without any appreciable difference in the $f_{\rm NL}$ results. The effect on the linear term is harder to assess, and to some

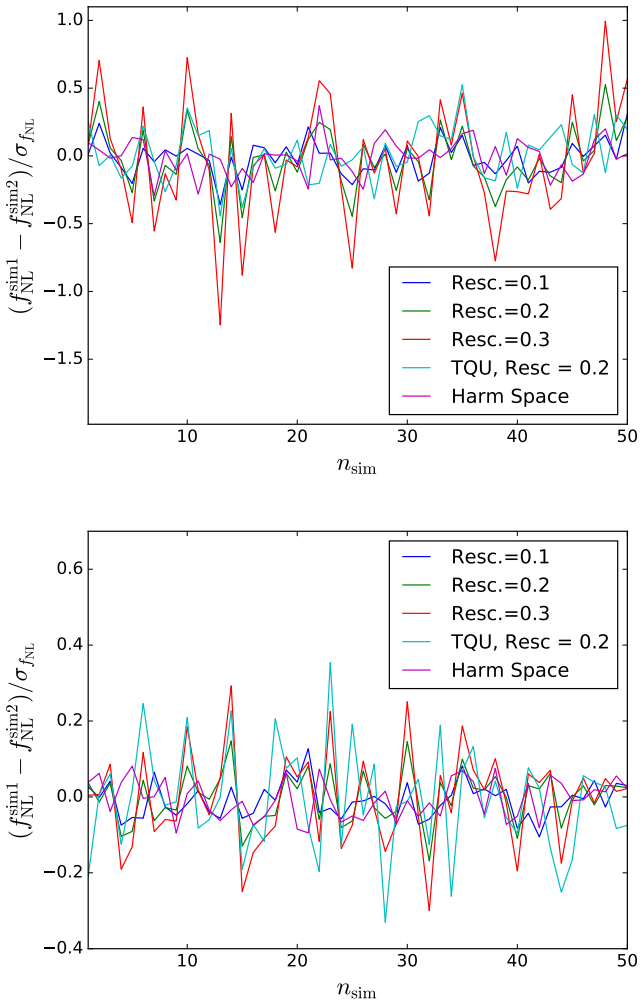


**Fig. 16.** Effects on $f_{\rm NL}$ scatter, simulation-by-simulation, of different levels of noise mismatch, as described in Sect. 6.2. *Top*: *EEE*-only results. *Bottom*: combined $T$+$E$ results. The magenta line represents the case in which we generate an extra-noise component in harmonic space, both in temperature and polarization, with variance equal to the difference between the noise spectra in the data and the simulations. All other lines represent cases in which we generate extra-noise maps in pixel space, with the rms per pixel extracted from simulations and rescaled by different factors, as specified in the legend; the label "TQU" further specifies the pixel-space test in which extra noise is added to both temperature and polarization data, unlike in the other three pixel-space cases, in which we include only a polarization extra-noise component.

extent it depends on the specific noise correlation properties in the data, which are possibly not fully captured in the simulations. In general, it is reasonable to expect a power spectrum mismatch of a few percent to have only a small impact, unless significant spatial correlations between large and small scales are present in the data and not captured by the simulation noise model. The linear term correction is generally dominated by the mask, in the mostly signal-dominated regime we consider for our analysis.

As a first test, we generate "extra-noise" multipoles, drawing them from a zero-mean Gaussian distribution having as power spectrum the difference between the simulation and data noise power spectra. We then add these realizations to the original noise simulations and apply our estimators to this extra-noise mock data set. However, we calibrate our linear term, estimator normalization, and weights using the original noise maps. We check the effect of the mis-calibration through a comparison

between the $f_{\rm NL}$ measurements obtained in this way and those extracted from the original realizations, without extra noise included. As a figure of merit, we consider the local $f_{\rm NL}$ results, since they are generally most sensitive to noise features. At the end of this analysis, we verify that the impact of the noise mismatch is very small: the uncertainties change by a negligible amount with respect to the exactly calibrated case, while the scatter between $f_{\rm NL}$ measured with and without extra noise in the maps is zero on average, as expected, and has a standard deviation $\sigma_{\Delta f_{\rm NL}} \approx \sigma_{f_{\rm NL}}/10$.

Our results are shown in detail in the first row of Table 18 and in Fig. 16. The former reports the $f_{\rm NL}$ error bars and the standard deviation of map-by-map $f_{\rm NL}$ differences, while the latter shows the map-by-map $f_{\rm NL}$ scatter for 50 realizations. These very small deviations were largely expected, given the small change in the amplitude of the noise power spectrum we are considering.

In the test we have just described we generate uncorrelated extra-noise multipoles (with a non-white spectrum). On the other hand, we know that the estimator is most sensitive to couplings between large and small scales, which require a properly calibrated linear term correction. Therefore, we decided to also test the impact of a possible linear term miscalibration of this type, by generating and studying an extra-noise component directly in pixel space. In this case, $\ell$-space correlations between large and small scales arise due the spatially anisotropic distribution of the noise. We proceed as follows. After extracting the noise rms of the `SMICA` FFP10 polarization simulations, we rescale it by a fixed factor $A$, and we use this rescaled rms map to generate new Gaussian "extra-noise" realizations in pixel space, which we add to the original noise maps. We consider different cases. Firstly, we take a rescaling factor $A_{TQU} = 0.2$ for both temperature and polarization noise maps. We then perform a more detailed study of the effect on polarization maps only. For this purpose, we leave the temperature noise unchanged and rescale the polarization rms noise by factors ranging from $A = 0.1$ up to $A = 0.3$.

We see again from the summary of results reported in Table 18 that the $f_{\rm NL}$ error change is always very small. The same can be said of the standard deviation of the $f_{\rm NL}$ scatter between realizations with and without extra noise, which reaches at most a value $\sigma_{\Delta f_{\rm NL}} \approx \sigma_{f_{\rm NL}}/3$, for a large $A = 0.3$. These results provide a good indication that a noise mismatch between simulations and data is not a concern for primordial NG estimation, unless the mismatch itself is well above the estimated percent level in the noise power spectrum and produces large correlations between small and large scales. It is also worth mentioning that, in the very early stages of our primordial NG analysis of *Planck* data, when accurate simulations were not yet available, we calibrated our estimators using very simple noise models (e.g. we initially generated noise in harmonic space, with a non-flat power spectrum consistent with the data, but neglecting any correlations between different scales; we then went to pixel space and modulated the noise map with the hit-count map to anisotropize it). Despite the simplicity of this approach, we were able to verify later, using FFP simulations, that such simple models already produced an accurate linear-term correction. This further reinforces our confidence in the robustness of $f_{\rm NL}$ estimators to imperfect modelling of the noise.

### 6.3. Effects of foregrounds

Here we look at two non-primordial contributions to the *Planck* bispectrum, namely Galactic dust and the Sunyaev-Zeldovich effect, which could potentially be present in the cleaned maps, but which we do not detect in the end. Another potential contaminant for both primordial and lensing bispectrum results is the "intrinsic bispectrum," induced in the CMB by weak (second-order) non-linearities from gravity in general relativity and by non-linearities in the recombination physics. This would set the minimal level of CMB NG present even for Gaussian initial conditions of the primordial curvature perturbation. However, this is of no particular concern to us, since for the *Planck* data set its expected impact is very small both in temperature (Bartolo et al. 2004b, 2005, 2010c, 2012; Creminelli & Zaldarriaga 2004; Boubekeur et al. 2009; Nitta et al. 2009; Bartolo & Riotto 2009; Senatore et al. 2009; Khatri & Wandelt 2009, 2010; Creminelli et al. 2011; Huang & Vernizzi 2013; Su et al. 2012; Pettinari et al. 2014) and polarization (Lewis 2012; Pettinari et al. 2014). Non-primordial contributions that we *do* detect (lensing and extragalactic point sources) were discussed in Sect. 4.

#### 6.3.1. Non-Gaussianity of the thermal dust emission

In principle, the foreground-cleaned maps should not contain any noticeable NG of Galactic origin. However, in raw observations the strongest contamination to the primordial NG is due to the thermal dust emission, which induces a large negative bias in the measurements of $f_{\rm NL}^{\rm local}$. Therefore, it is important to verify that this contamination has been removed entirely through the different component-separation methods.

There is no analytical template for the dust bispectrum, unlike the extragalactic templates discussed in Sect. 4.2. A simple method using instead numerical templates for the different Galactic foregrounds has been described in Jung et al. (2018) (see also Coulton & Spergel 2019). First, using the Binned bispectrum estimator, the dust bispectral shape is computed from the thermal dust emission map at 143 GHz (the dominant frequency channel in the cleaned CMB maps) produced by the `Commander` technique. This is actually a map determined at higher frequencies, where the dust dominates, and then rescaled to 143 GHz. Since there was no improvement for the temperature map of this foreground in the latest release, we use the *Planck* 2015 temperature map here. Then, this numerical dust bispectrum can be used as a theoretical template in the analysis of any other data map with the Binned bispectrum estimator. The only condition is that the same mask, beam, and binning are used for both the determination of the dust bispectrum and the analysis itself.

An illustration of the large bias induced by the presence of dust in the map is given in Table 19. It shows the $f_{\rm NL}$ parameters of the primordial local shape and the thermal dust bispectrum for both an independent and a joint analysis of the raw 2018 143-GHz *Planck* temperature map. As a reminder, in an independent analysis we assume that only one of the templates is present in the data (and we repeat the analysis for each individual template), while in a fully joint analysis we assume that all templates are present, so that we have to take their correlations into account. In the independent case, there is a strong detection of local NG (at about the $5\sigma$ level), while the dust is observed at the expected level (within the $1\sigma$ interval centred on the expected value: $f_{\rm NL}^{\rm dust} = 1$). The joint analysis shows that indeed the large negative $f_{\rm NL}^{\rm local}$ of the independent case is entirely due to the presence of dust in the map, while the dust is still detected at the expected amount.

The analysis of the $T$-only maps produced by the four component-separation methods is the main result of this

**Table 19.** Independent and joint estimates (see the main text for a definition) of the $f_{\rm NL}$ parameters of the primordial local shape and the thermal dust bispectrum in the raw *Planck* 143-GHz temperature map, determined using the Binned bispectrum estimator.

| Shape | Independent | Joint |
|---|---|---|
| Local . . . . . | −64 ± 13 | 8 ± 20 |
| Dust . . . . . | 1.21 ± 0.35 | 1.32 ± 0.52 |

**Notes.** Uncertainties are 68 % CL. Uncertainties in this table only are Fisher forecasts multiplied by a factor larger than 1, which depends on the shape, due to the small breaking of the weak non-Gaussianity approximation in this map. Here, we use the same factors as in Jung et al. (2018), which were determined by comparing the observed errors and Fisher forecasts in the analysis of 100 Gaussian CMB maps to which the dust map was added.

subsection. Table 20 gives the values of $f_{\rm NL}$ for several primordial NG shapes (local, equilateral, and orthogonal), the amplitudes of some extragalactic foreground bispectra (unclustered point sources and the CIB, as defined in Sect. 4.2) and the $f_{\rm NL}$ of the thermal dust emission bispectrum, after subtracting the lensing bias, in both an independent and a fully joint analysis. There is no significant detection of dust in any of the four maps (the worst case being SMICA with slightly more than a $1\sigma$ deviation)[11]. Given the small size of the $f_{\rm NL}^{\rm dust}$ errors, this non-detection means that there is at most a few percent of dust contamination in the cleaned maps (outside the mask). However, the errors of the local shape increase significantly in the joint analysis because the dust and the local shapes are quite correlated (more than 60%).

#### 6.3.2. Impact of the tSZ effect

The SMICA component-separation method also produces a foreground-cleaned temperature map, "SMICA no-SZ," where, in addition to the usual foregrounds, contamination by the thermal Sunyaev-Zeldovich (tSZ) effect has been subtracted as well (see Planck Collaboration IV 2020). This allows us to test if the tSZ contamination has any significant impact on our primordial results (we already saw in Sect. 4.1 that it does seem to have an impact on the determination of the lensing NG). This is an important test, since there have been recent claims in the literature (Hill 2018) that it might have an effect.

The results of the analysis can be found in Table 21. Because this effect is only important in temperature, we restrict ourselves to a $T$-only analysis. Results have been determined with the KSW estimator, with the Binned estimator (this time using only 150 maps for the linear correction and the error bars instead of 300, which is enough for the purposes of this test), and with the Modal 1 estimator. The mask used is the same as for the main analysis. The table also contains the difference with the result determined from the normal SMICA temperature map (without tSZ removal) and the uncertainties on this difference.

We see a shift of about $1\sigma_{\Delta f_{\rm NL}}$ (hence insignificant) in the local shape result, which together with orthogonal is the shape predicted to be most contaminated by the tSZ effect (see Hill 2018). We actually see a larger shift in the equilateral result, which is supposed to be almost unaffected by this effect, while the orthogonal shift is the largest, at more than $2\sigma_{\Delta f_{\rm NL}}$ for all estimators. However, such a marginal effect in the orthogonal shape without a corresponding effect in the local shape leads us to conclude that we do not detect any significant impact of contamination of the usual foreground-corrected maps by the tSZ effect on our primordial NG results. In other words, while some tSZ contamination, peaking in the squeezed limit, is expected to be present in the standard temperature maps, this is too small to be clearly disentangled from the statistical fluctuations in the $f_{\rm NL}$ results; in other words the tSZ contamination is not bigger than effects due to the different processing of the data when tSZ is included in the foreground components for the SMICA analysis. Furthermore, all shifts discussed here are much smaller than the uncertainties on the $f_{\rm NL}$ values themselves, due to the fact that the $f_{\rm NL}$ scatter between different cleaned maps ($\sigma_{\Delta f_{\rm NL}}$) is much smaller than the $f_{\rm NL}$ error ($\sigma_{f_{\rm NL}}$).

Finally, it should also be pointed out that, using the same criteria, we cannot strictly call the shift in $f_{\rm NL}^{\rm lens}$ discussed in Sect. 4.1 significant either. The observation that all $TTT$ measurements are below the expected value $f_{\rm NL}^{\rm lens} = 1$ and that adding polarization systematically shifts them up, does point to some tSZ contamination in $T$-only results for the lensing bispectrum. However, the analysis discussed in this section finds again only a $2\sigma_{\Delta f_{\rm NL}}$ effect in this case; this is slightly larger or smaller than the significance of the orthogonal $f_{\rm NL}$ shift, depending on the estimator. Again, intrinsic statistical uncertainties make it hard to detect this systematic effect.

### 6.4. Dependence on sky coverage

The temperature and polarization mask we are using have been determined to be the optimal masks for use on the maps produced by the component-separation pipelines, according to criteria explained in Planck Collaboration IV (2020), and hence are used for all *Planck* analyses on those maps. However, the choice of mask can have an impact on the results for $f_{\rm NL}$. In the first place the sky fraction of the mask has a direct effect on the size of the uncertainties. Potentially more important, however, is the effect the mask might have on the amount of foreground residuals. Hence we judge it important to investigate the impact of the choice of mask on our results by comparing results for several different masks. All tests are performed on the SMICA maps (with one exception, detailed below) using the Binned bispectrum estimator, using 150 maps for the linear correction and the errors.

As a first test, performed on the temperature map only, we take the union of the mask used in this paper ($f_{\rm sky} = 0.78$) with the mask we used in our 2015 analysis ($f_{\rm sky} = 0.76$), leading to a mask that leaves a fraction $f_{\rm sky} = 0.72$ of the sky uncovered. The results for the standard primordial shapes are given in Table 22, as well as their differences with the results using the standard mask (see Table 5). The errors on the differences have been determined from the scatter among 150 simulations when analysed with the two different masks. This particular test is performed both on the SMICA and the Commander maps, since its initial purpose was a further check on the discrepancy between those two component-separation methods regarding the $T$-only orthogonal result, discussed in detail in Sect. 6.1.1. We see, however, that the orthogonal result is very stable for both methods, while the local result shifts by about $1\sigma_{\Delta f_{\rm NL}}$ and the equilateral result moves by about $2\sigma_{\Delta f_{\rm NL}}$, which corresponds to

[11] While no dust is detected in the cleaned maps using the template determined from the dust map, it should be pointed out that there is no guarantee that the dust residuals (or negative dust residuals in the case of an oversubtraction), after passing through the component-separation pipelines, have exactly the same form as the original dust bispectrum. However, it seems reasonable to assume that the resulting shape would still be highly correlated with the original dust template, so that this remains a meaningful test.

**Table 20.** Independent and joint estimates of the $f_{\rm NL}$ parameters of the indicated shapes, including in particular the dust template, for the cleaned maps produced by the four component-separation methods, as determined with the Binned bispectrum estimator.

| | SMICA | | SEVEM | | NILC | | Commander | |
|---|---|---|---|---|---|---|---|---|
| Shape | Independent | Joint | Independent | Joint | Independent | Joint | Independent | Joint |
| $f_{\rm NL}^{\rm local}$ | $-0.1 \pm 5.6$ | $5.0 \pm 8.4$ | $0.0 \pm 5.7$ | $1.7 \pm 8.7$ | $0.0 \pm 5.6$ | $5.2 \pm 8.5$ | $-1.3 \pm 5.6$ | $3.1 \pm 8.3$ |
| $f_{\rm NL}^{\rm equil}$ | $26 \pm 69$ | $5 \pm 73$ | $43 \pm 70$ | $30 \pm 74$ | $5 \pm 69$ | $-12 \pm 73$ | $32 \pm 69$ | $20 \pm 73$ |
| $f_{\rm NL}^{\rm ortho}$ | $-11 \pm 39$ | $-5 \pm 44$ | $8 \pm 39$ | $13 \pm 45$ | $4 \pm 39$ | $13 \pm 45$ | $29 \pm 39$ | $35 \pm 44$ |
| $b_{\rm PS}/(10^{-29})$ | $6.3 \pm 1.0$ | $5.0 \pm 2.7$ | $9.7 \pm 1.1$ | $7.1 \pm 2.9$ | $5.7 \pm 1.1$ | $5.4 \pm 2.7$ | $5.4 \pm 1.0$ | $3.6 \pm 2.6$ |
| $A_{\rm CIB}/(10^{-27})$ | $3.0 \pm 0.5$ | $0.6 \pm 1.3$ | $4.6 \pm 0.5$ | $1.3 \pm 1.4$ | $2.6 \pm 0.5$ | $0.1 \pm 1.3$ | $2.6 \pm 0.5$ | $0.9 \pm 1.3$ |
| $f_{\rm NL}^{\rm dust}/(10^{-2})$ | $6.6 \pm 4.4$ | $6.7 \pm 5.9$ | $4.8 \pm 4.6$ | $1.9 \pm 6.1$ | $4.8 \pm 4.4$ | $5.1 \pm 5.9$ | $4.4 \pm 4.3$ | $3.1 \pm 5.7$ |

**Table 21.** Impact of the thermal Sunyaev-Zeldovich (tSZ) effect on $f_{\rm NL}$ estimators for temperature data.

| | SMICA no-SZ | | | Difference | | |
|---|---|---|---|---|---|---|
| Shape | KSW | Binned | Modal 1 | KSW | Binned | Modal 1 |
| Local | $1.3 \pm 5.6$ | $1.1 \pm 5.9$ | $0.9 \pm 5.8$ | $-1.8 \pm 1.4$ | $-1.3 \pm 1.6$ | $-1.5 \pm 1.3$ |
| Equilateral | $-6 \pm 67$ | $-6 \pm 67$ | $5 \pm 64$ | $14 \pm 14$ | $32 \pm 17$ | $19 \pm 14$ |
| Orthogonal | $-36 \pm 37$ | $-31 \pm 42$ | $-23 \pm 39$ | $21 \pm 8$ | $20 \pm 9$ | $19 \pm 7$ |

**Notes.** First three columns: results for the $f_{\rm NL}$ parameters of the primordial local, equilateral, and orthogonal shapes, determined by the indicated estimators from the SMICA foreground-cleaned temperature map with additional removal of the tSZ contamination. Last three columns: difference with the result from the normal SMICA map (see Table 5), where the error bar is the standard deviation of the differences in $f_{\rm NL}$ for the Gaussian FFP10 simulations when processed through the two different SMICA foreground-cleaning procedures. Results have been determined using an independent single-shape analysis and are reported with subtraction of the lensing bias; error bars are 68 % CL.

**Table 22.** Tests of dependence on choice of mask.

| Shape | Combined mask | Difference |
|---|---|---|
| | SMICA $T$ | |
| Local | $-1.8 \pm 6.0$ | $-1.7 \pm 1.6$ |
| Equilateral | $-23 \pm 71$ | $-49 \pm 24$ |
| Orthogonal | $-13 \pm 42$ | $-1 \pm 12$ |
| | Commander $T$ | |
| Local | $-3.2 \pm 6.0$ | $-1.9 \pm 1.6$ |
| Equilateral | $-11 \pm 72$ | $-42 \pm 24$ |
| Orthogonal | $26 \pm 43$ | $-3 \pm 12$ |

**Notes.** The first column gives results for the $f_{\rm NL}$ parameters of the primordial local, equilateral, and orthogonal shapes, determined using the Binned estimator on the SMICA and Commander foreground-cleaned temperature maps, using as a mask the union of the 2018 and 2015 common masks. The second column gives the differences from the results using the 2018 common mask (see Table 5), where the error is the standard deviation of the differences in $f_{\rm NL}$ for 150 simulations when analysed with the two masks. Results have been determined using an independent single-shape analysis and are reported with subtraction of the lensing bias.

about $(2/3)\sigma_{f_{\rm NL}}$. Checking the 150 Gaussian simulations used for determining the errors and the linear correction, we see that there are 20 that have at least one fluctuation larger than $2\sigma_{\Delta f_{\rm NL}}$ in at least one of the three shapes, which corresponds to a 13.3% probability. This is large enough that we consider the shift consistent with a statistical fluctuation. In any case, all results remain consistent with zero. It is interesting to note that this result for the equilateral shape is much closer to the one determined for the 2015 *Planck* data.

As a second test, to check the impact of the size of the polarization mask, we return to the standard temperature mask, but this time we change the polarization mask. It is altered as follows: each hole in the common mask is grown by a region 20 pixels in width. This reduces the sky fraction from 0.78 to 0.73. Results for this test can be found in Table 23, including the differences with the results determined with the standard mask. We see that many results shift around somewhat, but nothing appears very significant (all less than $1.5\sigma_{\Delta f_{\rm NL}}$). The largest is again a $(2/3)\sigma_{f_{\rm NL}}$ shift for the $E$-only equilateral case, moving it closer to zero. This test was also performed with the KSW and Modal 1 estimators, giving consistent (but not identical) results. In particular, while some values shift a bit more, the shift in the equilateral shape is smaller for both the KSW and Modal 1 estimators. However, all estimators agree on the signs of the shifts.

Our third and final test of the effects of mask choice is very similar to the previous one, except that this time the polarization mask is enlarged by an additional 40 (instead of 20) pixels around every hole. This further reduces $f_{\rm sky}$ to 0.66. Results can also be found in Table 23, and we find similar conclusions as for the previous test.

To summarize, while we see some effects on our $f_{\rm NL}$ results when considering different masks, none of these appear to be significant (further reinforced by the fact that the shifts vary somewhat between estimators), nor do they change the conclusions of our paper.

### 6.5. Dependence on multipole number

As in previous releases we also test the dependence of the results for $f_{\rm NL}$ on the choice of $\ell_{\rm min}$ and $\ell_{\rm max}$ used in the analysis. This test is most easily performed using the Binned estimator. Results are shown in Fig. 17 for the dependence on $\ell_{\rm min}$ and in Fig. 18 for the dependence on $\ell_{\rm max}$.

Considering first Fig. 17, for $\ell_{\rm min}$, the plots look very similar to the ones in the paper investigating the 2015 *Planck* data (PCNG15), with increased stability for the $E$-only local results.

**Table 23.** Tests of growing the polarization mask size.

| Shape | Extra mask | Difference |
|---|---|---|
| | Extra 20, $E$ | |
| Local | $36 \pm 28$ | $-12 \pm 17$ |
| Equilateral | $70 \pm 150$ | $-96 \pm 65$ |
| Orthogonal | $-180 \pm 91$ | $-3 \pm 43$ |
| | Extra 20, $T$+$E$ | |
| Local | $-5.0 \pm 5.2$ | $-2.4 \pm 1.8$ |
| Equilateral | $-31 \pm 48$ | $-12 \pm 13$ |
| Orthogonal | $-23 \pm 25$ | $11 \pm 7$ |
| | Extra 40, $E$ | |
| Local | $41 \pm 33$ | $-7 \pm 22$ |
| Equilateral | $16 \pm 170$ | $-150 \pm 85$ |
| Orthogonal | $-160 \pm 100$ | $20 \pm 61$ |
| | Extra 40, $T$+$E$ | |
| Local | $-2.4 \pm 5.2$ | $0.2 \pm 2.2$ |
| Equilateral | $-30 \pm 50$ | $-11 \pm 17$ |
| Orthogonal | $-35 \pm 26$ | $-1 \pm 10$ |

**Notes.** The first column gives results for the $f_{\rm NL}$ parameters of the primordial local, equilateral, and orthogonal shapes, determined using the Binned estimator on the SMICA foreground-cleaned maps. The standard temperature mask is used, but the polarization mask has been enlarged by surrounding every hole by a region either 20 or 40 pixels in width ("Extra 20" or "Extra 40"). The second column gives the difference with the result using the 2018 common mask (see Table 5), where the error is the standard deviation of the differences in $f_{\rm NL}$ for 150 simulations when analysed with the two masks. Results have been determined using an independent single-shape analysis and are reported with subtraction of the lensing bias.

As explained in Sect. 3.2.2, it was decided to use $\ell_{\rm min} = 4$ for the polarization maps, since there was an issue with bias and increased variance when the two lowest multipoles for $E$ were included. However, this still allows us to use many more low-$\ell$ polarization modes than for the 2015 data, when it was necessary to adopt $\ell_{\rm min} = 40$ for polarization.

Turning to Fig. 18, for $\ell_{\rm max}$, we also notice good agreement with the analysis of the 2015 data, and slightly more stable results (e.g. for $T$-only equilateral). We see that the $T$-only results have stabilized by $\ell = 2000$ and the $E$-only results by $\ell = 1500$, so that there is no problem with the KSW and Modal estimators using these lower values for $\ell_{\rm max}$. As before we also confirm the "WMAP excess" for the local shape at $\ell \approx 500$ (Bennett et al. 2013), even more clearly than in PCNG15. However, this does not appear to be a major outlier when considering all values of $\ell_{\rm max}$ and all possible shapes.

### 6.6. Summary of validation tests

Throughout this section we have discussed a set of tests aimed at evaluating the robustness of our results. For convenience, we summarize here our main findings.

- We find good consistency for $f_{\rm NL}$ local, equilateral, and orthogonal measurements, between all component-separation methods and with all bispectrum estimators, separately considering $T$-only, $E$-only, and $T$+$E$ results. The agreement for $E$-only results has significantly improved with respect to the previous release.
- The possible exception is provided by the orthogonal $T$-only estimates. In this case, a comparison with simulations shows significant differences in specific cases, namely SMICA – Commander and, to a lesser extent, SMICA – SEVEM. However, these differences are coming from fluctuations in $f_{\rm NL}$ for Commander and SEVEM with respect to the two previous *Planck* releases, when all temperature maps were always in excellent agreement. Our main SMICA results are, on the other hand, completely stable. Moreover, the discrepancy becomes much less significant when adding polarization, it is limited to one specific shape, and it is in any case at a level that does not alter the interpretation of the results. Therefore, in the end, we do not consider this issue to be problematic.
- Nevertheless, in light of this test, we are led to a slight preference for SMICA and NILC as methods of choice for primordial NG analysis. Given that SMICA also gave a slightly better performance for NG analysis in the two previous releases, and that results extracted from SMICA maps have been quite stable over time, we maintain SMICA as our final choice, as already justified in detail in the papers analysing the 2013 and 2015 *Planck* data (PCNG13; PCNG15).
- We find very good consistency between fully reconstructed bispectra, for different component-separation methods, using both the Modal and the Binned approaches. Polarization-only bispectra again show a large improvement compared to the previous analysis of *Planck* data.
- The observed noise mismatch between the data and the FFP10 simulations does not seem to impact our results. Our cubic statistics cannot be biased by this mismatch, and tests on simulations show that the effect on $f_{\rm NL}$ errors is negligible.
- Results are stable to changes in sky coverage and different cuts in the multipole domain. Restricting the analysis to the multipole range probed by WMAP shows good agreement between WMAP and *Planck*.
- We find no sign of any residual Galactic thermal dust contamination in the *Planck* component-separated CMB maps.
- Contamination from the thermal SZ effect on the standard primordial and lensing $T$-only results is not at a significant level, compared to statistical errors.

Overall, the results display a high level of internal consistency, and are notably characterized by a large improvement in the quality of polarization-only bispectra with respect to the previous release. Whereas in 2015 we cautioned the reader to take polarization-based $f_{\rm NL}$ estimates as preliminary, we can now state that our $T$+$E$-based constraints are fully robust. This is one of the main conclusions of this paper.

## 7. Limits on the primordial trispectrum

We now present constraints on three shapes for the primordial four-point function or trispectrum, denoted $g_{\rm NL}^{\rm local}$, $g_{\rm NL}^{\dot\sigma^4}$, and $g_{\rm NL}^{(\partial\sigma)^4}$, and describe them below. The details of the analysis are mostly unchanged from the paper analysing the 2015 *Planck* data (see Sect. 9 of PCNG15). Nevertheless, we briefly review the four-point analysis here; for more details, see PCNG15, or Smith et al. (2015), which contains technical details of the pipeline.

First, we describe the three different signals of interest. The local-type trispectrum $g_{\rm NL}^{\rm local}$ arises if the initial adiabatic curvature $\zeta$ is given by the following non-Gaussian model:

$$\zeta(\boldsymbol{x}) = \zeta_{\rm G}(\boldsymbol{x}) + \frac{9}{25}\, g_{\rm NL}^{\rm local}\, \zeta_{\rm G}(\boldsymbol{x})^3 \tag{42}$$

where $\zeta_{\rm G}$ is a Gaussian field. The trispectrum in the local model is given by

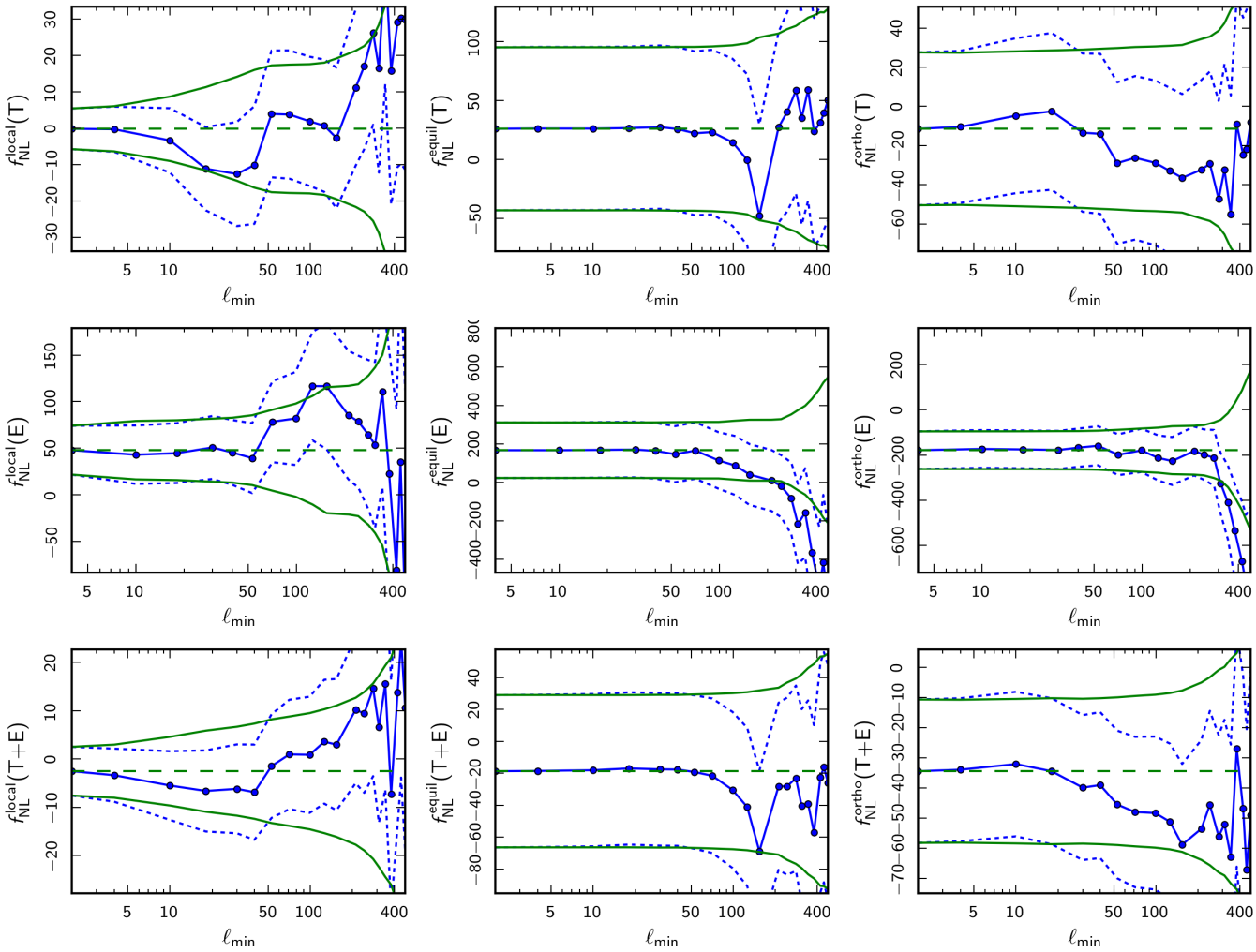


**Fig. 17.** Evolution of the $f_{\rm NL}$ parameters (solid blue line with data points) and their uncertainties (dotted blue lines) for the three primordial bispectrum templates as a function of the minimum multipole number $\ell_{\rm min}$ used in the analysis. *From left to right* the panels show local, equilateral, and orthogonal shape results, while the different rows *from top to bottom* show results for $T$ only, $E$ only, and full $T$+$E$ data. To indicate more clearly the evolution of the uncertainties, they are also plotted around the final value of $f_{\rm NL}$ (solid green lines without data points, around the horizontal dashed green line). The results here have been determined with the Binned bispectrum estimator for the SMICA map, assume all shapes to be independent, and the lensing bias has been subtracted.

$$\langle\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}\zeta_{\boldsymbol{k}_4}\rangle' = \frac{54}{25}\, g_{\rm NL}^{\rm local}\left[P_\zeta(k_1)P_\zeta(k_2)P_\zeta(k_3) + 3\,\text{perms.}\right]. \tag{43}$$

In this equation and throughout this section, a "primed" four-point function denotes the four-point function without its momentum-conserving delta function, that is,

$$\langle\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}\zeta_{\boldsymbol{k}_4}\rangle = \langle\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}\zeta_{\boldsymbol{k}_4}\rangle'(2\pi)^3\delta^{(3)}\Big(\sum \boldsymbol{k}_i\Big) + \text{disc.}, \tag{44}$$

where "+ disc." denotes disconnected contributions to the 4-point function. It can be shown that the local-type trispectrum in Eq. (43) is always negligibly small in single-field inflation (Senatore & Zaldarriaga 2012a); however, it can be large in multifield models of inflation in which a large bispectrum is forbidden by symmetry (Senatore & Zaldarriaga 2012b).

The next two shapes $g_{\rm NL}^{\dot\sigma^4}$, $g_{\rm NL}^{(\partial\sigma)^4}$ are generated by the operators $\dot\sigma^4$ and $(\partial_i\sigma)^2(\partial_j\sigma)^2$ in the effective field theory (EFT) of inflation (Bartolo et al. 2010b; Senatore & Zaldarriaga 2012b; Smith et al. 2015). For data analysis purposes, they can be defined by the following trispectra:

$$\begin{aligned}\langle\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}\zeta_{\boldsymbol{k}_4}\rangle' &= \frac{9216}{25} g_{\rm NL}^{\dot\sigma^4} A_\zeta^3 \int_{-\infty}^{0} {\rm d}\tau_E\, \tau_E^4 \left(\prod_{i=1}^{4}\frac{e^{k_i\tau_E}}{k_i}\right)\\ &= \frac{221\,184}{25} g_{\rm NL}^{\dot\sigma^4} A_\zeta^3 \frac{1}{k_1k_2k_3k_4K^5}\,;\end{aligned} \tag{45}$$

$$\begin{aligned}\langle\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}\zeta_{\boldsymbol{k}_4}\rangle' &= \frac{82\,944}{2575} g_{\rm NL}^{(\partial\sigma)^4} A_\zeta^3 \int_{-\infty}^{0} {\rm d}\tau_E \left[\prod_{i=1}^{4}\frac{(1-k_i\tau_E)e^{k_i\tau_E}}{k_i^3}\right]\\ &\quad\times\left[(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)(\boldsymbol{k}_3\cdot\boldsymbol{k}_4) + 2\,\text{perms.}\right]\\ &= \frac{165\,888}{2575} g_{\rm NL}^{(\partial\sigma)^4} A_\zeta^3\\ &\quad\times\left(\frac{2K^4 - 2K^2\sum k_i^2 + K\sum k_i^3 + 12k_1k_2k_3k_4}{k_1^3k_2^3k_3^3k_4^3K^5}\right)\\ &\quad\times\left[(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)(\boldsymbol{k}_3\cdot\boldsymbol{k}_4) + 2\,\text{perms.}\right].\end{aligned} \tag{46}$$

Here $K = \sum_i k_i$, and numerical prefactors have been chosen so that the trispectra have the same normalization as the local shape

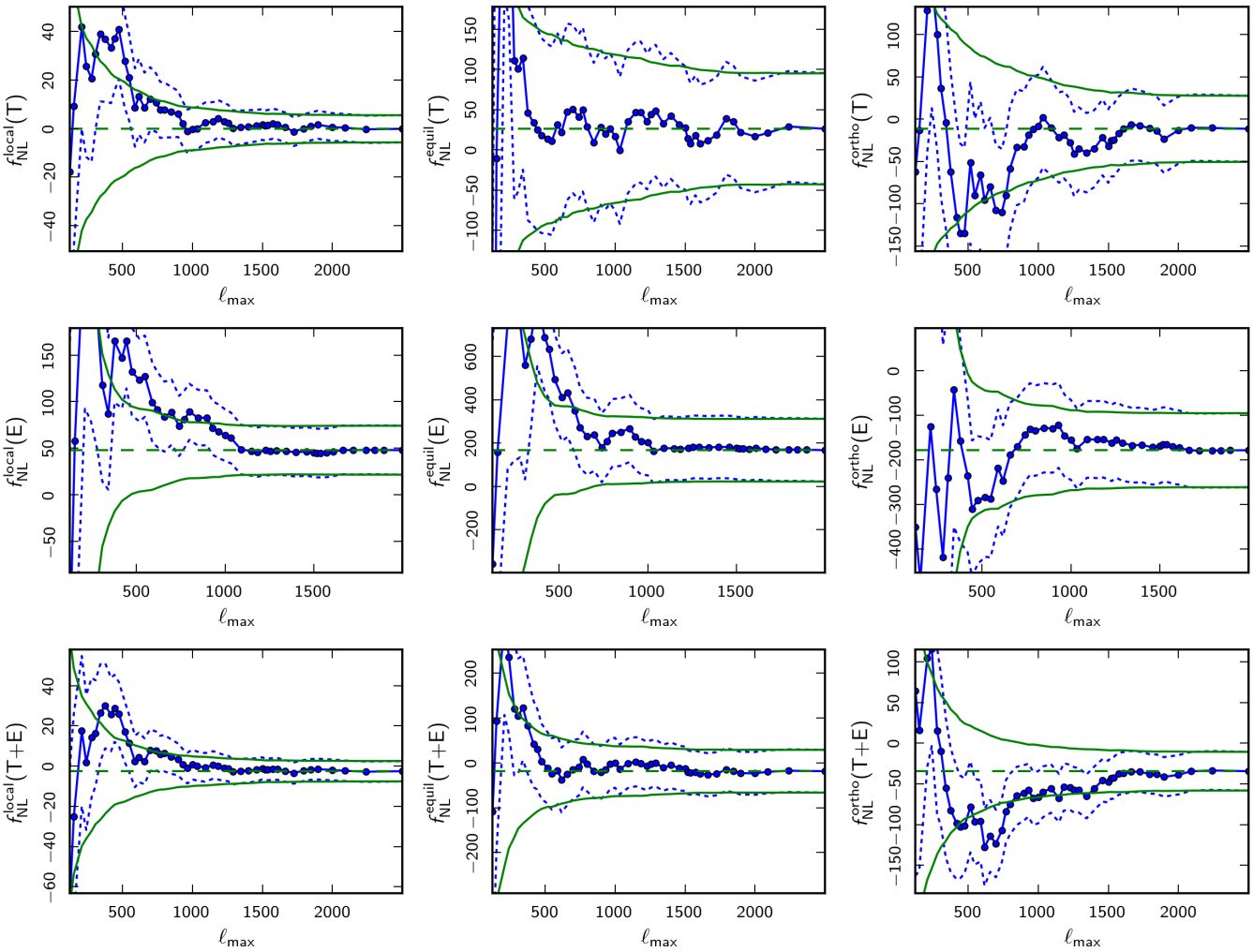


**Fig. 18.** Same as Fig. 17, but this time as a function of the maximum multipole number $\ell_{max}$ used in the analysis.

in Eq. (43) when restricted to tetrahedral wavenumber configurations with $|\boldsymbol{k}_i| = k$ and $(\boldsymbol{k}_i \cdot \boldsymbol{k}_j) = -k^2/3$. We mention in advance that there is another shape $g_{NL}^{\dot{\sigma}^2(\partial\sigma)^2}$ that arises at the same order in the EFT expansion, to be discussed at the end of this section.

In Eqs. (45) and (46), we have written each trispectrum in two algebraically equivalent ways, an integral representation and an "integrated" form. The integral representation arises naturally when evaluating the Feynman diagram for the EFT operator. It also turns out to be useful for data analysis, since the resulting "factorizable" representation for the trispectrum leads to an efficient algorithm for evaluating the CMB trispectrum estimator (Planck Collaboration XVII 2016; Smith et al. 2015).

For simplicity in Eqs. (45) and (46), we have assumed a scale-invariant power spectrum $P_\zeta(k) = A_\zeta/k^3$. For the *Planck* analysis, we slightly modify these trispectra to a power-law spectrum $P_\zeta(k) \propto k^{n_s-4}$, as described in Appendix C of Smith et al. (2015).

To estimate each $g_{NL}$ parameter from *Planck* data, we use the "pure-MC" trispectrum estimation pipeline from section IX.B of Smith et al. (2015). In this pipeline, the data are specified as a filtered harmonic space map $\tilde{d}_{\ell m}$, and its covariance is characterized via a set of 1000 filtered signal + noise simulations.

The "filter" is an experiment-specific linear operation whose input consists of one or more pixel-space maps, and whose output is a single harmonic-space map, $d_{\ell m}$. In the *Planck* trispectrum analysis, we define the filter as follows. First, we take the single-frequency pixel-space maps, and combine them to obtain a single component-separated pixel-space map, using one of the component-separation algorithms, `SMICA`, `SEVEM`, `NILC`, `Commander`, or `SMICA` no-SZ (a variant of the `SMICA` algorithm that guarantees zero response to Compton-$y$ sources, at the expense of slightly higher noise; see Sect. 6.3.2). Second, we subtract the best-fit monopole and dipole, inpaint masked point sources, and apodize the Galactic plane boundary. The details of these steps are unchanged from the 2015 *Planck* data analysis, and are described in Sect. 9.1 of PCNG15. Third, we take the spherical transform of the pixel-space map out to $l_{max} = 1600$, obtaining a harmonic-space map $d_{\ell m}$. Finally, we define the filtered map $\tilde{d}_{\ell m}$ by applying the multiplicative factor:

$$\tilde{d}_{\ell m} = \frac{d_{\ell m}}{b_\ell C_\ell + b_\ell^{-1} N_\ell}. \tag{47}$$

This sequence of steps defines a linear operation, whose input is a set of single-frequency pixel-space maps, and whose output is a filtered harmonic-space map $\tilde{d}_{\ell m}$. This filtering operation is used as a building block in the trispectrum pipeline described in Smith et al. (2015), and is the only part of the pipeline that is *Planck*-specific.

The results of the analysis, for all three trispectrum shapes and five different component-separation algorithms, are

**Table 24.** *Planck* 2018 constraints on the trispectrum parameters $g_{\rm NL}^{\rm local}$, $g_{\rm NL}^{\dot\sigma^4}$, and $g_{\rm NL}^{(\partial\sigma)^4}$ from different component-separated maps.

| | $g_{\rm NL}^{\rm local}$ | $g_{\rm NL}^{\dot\sigma^4}$ | $g_{\rm NL}^{(\partial\sigma)^4}$ |
|---|---|---|---|
| SMICA . . . . . | $(-5.8 \pm 6.5) \times 10^4$ | $(-0.8 \pm 1.9) \times 10^6$ | $(-3.9 \pm 3.9) \times 10^5$ |
| SMICA no-SZ . . . . . | $(-12.3 \pm 6.6) \times 10^4$ | $(-0.6 \pm 1.9) \times 10^6$ | $(-3.5 \pm 3.9) \times 10^5$ |
| SEVEM . . . . . | $(-5.5 \pm 6.5) \times 10^4$ | $(-0.8 \pm 1.9) \times 10^6$ | $(-3.2 \pm 3.9) \times 10^5$ |
| NILC . . . . . | $(-3.6 \pm 6.3) \times 10^4$ | $(-0.8 \pm 1.9) \times 10^6$ | $(-4.0 \pm 3.9) \times 10^5$ |
| Commander . . . . . | $(-8.1 \pm 6.5) \times 10^4$ | $(-0.8 \pm 1.9) \times 10^6$ | $(-3.5 \pm 3.9) \times 10^5$ |

presented in Table 24. We do not find evidence for a nonzero primordial trispectrum.

Each entry in Table 24 is a constraint on a single $g_{\rm NL}$-parameter with the others held fixed. We next consider joint constraints involving multiple $g_{\rm NL}$-parameters. In this case, we need to know the covariance matrix between $g_{\rm NL}$-parameters. We find that

$$\mathrm{Corr}(g_{\rm NL}^{\dot\sigma^4}, g_{\rm NL}^{(\partial\sigma)^4}) = 0.61, \tag{48}$$

and the correlation between $g_{\rm NL}^{\rm local}$ and the other two $g_{\rm NL}$-parameters is negligible.

Multifield models of inflation generally predict a linear combination of the quartic operators $\dot\sigma^4$ and $(\partial_i\sigma)^2(\partial_j\sigma)^2$. In addition, there is a third operator $\dot\sigma^2(\partial_i\sigma)^2$ that arises at the same order in the EFT expansion. For completeness, its trispectrum is given by

$$\begin{aligned}\langle\zeta_{\boldsymbol{k}_1}\zeta_{\boldsymbol{k}_2}\zeta_{\boldsymbol{k}_3}\zeta_{\boldsymbol{k}_4}\rangle' &= -\frac{13\,824}{325} g_{\rm NL}^{\dot\sigma^2(\partial\sigma)^2} A_\zeta^3 \int_{-\infty}^{0} \mathrm{d}\tau_E\, \tau_E^2 \\ &\quad\times \left[\frac{(1-k_3\tau_E)(1-k_4\tau_E)}{k_1k_2k_3^3k_4^3}(\boldsymbol{k}_3\cdot\boldsymbol{k}_4)\, e^{\sum k_i\tau_E} + 5\,\text{perms.}\right] \\ &= -\frac{27\,648}{325} g_{\rm NL}^{\dot\sigma^2(\partial\sigma)^2} A_\zeta^3 \\ &\quad\times \left[\frac{K^2+3(k_3+k_4)K+12k_3k_4}{k_1k_2k_3^3k_4^3K^5}(\boldsymbol{k}_3\cdot\boldsymbol{k}_4) + 5\,\text{perms.}\right].\end{aligned} \tag{49}$$

However, a Fisher matrix analysis shows that this trispectrum is nearly 100% correlated with the $\dot\sigma^4$ and $(\partial_i\sigma)^2(\partial_j\sigma)^2$ trispectra. If the parameter $g_{\rm NL}^{\dot\sigma^2(\partial\sigma)^2}$ is non-zero, then we can absorb it into the "effective" values of the parameters $g_{\rm NL}^{\dot\sigma^4}$ and $g_{\rm NL}^{(\partial\sigma)^4}$ as

$$\begin{aligned}(g_{\rm NL}^{\dot\sigma^4})_{\rm eff} &= 0.59\, g_{\rm NL}^{\dot\sigma^2(\partial\sigma)^2}, \\ (g_{\rm NL}^{(\partial\sigma)^4})_{\rm eff} &= 0.091\, g_{\rm NL}^{\dot\sigma^2(\partial\sigma)^2}.\end{aligned} \tag{50}$$

Therefore, to study joint constraints involving multiple $g_{\rm NL}$ parameters, it suffices to consider the two parameters, $g_{\rm NL}^{\dot\sigma^4}$ and $g_{\rm NL}^{(\partial\sigma)^4}$, with correlation coefficient given in Eq. (48).

We define the two-component parameter vector

$$g_i = \begin{pmatrix} g_{\rm NL}^{\dot\sigma^4} \\ g_{\rm NL}^{(\partial\sigma)^4} \end{pmatrix}. \tag{51}$$

and let $\hat g_i$ denote the two-component vector of single-$g_{\rm NL}$ estimates from the SMICA maps (Table 24):

$$\hat g_i = \begin{pmatrix} -8.0 \\ -3.9 \end{pmatrix} \times 10^{-5}. \tag{52}$$

We also define a two-by-two Fisher matrix $F_{ij}$, whose diagonal is given by $F_{ii} = 1/\sigma_i^2$, where $\sigma_i$ is the single-$g_{\rm NL}$ statistical error in Table 24, and whose off-diagonal is $F_{12} = rF_{11}^{1/2}F_{22}^{1/2}$, where $r$ is the correlation in Eq. (48). This gives:

$$F_{ij} = \begin{pmatrix} 2.8 & 8.3 \\ 8.3 & 66.5 \end{pmatrix} \times 10^{-13}. \tag{53}$$

Now, given a set of "theory" $g_{\rm NL}$ values, represented by a two-vector $g_i$, we compare to the *Planck* data by computing the following quantity for the trispectrum:

$$\chi^2(g_i) = [F_{ii}\hat g_i - (Fg)_i\,]\, F_{ij}^{-1}\, [F_{jj}\hat g_j - (Fg)_j]. \tag{54}$$

In a model-building context where the $g_{\rm NL}$ quantities $g_i$ depend on model parameters, confidence regions on model parameters can be obtained by appropriately thresholding $\chi^2$. We give some examples in Sect. 8.

## 8. Implications for early-Universe physics

We now want to convert constraints on primordial NG into constraints on parameters of various models of inflation. This allows us to highlight the constraining power of NG measurements, as an additional complementary observable beyond the CMB power spectra. In particular NG constraints can severely limit the parameter space of models that are alternatives to the standard single-field models of slow-roll inflation, since they typically feature a higher level of NG.

Unless stated otherwise, we follow the same procedures adopted in PCNG13 and PCNG15. A posterior of the model parameters is built based on the following steps: we start from the assumption that the sampling distribution is Gaussian (which is supported by Gaussian simulations); the likelihood is approximated by the sampling distribution, but centred on the NG estimate (see Elsner & Wandelt 2009); we use uniform or Jeffreys' priors, over intervals of the model parameter space that are physically meaningful (or as otherwise stated); and in some cases where two or more parameters are involved, we marginalize the posterior to provide one-dimensional limits on the parameter under consideration.

### 8.1. General single-field models of inflation

*DBI models.* Dirac-Born-Infeld (DBI) models of inflation (Silverstein & Tong 2004; Alishahiha et al. 2004) arise from high-energy string-theory constructions and generate a non-linearity parameter $f_{\rm NL}^{\rm DBI} = -(35/108)(c_{\rm s}^{-2} - 1)$, where $c_{\rm s}$ is the sound speed of the inflaton perturbations (Silverstein & Tong 2004; Alishahiha et al. 2004; Chen et al. 2007b). The enhancement of the NG amplitude due to a possible sound speed $c_{\rm s} < 1$ arises from a non-standard kinetic term of the inflaton field. Notice that we have constrained the exact theoretical (non-separable) shape (see Eq. (7) of PCNG13), even though it is very

similar to the equilateral type. Our constraint $f_{\rm NL}^{\rm DBI} = 46 \pm 58$ from temperature data ($f_{\rm NL}^{\rm DBI} = 14 \pm 38$ from temperature and polarization) at 68% CL (with lensing and point sources subtracted, see Table 11) implies

$$c_{\rm s}^{\rm DBI} \geq 0.079 \qquad (95\%,\ T\,\text{only}), \tag{55}$$

and

$$c_{\rm s}^{\rm DBI} \geq 0.086 \qquad (95\%,\ T+E)\,. \tag{56}$$

*Implications for the effective field theory of inflation.* Now we can update CMB limits on the speed of sound $c_{\rm s}$ at which inflaton fluctuations propagated in the very early Universe. A very general constraint on this inflationary parameter can be obtained by employing the EFT approach to inflation (Cheung et al. 2008; Weinberg 2008, and see Sect. 7). This approach allows us to obtain predictions for the parameter space of primordial NG through a general characterization of the inflaton field interactions. The Lagrangian of the system is expanded into the dominant operators that respect some underlying symmetries. The procedure thus determines a unifying scheme for classes of models featuring deviations from single-field slow-roll inflation. Typically the equilateral and orthogonal templates represent an accurate basis to describe the full parameter space of EFT single-field models of inflation, and therefore we used the constraints on $f_{\rm NL}^{\rm equil}$ and $f_{\rm NL}^{\rm ortho}$.

As a concrete example, let us consider the Lagrangian of general single-field models of inflation (of the form $P(X,\varphi)$ models, where $X = g^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi$) written with the EFT approach:

$$S = \int {\rm d}^4x\,\sqrt{-g}\left[-\frac{M_{\rm Pl}^2\dot{H}}{c_{\rm s}^2}\left(\dot{\pi}^2 - c_{\rm s}^2\frac{(\partial_i\pi)^2}{a^2}\right)\right.$$
$$\left. -M_{\rm Pl}^2\dot{H}(1-c_{\rm s}^{-2})\dot{\pi}\frac{(\partial_i\pi)^2}{a^2} + \left(M_{\rm Pl}^2\dot{H}(1-c_{\rm s}^{-2}) - \frac{4}{3}M_3^4\right)\dot{\pi}^3\right]. \tag{57}$$

The scalar perturbation $\pi$ generates the curvature perturbation $\zeta = -H\pi$. In this case there are two relevant inflaton interactions, $\dot{\pi}(\partial_i\pi)^2$ and $(\dot{\pi})^3$, producing two specific bispectra with amplitudes $f_{\rm NL}^{\rm EFT1} = -(85/324)(c_{\rm s}^{-2}-1)$ and $f_{\rm NL}^{\rm EFT2} = -(10/243)(c_{\rm s}^{-2}-1)\left[\tilde{c}_3 + (3/2)c_{\rm s}^2\right]$, respectively.

Here $M_3$ is the amplitude of the operator $\dot{\pi}^3$ (see Senatore et al. 2010; Chen et al. 2007b; Chen 2010b), with the dimensionless parameter $\tilde{c}_3(c_{\rm s}^{-2}-1) = 2M_3^4c_{\rm s}^2/(\dot{H}M_{\rm Pl}^2)$ (Senatore et al. 2010). The two EFT shapes can be projected onto the equilateral and orthogonal shapes, with the mean values of the estimators for $f_{\rm NL}^{\rm equil}$ and $f_{\rm NL}^{\rm ortho}$ expressed in terms of $c_{\rm s}$ and $\tilde{c}_3$ as

$$f_{\rm NL}^{\rm equil} = \frac{1-c_{\rm s}^2}{c_{\rm s}^2}\left[-0.275 - 0.0780c_{\rm s}^2 - (2/3)\times 0.0780\tilde{c}_3\right],$$
$$f_{\rm NL}^{\rm ortho} = \frac{1-c_{\rm s}^2}{c_{\rm s}^2}\left[0.0159 + 0.0167c_{\rm s}^2 + (2/3)\times 0.0167\tilde{c}_3\right], \tag{58}$$

where the coefficients come from the Fisher matrix between the theoretical bispectra predicted by the two operators $\dot{\pi}(\nabla\pi)^2$ and $\dot{\pi}^3$ and the equilateral and orthogonal templates. Notice that DBI models reduce to the condition $\tilde{c}_3 = 3(1-c_{\rm s}^2)/2$, while the non-interacting (vanishing NG) case corresponds to $c_{\rm s} = 1$ and $M_3 = 0$ (or $\tilde{c}_3(c_{\rm s}^{-2}-1) = 0$).

We then proceed as in the two previous analyses (PCNG13; PCNG15). We employ a $\chi^2$ statistic computed as $\chi^2(\tilde{c}_3, c_{\rm s}) = \boldsymbol{v}^{\rm T}(\tilde{c}_3, c_{\rm s})\mathsf{C}^{-1}\boldsymbol{v}(\tilde{c}_3, c_{\rm s})$, with $v^i(\tilde{c}_3, c_{\rm s}) = f^i(\tilde{c}_3, c_{\rm s}) - f_P^i$ ($i = \{\text{equilateral, orthogonal}\}$), where $f_P^i$ are the joint estimates of equilateral and orthogonal $f_{\rm NL}$ values (see Table 6), while $f^i(\tilde{c}_3, c_{\rm s})$ are provided by Eq. (58) and $\mathsf{C}$ is the covariance matrix of the joint estimators. Figure 19 shows the 68%, 95%, and 99.7% confidence regions for $f_{\rm NL}^{\rm equil}$ and $f_{\rm NL}^{\rm ortho}$, as derived from from the $T+E$ constraints, with the requirement $\chi^2 \leq 2.28$, 5.99, and 11.62, respectively (corresponding to a $\chi^2$ variable with two degrees of freedom). In Fig. 20 we show the corresponding confidence regions in the $(\tilde{c}_3, c_{\rm s})$ parameter space. Marginalizing over $\tilde{c}_3$ we find

$$c_{\rm s} \geq 0.021 \qquad (95\%,\ T\,\text{only})\,, \tag{59}$$

and

$$c_{\rm s} \geq 0.021 \qquad (95\%,\ T+E). \tag{60}$$

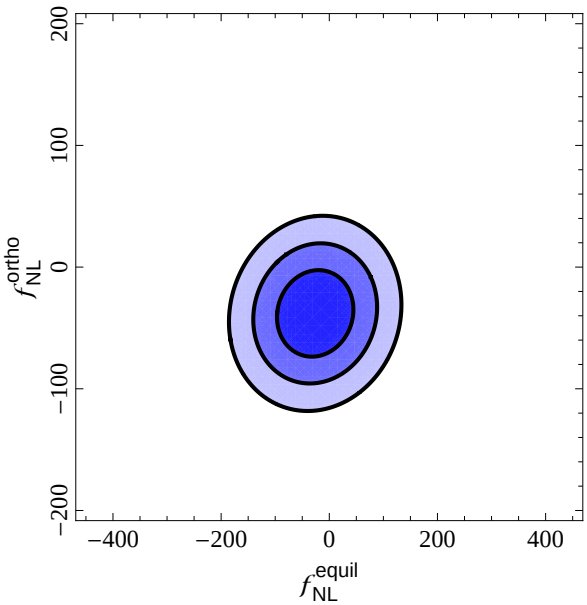


**Fig. 19.** 68%, 95%, and 99.7% confidence regions in the parameter space $(f_{\rm NL}^{\rm equil}, f_{\rm NL}^{\rm ortho})$, defined by thresholding $\chi^2$, as described in the text.

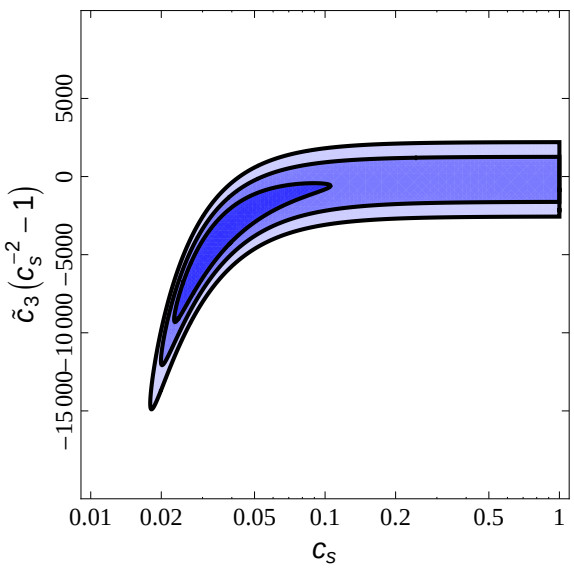


**Fig. 20.** 68%, 95%, and 99.7% confidence regions in the single-field inflation parameter space $(c_{\rm s}, \tilde{c}_3)$, obtained from Fig. 19 via the change of variables in Eq. (58).

There is a slight improvement in comparison with the constraints obtained in PCNG15 coming from the $T + E$ data. Notice that, contrary to the situation in 2015, the final $c_s$ constraint here is essentially determined using temperature data, with no significant improvement when adding the polarization data set. There are two reasons for this. Firstly, the errors on $f_{\rm NL}^{\rm equil}$ and $f_{\rm NL}^{\rm ortho}$ display less improvement in going from $T$ to $T + E$, in comparison to the previous release (see Sect. 9 for further comments on this point). Secondly, the shift of the central value of $f_{\rm NL}^{\rm ortho}$ (and also of $f_{\rm NL}^{\rm equil}$) when passing from $T$ to $T + E$ is somewhat larger than in 2015. Such a shift is now towards more negative values and has the net effect of counterbalancing the improvement of the error bars on $f_{\rm NL}^{\rm equil}$ and $f_{\rm NL}^{\rm ortho}$ when adding polarization, thus leaving the constraints on $c_s$ unchanged with respect to the one from temperature data only.

### 8.2. Multi-field models

Constraints on primordial NG of the local type lead to strong implications for models of inflation where scalar fields (different from the inflaton) are dynamically important for the generation of the primordial curvature perturbation. In the following we test two scenarios for curvaton models.

*Basic curvaton models.* The simplest adiabatic curvaton models predict primordial NG of the local shape with a non-linearity parameter (Bartolo et al. 2004c,d)

$$f_{\rm NL}^{\rm local} = \frac{5}{4r_{\rm D}} - \frac{5r_{\rm D}}{6} - \frac{5}{3}\,, \tag{61}$$

in the case where the curvaton field has a quadratic potential (Lyth & Wands 2002; Lyth et al. 2003; Lyth & Rodriguez 2005; Malik & Lyth 2006; Sasaki et al. 2006). Here the parameter $r_{\rm D} = [3\rho_{\rm curvaton}/(3\rho_{\rm curvaton} + 4\rho_{\rm radiation})]_{\rm D}$ is the "curvaton decay fraction" at the time of the curvaton decay in the sudden decay approximation. We assume a uniform prior, $0 < r_{\rm D} < 1$. It is worth recalling that these models predict a lower bound for the level of NG, of the order of unity (corresponding precisely to $f_{\rm NL}^{\rm local} = -5/4$), which is considered as a typical threshold to distinguish between standard single-field and multi-field scenarios. Our constraint $f_{\rm NL}^{\rm local} = -0.5 \pm 5.6$ at 68% CL (see Table 6) implies

$$r_{\rm D} \geq 0.19 \qquad (95\%,\ T\ \text{only})\,. \tag{62}$$

The constraint $f_{\rm NL}^{\rm local} = -0.9 \pm 5.1$ at 68% CL obtained from temperature and polarization data yields the constraint

$$r_{\rm D} \geq 0.21 \qquad (95\%,\ T + E)\,. \tag{63}$$

These limits indicate that in such scenarios the curvaton field has a non-negligible energy density when it decays. Meaningful improvements are achieved with respect to previous bounds in PCNG15, namely an almost 20% improvement from $T$-only data and a 10% improvement when including $E$-mode polarization.

*Decay into curvaton particles.* We reach similar improvements on the parameters of the second scenario of the curvaton models we consider. In this case one accounts for the possibility that the inflaton field can decay into curvaton particles (Linde & Mukhanov 2006), a possibility that is neglected in the above expression (Eq. (61)) for $f_{\rm NL}^{\rm local}$. It might be the case that the classical curvaton field survives and begins to dominate. In this case also the curvaton particles produced during reheating are expected to survive and dominate over other species at the epoch of their decay (since thay have the same equation of state as the classical curvaton field). Primordial adiabatic perturbations are generated given that the classical curvaton field and the curvaton particles decay at the same time (see Linde & Mukhanov 2006 for a detailed discussion). To interpret $f_{\rm NL}$ in this scenario we employ the general formula for $f_{\rm NL}^{\rm local}$ derived in Sasaki et al. (2006), which takes into account the possibility that the inflaton field decays into curvaton particles:

$$f_{\rm NL}^{\rm local} = (1 + \Delta_{\rm s}^2)\frac{5}{4r_{\rm D}} - \frac{5r_{\rm D}}{6} - \frac{5}{3}\cdot \tag{64}$$

Here the parameter $\Delta_{\rm s}^2$ is the ratio of the energy density of curvaton particles to the energy density of the classical curvaton field (Linde & Mukhanov 2006; Sasaki et al. 2006), while now $\rho_{\rm curv}$ in the expression for $r_{\rm D}$ must be replaced by the sum of the densities of the curvaton particles and curvaton field. As in PCNG15 we use uniform priors $0 < r_{\rm D} < 1$ and $0 < \Delta_{\rm s}^2 < 10^2$. Our limits on $f_{\rm NL}^{\rm local}$ constrain

$$\Delta_{\rm s}^2 \leq 6.9 \qquad (95\%,\ T\ \text{only}), \tag{65}$$

and

$$\Delta_{\rm s}^2 \leq 6.2 \qquad (95\%,\ T + E), \tag{66}$$

which does not exclude a contribution of curvaton particles comparable to the one from the classical curvaton field.

### 8.3. Non-standard inflation models

*Directional-dependent NG.* Table 13 shows the constraints on directionally-dependent bispectra (Eq. (20)). This kind of NG is predicted by several different inflationary models. For example, it is a robust and (almost unavoidable) outcome of models of inflation where scale-invariant gauge fields are present during inflation. As summarized in Sect. 2 they are also produced from partially massless higher-spin particles (Franciolini et al. 2018) or from models of solid inflation (Endlich et al. 2013, 2014; Shiraishi et al. 2013a), as well as in models of inflation that break both rotational and parity invariance (Bartolo et al. 2015). To compare with the constraints obtained in the analysis of the 2015 *Planck* data (PCNG15), we reconsider the specific model where the inflaton is coupled to the kinetic term $F^2$ of a gauge field via a term $\mathcal{L} = -I^2(\phi)F^2/4$, where $I(\phi)$ is a function that depends on the inflaton field, having an appropriate time evolution during inflation (see, e.g. Ratra 1992). Specifically in these models the production of super-horizon vector field perturbations switches on the $L = 0$ and $L = 2$ modes in the bispectrum, with non-linearity parameters $f_{\rm NL}^{L} = X_L(|g_*|/0.1)\,(N_{k_3}/60)$, with $X_{L=0} = (80/3)$ and $X_{L=2} = -(10/6)$, respectively (Barnaby et al. 2012a; Bartolo et al. 2013a; Shiraishi et al. 2013a). In these expressions $g_*$ is a parameter that measures the amplitude of a quadrupolar anisotropy in the power spectrum (see, e.g. Ackerman et al. 2007), while $N$ is the number of e-folds (from the end of inflation) at which the relevant scales cross outside the Hubble scale. It is therefore interesting to set some limits on the parameter $g_*$ exploiting the constraints from primordial NG of this type. Using the SMICA constraints from $T$ (or $T$+$E$) in Table 13, marginalizing over a uniform prior $50 \leq N \leq 70$, and assuming uniform priors on $-1 \leq g_* \leq 1$, we obtain the 95% bounds $-0.041 < g_* < 0.041$ ($-0.036 < g_* < 0.036$), and $-0.35 < g_* < 0.35$ ($-0.30 < g_* < 0.30$), from the $L = 0$ and $L = 2$ modes, respectively (considering $g_*$ to be scale independent).

*Tensor NG and pseudoscalars.* Using the SMICA $T$+$E$ result $f_{\rm NL}^{\rm tens} = (8 \pm 11) \times 10^2$ (68% CL), we here place constraints on two specific inflation models, including either a U(1)-axion coupling or an SU(2)-axion one. The former U(1) model results in $f_{\rm NL}^{\rm tens} \approx 6.4 \times 10^{11} \mathcal{P}^3 \epsilon^3 e^{6\pi\xi}/\xi^9$, where $\mathcal{P}$ is the vacuum-mode curvature power spectrum, $\epsilon$ is a slow-roll parameter of the inflaton field, and $\xi$ expresses the strength of the U(1)-axion coupling (Cook & Sorbo 2013; Shiraishi et al. 2013b). We then fix $\epsilon$ to be 0.01 and marginalize $\mathcal{P}$ with the prior, $1.5 \times 10^{-9} < \mathcal{P} < 3.0 \times 10^{-9}$; assuming a prior, $0.1 < \xi < 7.0$, the upper bound on $\xi$ is derived as $\xi < 3.3$ (95% CL). In the latter SU(2) model, under one specific condition, the tensor non-linear parameter is related to the tensor-to-scalar ratio $r$ and the energy density fraction of the gauge field $\Omega_{\rm A}$ as $f_{\rm NL}^{\rm tens} \approx 2.5 r^2/\Omega_{\rm A}$ (Agrawal et al. 2018). The lower bound on $\Omega_{\rm A}$ is estimated under a prior, $0 < \Omega_{\rm A} < 1$. We find $\Omega_{\rm A} > 2.3 \times 10^{-7}$ and $2.7 \times 10^{-9}$ (95% CL) for $r = 10^{-2}$ and $10^{-3}$, respectively.

*Warm inflation.* As in previous analyses (PCNG13; PCNG15), we adopt the expression $f_{\rm NL}^{\rm warm} = -15 \ln\left(1 + r_{\rm d}/14\right) - 5/2$ (Moss & Xiong 2007). This is valid when dissipative effects are strong, this is for values $r_{\rm d} \gtrsim 2.5$ of the dissipation parameter $r_{\rm d} = \Gamma/(3H)$ (measuring the effectiveness of the energy transfer from the inflaton field to radiation)[12]. Assuming a constant prior $0 \leq \log_{10} r_{\rm d} \leq 4$, the SMICA constraints $f_{\rm NL}^{\rm warmS} = -39 \pm 44$ (at 68% CL) from $T$ and $f_{\rm NL}^{\rm warmS} = -48 \pm 27$ from $T$+$E$ (see Table 12), yield $\log_{10} r_{\rm d} \leq 3.6$ and $\log_{10} r_{\rm d} \leq 3.5$, respectively, at 95 % CL. The results show that strong-dissipative effects in warm inflation models remain allowed. This is a regime where gravitino overproduction problems can be evaded (for a discussion see Hall & Peiris 2008).

### 8.4. Alternatives to inflation

As an example we update the constraints on some ekpyrotic/cyclic models (e.g. see Lehners 2010 for a review). Typically local NG is produced through a conversion of "intrinsic" non-Gaussianity in the entropy fluctuations into the curvature perturbation. This conversion can proceed in different ways. The "ekpyrotic conversion" models, for which the conversion acts during the ekpyrotic phase, have already been ruled out (Koyama et al. 2007, PCNG13). On the other hand, in the "kinetic conversion" models the conversion takes place after the ekpyrotic phase and a local bispectrum is generated with an amplitude $f_{\rm NL}^{\rm local} = (3/2)\,\kappa_3 \sqrt{\epsilon} \pm 5$[13]. The sign depends on the details of the conversion process (Lehners & Steinhardt 2008, 2013; Lehners 2010), and typical values of the parameter $\epsilon$ are $\epsilon \approx 50$ or greater. Assuming $\epsilon \approx 100$ and using a uniform prior on $-5 < \kappa_3 < 5$ the constraints on $f_{\rm NL}^{\rm local}$ from $T$ only (see Table 6), implies $-1.1 < \kappa_3 < 0.36$ and $-0.43 < \kappa_3 < 1.0$ at 95% CL, for the plus and minus sign in $f_{\rm NL}^{\rm local}$, respectively. The $T$+$E$ constraints on $f_{\rm NL}^{\rm local}$ (Table 6) yield $-1.05 < \kappa_3 < 0.27$ and $-0.38 < \kappa_3 < 0.94$ at 95% CL, for the plus and minus sign, respectively. If we take $\epsilon \approx 50$ as an example, we obtain the following limits: $-1.6 < \kappa_3 < 0.51$ and $-0.62 < \kappa_3 < 1.5$ at 95% CL from $T$ only; and $-1.5 < \kappa_3 < 0.39$ and $-0.54 < \kappa_3 < 1.3$ at 95% CL from $T$+$E$ constraints.

### 8.5. Inflationary interpretation of CMB trispectrum results

We briefly analyse inflationary implications of the *Planck* trispectrum constraints, using the SMICA limits on $g_{NL}$ parameters from Table 24.

First, we consider single-field inflationary models, using the effective action for the Goldstone boson $\pi$ (see, e.g., Smith et al. 2015):

$$\begin{aligned} S_\pi = \int {\rm d}^4x\,\sqrt{-g}\Big\{ & -M_{\rm Pl}^2 \dot{H}\left(\partial_\mu\pi\right)^2 \\ & + 2M_2^4\left[\dot{\pi}^2 + \dot{\pi}^3 - \dot{\pi}\frac{(\partial_i\pi)^2}{a^2} + (\partial_\mu\pi)^2(\partial_\nu\pi)^2\right] \\ & - \frac{M_3^4}{3!}\left[8\,\dot{\pi}^3 + 12\dot{\pi}^2(\partial_\mu\pi)^2 + \cdots\right] \\ & + \frac{M_4^4}{4!}\left[16\,\dot{\pi}^4 + 32\dot{\pi}^3(\partial_\mu\pi)^2 + \cdots\right] + \cdots\Big\}, \end{aligned} \tag{67}$$

The mass scale $M_4$ is related to our previously-defined $g_{\rm NL}$ parameters by:

$$g_{\rm NL}^{\dot{\sigma}^4} = \frac{25}{288}\frac{M_4^4}{H^4} A_\zeta\, c_{\rm s}^3\,. \tag{68}$$

Therefore, using the $g_{\rm NL}^{\dot{\sigma}^4}$ limit from Table 24, we get the following constraint on single-field models:

$$-12.8 \times 10^{14} < \frac{M_4^4}{H^4 c_{\rm s}^3} < 8.2 \times 10^{14} \quad (95\%\ {\rm CL}). \tag{69}$$

Next consider the case of multifield inflation. Here, we consider an action of the more general form:

$$\begin{aligned} S_\sigma = \int {\rm d}^4x\,\sqrt{-g}\,\Big[ & \frac{1}{2}(\partial_\mu\sigma)^2 + \frac{1}{\Lambda_1^4}\dot{\sigma}^4 \\ & + \frac{1}{\Lambda_2^4}\dot{\sigma}^2(\partial_i\sigma)^2 + \frac{1}{\Lambda_3^4}(\partial_i\sigma)^2(\partial_j\sigma)^2\Big], \end{aligned} \tag{70}$$

where $\sigma$ is a light field that acquires quantum fluctuations with power spectrum $P_\sigma(k) = H^2/(2k^3)$. We assume that $\sigma$ converts to adiabatic curvature $\zeta$, i.e. $\zeta = (2A_\zeta)^{1/2} H^{-1}\sigma$. The model parameters $\Lambda_i$ are related to our previously-defined $g_{\rm NL}$ parameters by:

$$\begin{aligned} g_{\rm NL}^{\dot{\sigma}^4} A_\zeta &= \frac{25}{768}\frac{H^4}{\Lambda_1^4}, \\ g_{\rm NL}^{\dot{\sigma}^2(\partial\sigma)^2} A_\zeta &= -\frac{325}{6912}\frac{H^4}{\Lambda_2^4}, \\ g_{\rm NL}^{(\partial\sigma)^4} A_\zeta &= \frac{2575}{20\,736}\frac{H^4}{\Lambda_3^4}, \end{aligned} \tag{71}$$

and can be constrained by thresholding the $\chi^2$-statistic defined in Eq. (54). For example, to constrain the parameter $\Lambda$ in the Lorentz invariant model

$$S = \int {\rm d}^4x\,\sqrt{-g}\left[\frac{1}{2}(\partial_\mu\sigma)^2 + \frac{1}{\Lambda^4}(\partial_\mu\sigma)^2(\partial_\nu\sigma)^2\right], \tag{72}$$

[12] The intermediate and weak dissipative regimes ($r_{\rm d} \leq 1$) predict an NG amplitude with a strong dependence on the microscopic parameters ($T/H$ and $r_{\rm d}$), giving rise to a different additional bispectrum shape (see Bastero-Gil et al. 2014).

[13] There might also be the case where the intrinsic NG is vanishing and primordial NG is generated only by non-linearities in the conversion process, reaching an amplitude $f_{\rm NL}^{\rm local} \approx \pm 5$ (Qiu et al. 2013; Li 2013; Fertig et al. 2014).

we set $\Lambda_1^4 = -2\Lambda_2^4 = \Lambda_3^4 = \Lambda^4$, obtain $g_{\rm NL}$-values using Eq. (71), and use Eq. (54) to obtain $\chi^2$ as a function of $\Lambda$. We then threshold at $\Delta\chi^2 = 4$ (as appropriate for one degree of freedom), to obtain the following constraint:

$$-0.33 < \frac{H^4}{\Lambda^4} < 0.11 \qquad (95\%). \tag{73}$$

*DBI trispectrum.* We use the trispectrum constraints on the shape $\dot{\sigma}^4$ in Table 24 to determine a lower bound on the sound speed of the inflaton field in DBI models. In fact in these models the dominant contribution in the small-sound-speed limit (Chen et al. 2009; Arroja et al. 2009) to the contact interaction trispectrum (Huang & Shiu 2006) produces such a shape, with an amplitude $g_{\rm NL}^{\dot{\sigma}^4} = -25/(768\, c_{\rm s}^4)$. We employ the same procedure described at the beginning of this section and, assuming a uniform prior in the range $0 \le c_{\rm s} \le 1/5$, we derive the following constraint on $c_{\rm s}$:

$$c_{\rm s}^{\rm DBI} \ge 0.015 \qquad (95\%)\,. \tag{74}$$

This constraint is independent from and consistent with the bounds of Eqs. (55) and (56) obtained from the bispectrum measurements. Notice that in the trispectrum case we are ignoring the scalar exchange contribution, which turns out to be of the same order in $c_{\rm s}$.

*Curvaton trispectrum.* A generic prediction of the simplest adiabatic curvaton scenario is also a local-type trispectrum with an amplitude $g_{\rm NL}^{\rm local}$ given by (Sasaki et al. 2006)

$$g_{\rm NL}^{\rm local} = \frac{25}{54}\left(-\frac{9}{r_{\rm D}} + \frac{1}{2} + 10 r_{\rm D} + 3 r_{\rm D}^2\right). \tag{75}$$

Following the procedure described at the beginning of this section, we use the observational constraint obtained in Sect. 7 (see Table 24), and the same uniform prior $(0 < r_{\rm D} < 1)$ as in Sect. 8.2, to obtain a lower bound on the curvaton decay fraction

$$r_{\rm D} \ge 0.05 \qquad (95\%)\,. \tag{76}$$

This limit is consistent with the previous ones derived using the bispectrum measurements and it is about a factor of 4 weaker.

## 9. Conclusions

In this paper, we have presented constraints on primordial NG, using the *Planck* full-mission CMB temperature and *E*-mode polarization maps. Compared to the *Planck* 2015 release we now include the low-$\ell$ ($4 \le \ell < 40$) polarization multipole range.

Our analysis produces the following final results (68% CL, statistical): $f_{\rm NL}^{\rm local} = -0.9 \pm 5.1$; $f_{\rm NL}^{\rm equil} = -26 \pm 47$; and $f_{\rm NL}^{\rm ortho} = -38 \pm 24$. These results are overall stable with respect to our constraints from the 2015 *Planck* data. They show no real improvement in errors, despite the additional polarization modes. This is due to a combination of two factors. Firstly, the local shape, which is most sensitive to low-$\ell$ modes and where one would naively expect an improvement, is actually less sensitive to polarization than the equilateral and orthogonal shapes. This means that in the end none of the three shapes are very sensitive to low-$\ell$ polarization modes. Secondly, the temperature and polarization simulations used to determine the errors have a more realistic but slightly higher noise level than in the previous release.

On the other hand, the quality of polarization data shows a clear improvement with respect to our previous analysis. This is confirmed by a large battery of tests on our data set, including comparisons between different estimator implementations (KSW, Binned, and two Modal estimators) and foreground-cleaning methods (`SMICA`, `SEVEM`, `NILC`, and `Commander`), studies of robustness under changes in sky coverage and multipole range, and an analysis of the impact of noise-related systematics. While in our previous release we cautioned the reader to take polarization bispectra and related constraints as preliminary, in light of these tests we now consider our results based on the combined temperature and polarization data set to be fully reliable. This also implies that polarization-only, *EEE* bispectra can now be used for independent tests, leading to primordial NG constraints at a sensitivity level comparable to that of WMAP from temperature bispectra, and yielding statistical agreement.

As in the previous analyses, we go beyond the local, equilateral, and orthogonal $f_{\rm NL}$ constraints by considering a large number of additional cases, such as scale-dependent feature and resonance bispectra, running $f_{\rm NL}$ models, isocurvature primordial NG, and parity-breaking models. We set tight constraints on all these scenarios, but do not detect any significant signals.

On the other hand, the non-primordial lensing bispectrum is now detected with an improved significance compared to 2015, excluding the null hypothesis at $3.5\sigma$. The amplitude of the signal is consistent with the expectation from the *Planck* best-fit cosmological parameters, further indicating the absence of significant foreground contamination or spurious systematic effects. We also explicitly checked for the presence of various non-primordial contaminants, like unclustered extragalactic point sources, CIB, Galactic thermal dust, and the thermal SZ effect, but apart from the first, none of these were detected. The small amount of remaining point-source signal in the cleaned maps has no impact on our other constraints because of its negligible correlations.

We update our trispectrum constraints, now finding $g_{\rm NL}^{\rm local} = (-5.8 \pm 6.5) \times 10^4$ (68% CL, statistical), while also constraining additional shapes, generated by different operators in an effective field-theory approach to inflation.

In addition to estimates of bispectrum and trispectrum amplitudes, we produce model-independent reconstructions and analyses of the *Planck* CMB bispectrum. Finally, we use our measurements to obtain constraints on early-Universe scenarios that can generate primordial NG. We consider, for example, general single-field models of inflation, curvaton models, models with axion fields producing parity-violating tensor bispectra, and inflationary scenarios generating directionally-dependent bispectra (such as those involving vector fields).

In our data analysis efforts, which started with the 2013 release, we achieved a number of crucial scientific goals. In particular we reached an unprecedented level of sensitivity in the determination of the bispectrum and trispectrum amplitude parameters ($f_{\rm NL}$, $g_{\rm NL}$) and significantly extended the standard local, equilateral, and orthogonal analysis, encompassing a large number of additional shapes motivated by a variety of inflationary models. Moreover, we produced the first polarization-based CMB bispectrum constraints and the first detection of the (non-primordial) bispectrum induced by correlations between CMB lensing and secondary anisotropies. Our stringent tests of many types of non-Gaussianity are fully consistent with expectations from the standard single-field slow-roll paradigm and provide strong constraints on alternative scenarios. Nevertheless, the current level of sensitivity does not allow us to rule out or confirm most alternative scenarios. It is natural at this stage to

ask ourselves what should be the $f_{\rm NL}$ sensitivity goal for future cosmological experiments. A number of studies has identified $f_{\rm NL} \sim 1$ as a target. Achieving such sensitivity for local-type NG would enable us to either confirm or rule out a large class of multi-field models. A similar target for equilateral, orthogonal, and scale-dependent shapes would allow us to distinguish standard slow-roll from more complex single-field scenarios, such as those characterized by higher-derivative kinetic terms or slow-roll-breaking features in the inflaton potential (see e.g., Alvarez et al. 2014; Finelli et al. 2018, and references therein). With this aim in mind, the challenge for future cosmological observations will be therefore that of reducing the $f_{\rm NL}$ errors from this paper by at least one order of magnitude.

*Acknowledgements.* The Planck Collaboration acknowledges the support of: ESA; CNES and CNRS/INSU-IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MINECO, JA, and RES (Spain); Tekes, AoF, and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); and ERC and PRACE (EU). A description of the Planck Collaboration and a list of its members, indicating which technical or scientific activities they have been involved in, can be found at http://www.cosmos.esa.int/web/planck. Some of the results in this paper have been derived using the HEALPix package. We also acknowledge the use of CAMB. Part of this work was undertaken on the STFC COSMOS@DiRAC HPC Facility at the University of Cambridge, funded by UK BIS NEI grants. We gratefully acknowledge the IN2P3 Computer Center (http://cc.in2p3.fr) for providing a significant amount of the computing resources and services needed for the analysis. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Some computations were performed on the GPC and Niagara cluster at the SciNet HPC Consortium; SciNet is funded by the Canada Foundation for Innovation under the auspices of Compute Canada, the Government of Ontario, and the University of Toronto.

## References

Achúcarro, A., Gong, J.-O., Hardeman, S., Palma, G. A., & Patil, S. P. 2011, JCAP, 1, 30
Ackerman, L., Carroll, S. M., & Wise, M. B. 2007, Phys. Rev. D, 75, 083502
Adshead, P., Dvorkin, C., Hu, W., & Lim, E. A. 2012, Phys. Rev. D, 85, 023531
Agrawal, A., Fujita, T., & Komatsu, E. 2018, Phys. Rev. D, 97, 103526
Agullo, I., & Parker, L. 2011, Gen. Rel. Grav., 43, 2541
Akita, Y., & Kobayashi, T. 2016, Phys. Rev. D, 93, 043519
Alishahiha, M., Silverstein, E., & Tong, D. 2004, Phys. Rev. D, 70, 123505
Allen, M. 1997, The Coefficient of Determination in Multiple Regression (Boston, MA: Springer, US), 91
Alvarez, M., Baldauf, T., Bond, J. R., et al. 2014, ArXiv e-prints [arXiv:1412.4671]
Anninos, D., De Luca, V., Franciolini, G., Kehagias, A., & Riotto, A. 2019, JCAP, 2019, 045
Arkani-Hamed, N., & Maldacena, J. 2015, ArXiv e-prints [arXiv:1503.08043]
Arkani-Hamed, N., Creminelli, P., Mukohyama, S., & Zaldarriaga, M. 2004, JCAP, 4, 1
Arkani-Hamed, N., Baumann, D., Lee, H., & Pimentel, G. L. 2018, ArXiv e-prints [arXiv:1811.00024]
Arroja, F., Mizuno, S., Koyama, K., & Tanaka, T. 2009, Phys. Rev. D, 80, 043527
Babich, D., & Zaldarriaga, M. 2004, Phys. Rev. D, 70, 083005
Bardeen, J. M. 1980, Phys. Rev. D, 22, 1882
Barnaby, N., & Cline, J. M. 2006, Phys. Rev. D, 73, 106012
Barnaby, N., Namba, R., & Peloso, M. 2012a, Phys. Rev. D, 85, 123523
Barnaby, N., Moxon, J., Namba, R., et al. 2012b, Phys. Rev. D, 86, 103508
Bartolo, N., & Orlando, G. 2017, JCAP, 2017, 034
Bartolo, N., & Riotto, A. 2009, JCAP, 3, 17
Bartolo, N., Matarrese, S., & Riotto, A. 2002, Phys. Rev. D, 65, 103505
Bartolo, N., Komatsu, E., Matarrese, S., & Riotto, A. 2004a, Phys. Rep., 402, 103
Bartolo, N., Matarrese, S., & Riotto, A. 2004b, JCAP, 1, 3
Bartolo, N., Matarrese, S., & Riotto, A. 2004c, Phys. Rev. D, 69, 043503
Bartolo, N., Matarrese, S., & Riotto, A. 2004d, Phys. Rev. Lett., 93, 231301
Bartolo, N., Matarrese, S., & Riotto, A. 2005, JCAP, 8, 10
Bartolo, N., Fasiello, M., Matarrese, S., & Riotto, A. 2010a, JCAP, 8, 8
Bartolo, N., Fasiello, M., Matarrese, S., & Riotto, A. 2010b, JCAP, 9, 35
Bartolo, N., Matarrese, S., & Riotto, A. 2010c, Adv. Astron., 2010, 157079
Bartolo, N., Fasiello, M., Matarrese, S., & Riotto, A. 2010d, JCAP, 12, 026
Bartolo, N., Matarrese, S., & Riotto, A. 2012, JCAP, 2, 17
Bartolo, N., Matarrese, S., Peloso, M., & Ricciardone, A. 2013a, Phys. Rev. D, 87, 023504
Bartolo, N., Matarrese, S., Peloso, M., & Ricciardone, A. 2013b, JCAP, 8, 22
Bartolo, N., Peloso, M., Ricciardone, A., & Unal, C. 2014, JCAP, 11, 9
Bartolo, N., Matarrese, S., Peloso, M., & Shiraishi, M. 2015, JCAP, 7, 039
Bartolo, N., Orlando, G., & Shiraishi, M. 2019, JCAP, 2019, 050
Bastero-Gil, M., Berera, A., Moss, I. G., & Ramos, R. O. 2014, JCAP, 12, 8
Battefeld, T., Niemeyer, J. C., & Vlaykov, D. 2013, JCAP, 5, 6
Baumann, D., & Green, D. 2012, Phys. Rev. D, 85, 103520
Baumann, D., Goon, G., Lee, H., & Pimentel, G. L. 2018, J. High Energy Phys., 4, 140
Becker, A., & Huterer, D. 2012, Phys. Rev. Lett., 109, 121302
Becker, A., Huterer, D., & Kadota, K. 2011, JCAP, 1, 006
Becker, A., Huterer, D., & Kadota, K. 2012, JCAP, 12, 034
Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJS, 208, 20
Bernardeau, F., & Uzan, J.-P. 2002, Phys. Rev. D, 66, 103506
Bernardeau, F., Kofman, L., & Uzan, J.-P. 2004, Phys. Rev. D, 70, 083004
Bond, J. R., Frolov, A. V., Huang, Z., & Kofman, L. 2009, Phys. Rev. Lett., 103, 071301
Boubekeur, L., Creminelli, P., D'Amico, G., Noreña, J., & Vernizzi, F. 2009, JCAP, 8, 29
Bucher, M., Moodley, K., & Turok, N. 2000, Phys. Rev. D, 62, 083508
Bucher, M., van Tent, B., & Carvalho, C. S. 2010, MNRAS, 407, 2193
Bucher, M., Racine, B., & van Tent, B. 2016, JCAP, 5, 055
Burrage, C., de Rham, C., Seery, D., & Tolley, A. J. 2011, JCAP, 1, 14
Byrnes, C. T., & Choi, K.-Y. 2010, Adv. Astron., 2010
Byrnes, C. T., Nurmi, S., Tasinato, G., & Wands, D. 2010a, JCAP, 2, 034
Byrnes, C. T., Gerstenlauer, M., Nurmi, S., Tasinato, G., & Wands, D. 2010b, JCAP, 10, 004
Chambers, A., & Rajantie, A. 2008, Phys. Rev. Lett., 100, 041302
Chen, X. 2005, Phys. Rev. D, 72, 123518
Chen, X. 2010a, JCAP, 12, 3
Chen, X. 2010b, Adv. Astron., 2010, 638979
Chen, X., & Wang, Y. 2010, JCAP, 4, 27
Chen, X., Easther, R., & Lim, E. A. 2007a, JCAP, 6, 23
Chen, X., Huang, M.-X., Kachru, S., & Shiu, G. 2007b, JCAP, 1, 002
Chen, X., Easther, R., & Lim, E. A. 2008, JCAP, 0804, 010
Chen, X., Hu, B., Huang, M.-X., Shiu, G., & Wang, Y. 2009, JCAP, 8, 8
Chen, X., Firouzjahi, H., Namjoo, M. H., & Sasaki, M. 2013, EPL (Europhys. Lett.), 102, 59001
Cheung, C., Fitzpatrick, A. L., Kaplan, J., Senatore, L., & Creminelli, P. 2008, J. High Energy Phys., 3, 14
Cicoli, M., Tasinato, G., Zavala, I., Burgess, C. P., & Quevedo, F. 2012, JCAP, 5, 39
Cook, J. L., & Sorbo, L. 2012, Phys. Rev. D, 85, 023534
Cook, J. L., & Sorbo, L. 2013, JCAP, 11, 47
Coulton, W. R., & Spergel, D. N. 2019, JCAP, 10, 56
Creminelli, P., & Zaldarriaga, M. 2004, Phys. Rev. D, 70, 083532
Creminelli, P., Nicolis, A., Senatore, L., Tegmark, M., & Zaldarriaga, M. 2006, JCAP, 5, 4
Creminelli, P., Pitrou, C., & Vernizzi, F. 2011, JCAP, 11, 25
Domènech, G., Hiramatsu, T., Lin, C., et al. 2017, JCAP, 2017, 034
Dvali, G., Gruzinov, A., & Zaldarriaga, M. 2004a, Phys. Rev. D, 69, 083505
Dvali, G., Gruzinov, A., & Zaldarriaga, M. 2004b, Phys. Rev. D, 69, 023505
Elsner, F., & Wandelt, B. D. 2009, ApJS, 184, 264
Elsner, F., & Wandelt, B. D. 2012, ArXiv e-prints [arXiv:1211.0585]
Enqvist, K., & Sloth, M. S. 2002, Nucl. Phys. B, 626, 395
Endlich, S., Nicolis, A., & Wang, J. 2013, JCAP, 10, 11
Endlich, S., Horn, B., Nicolis, A., & Wang, J. 2014, Phys. Rev. D, 90, 063506
Enqvist, K., Jokinen, A., Mazumdar, A., Multamäki, T., & Väihkönen, A. 2005, Phys. Rev. Lett., 94, 161301
Falk, T., Rangarajan, R., & Srednicki, M. 1993, ApJ, 403, L1
Fergusson, J. 2014, Phys. Rev. D, 90, 043533
Fergusson, J. R., Liguori, M., & Shellard, E. P. S. 2010, Phys. Rev. D, 82, 023502
Fergusson, J. R., Liguori, M., & Shellard, E. P. S. 2012, JCAP, 12, 32
Fergusson, J., Gruetjen, H., Shellard, E., & Liguori, M. 2015, Phys. Rev. D, 91, 023502
Fertig, A., Lehners, J.-L., & Mallwitz, E. 2014, Phys. Rev. D, 89, 103537
Finelli, F., Bucher, M., Achúcarro, A., et al. 2018, J. Cosmol. Astro-Part. Phys., 2018, 016
Flauger, R., & Pajer, E. 2011, JCAP, 1, 17
Flauger, R., McAllister, L., Pajer, E., Westphal, A., & Xu, G. 2010, JCAP, 6, 9
Flauger, R., McAllister, L., Silverstein, E., & Westphal, A. 2017, JCAP, 1710, 055

Franciolini, G., Kehagias, A., Riotto, A., & Shiraishi, M. 2018, Phys. Rev. D, 98, 043533
Gangui, A., & Martin, J. 2000, MNRAS, 313, 323
Gangui, A., Lucchin, F., Matarrese, S., & Mollerach, S. 1994, ApJ, 430, 447
Gao, X., Kobayashi, T., Yamaguchi, M., & Yokoyama, J. 2011, Phys. Rev. Lett., 107, 211301
Giannantonio, T., Porciani, C., Carron, J., Amara, A., & Pillepich, A. 2012, MNRAS, 422, 2854
Goldberg, D. M., & Spergel, D. N. 1999, Phys. Rev. D, 59, 103002
González-Nuevo, J., Toffolatti, L., & Argüeso, F. 2005, ApJ, 621, 1
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Green, D., Meyers, J., & van Engelen, A. 2017, JCAP, 1712, 005
Hall, L. M. H., & Peiris, H. V. 2008, JCAP, 1, 27
Hannestad, S., Haugbolle, T., Jarnhus, P. R., & Sloth, M. S. 2010, JCAP, 1006, 001
Hanson, D., & Lewis, A. 2009, Phys. Rev. D, 80, 063004
Heavens, A. F. 1998, MNRAS, 299, 805
Hikage, C., Koyama, K., Matsubara, T., Takahashi, T., & Yamaguchi, M. 2009, MNRAS, 398, 2188
Hikage, C., Kawasaki, M., Sekiguchi, T., & Takahashi, T. 2013a, JCAP, 7, 7
Hikage, C., Kawasaki, M., Sekiguchi, T., & Takahashi, T. 2013b, JCAP, 3, 20
Hill, J. C. 2018, Phys. Rev. D, 98, 083542
Holman, R., & Tolley, A. J. 2008, JCAP, 0805, 001
Hu, W. 2000, Phys. Rev. D, 62, 043007
Huang, M.-X., & Shiu, G. 2006, Phys. Rev. D, 74, 121301
Huang, Z., & Vernizzi, F. 2013, Phys. Rev. Lett., 110, 101303
Jung, G., & van Tent, B. 2017, JCAP, 1705, 019
Jung, G., Racine, B., & van Tent, B. 2018, JCAP, 1811, 047
Kawakami, E., Kawasaki, M., Miyamoto, K., Nakayama, K., & Sekiguchi, T. 2012, JCAP, 7, 37
Kawasaki, M., Nakayama, K., Sekiguchi, T., Suyama, T., & Takahashi, F. 2008, JCAP, 11, 19
Kawasaki, M., Nakayama, K., Sekiguchi, T., Suyama, T., & Takahashi, F. 2009, JCAP, 1, 42
Khatri, R., & Wandelt, B. D. 2009, Phys. Rev. D, 79, 023501
Khatri, R., & Wandelt, B. D. 2010, Phys. Rev. D, 81, 103518
Khoury, J., & Piazza, F. 2009, JCAP, 7, 026
Kofman, L. 2003, ArXiv e-prints [arXiv:astro-ph/0303614]
Kolb, E. W., Riotto, A., & Vallinotto, A. 2006, Phys. Rev. D, 73, 023522
Komatsu, E. 2002, ArXiv e-prints [arXiv:astro-ph/0206039]
Komatsu, E. 2010, Class. Quant. Grav., 27, 124010
Komatsu, E., & Spergel, D. N. 2001, Phys. Rev. D, 63, 063002
Komatsu, E., Spergel, D. N., & Wandelt, B. D. 2005, ApJ, 634, 14
Koyama, K., Mizuno, S., Vernizzi, F., & Wands, D. 2007, JCAP, 0711, 024
Kumar, J., Leblond, L., & Rajaraman, A. 2010, JCAP, 4, 024
Lacasa, F., Pénin, A., & Aghanim, N. 2014, MNRAS, 439, 123
Lagache, G., Puget, J.-L., & Dole, H. 2005, ARA&A, 43, 727
Langlois, D., & Lepidi, A. 2011, JCAP, 1, 8
Langlois, D., & van Tent, B. 2011, Class. Quant. Grav., 28, 222001
Langlois, D., & van Tent, B. 2012, JCAP, 7, 40
Langlois, D., Vernizzi, F., & Wands, D. 2008, JCAP, 12, 4
Lehners, J. L. 2010, Adv. Astron. , 2010
Lehners, J.-L., & Steinhardt, P. J. 2008, Phys. Rev. D, 77, 063533
Lehners, J.-L., & Steinhardt, P. J. 2013, Phys. Rev. D, 87, 123533
Lewis, A. 2012, JCAP, 6, 023
Lewis, A., Challinor, A., & Hanson, D. 2011, JCAP, 3, 18
Lewis, A., Challinor, A., & Lasenby, A. 2000, ApJ, 538, 473
Li, M. 2013, Phys. Lett. B, 724, 192
Liguori, M., Hansen, F. K., Komatsu, E., Matarrese, S., & Riotto, A. 2006, Phys. Rev. D, 73, 043505
Liguori, M., Sefusatti, E., Fergusson, J. R., & Shellard, E. P. S. 2010, Adv. Astron., 2010
Linde, A., & Mukhanov, V. 1997, Phys. Rev. D, 56, 535
Linde, A., & Mukhanov, V. 2006, JCAP, 4, 9
LoVerde, M., Miller, A., Shandera, S., & Verde, L. 2008, JCAP, 4, 014
Lyth, D. H. 2005, JCAP, 11, 6
Lyth, D. H., & Rodriguez, Y. 2005, Phys. Rev. Lett., 95, 121302
Lyth, D. H., & Riotto, A. 2006, Phys. Rev. Lett., 97, 121301
Lyth, D. H., & Wands, D. 2002, Phys. Lett. B, 524, 5
Lyth, D. H., Ungarelli, C., & Wands, D. 2003, Phys. Rev. D, 67, 023503
Maldacena, J. M., & Pimentel, G. L. 2011, J. High Energy Phys., 9, 45
Malik, K. A., & Lyth, D. H. 2006, JCAP, 9, 8
Mangilli, A., & Verde, L. 2009, Phys. Rev. D, 80, 123007
Mangilli, A., Wandelt, B., Elsner, F., & Liguori, M. 2013, A&A, 555, A82
Meerburg, P. D., van der Schaar, J. P., & Stefano Corasaniti, P. 2009, JCAP, 5, 18
Meerburg, P. D., Meyers, J., van Engelen, A., & Ali-Haïmoud, Y. 2016a, Phys. Rev. D, 93, 123511
Meerburg, P. D., Münchmeyer, M., & Wandelt, B. 2016b, Phys. Rev. D, 93, 043536
Mirbabayi, M., Senatore, L., Silverstein, E., & Zaldarriaga, M. 2015, Phys. Rev. D, 91, 063518
Mollerach, S. 1990, Phys. Rev. D, 42, 313
Moradinezhad Dizgah, A., Lee, H., Muñoz, J. B., & Dvorkin, C. 2018, JCAP, 5, 013
Moroi, T., & Takahashi, T. 2001, Phys. Lett. B, 522, 215
Moss, I. G., & Xiong, C. 2007, JCAP, 4, 7
Münchmeyer, M., Bouchet, F., Jackson, M., & Wandelt, B. 2014, A&A, 570, A94
Münchmeyer, M., Meerburg, P. D., & Wandelt, B. D. 2015, Phys. Rev. D, 91, 043534
Munshi, D., & Heavens, A. 2010, MNRAS, 401, 2406
Namba, R., Peloso, M., Shiraishi, M., Sorbo, L., & Unal, C. 2016, JCAP, 1, 041
Naskar, A., & Pal, S. 2018, Phys. Rev. D, 98, 083520
Nitta, D., Komatsu, E., Bartolo, N., Matarrese, S., & Riotto, A. 2009, JCAP, 5, 14
Noumi, T., Yamaguchi, M., & Yokoyama, D. 2013, J. High Energy Phys., 6, 51
Oppizzi, F., Liguori, M., Renzi, A., Arroja, F., & Bartolo, N. 2018, JCAP, 5, 045
Ozsoy, O., Mylova, M., Parameswaran, S., et al. 2019, JCAP, 9, 036
Peloton, J., Schmittfull, M., Lewis, A., Carron, J., & Zahn, O. 2017, Phys. Rev. D, 95, 043508
Pénin, A., Lacasa, F., & Aghanim, N. 2014, MNRAS, 439, 143
Pettinari, G. W., Fidler, C., Crittenden, R., et al. 2014, Phys. Rev. D, 90, 103010
Planck Collaboration XVII. 2014, A&A, 571, A17
Planck Collaboration XIX. 2014, A&A, 571, A19
Planck Collaboration XXII. 2014, A&A, 571, A22
Planck Collaboration XXIV. 2014, A&A, 571, A24
Planck Collaboration XVII. 2016, A&A, 594, A17
Planck Collaboration XX. 2016, A&A, 594, A20
Planck Collaboration XXI. 2016, A&A, 594, A21
Planck Collaboration I. 2020, A&A, 641, A1
Planck Collaboration II. 2020, A&A, 641, A2
Planck Collaboration III. 2020, A&A, 641, A3
Planck Collaboration IV. 2020, A&A, 641, A4
Planck Collaboration V. 2020, A&A, 641, A5
Planck Collaboration VI. 2020, A&A, 641, A6
Planck Collaboration VII. 2020, A&A, 641, A7
Planck Collaboration VIII. 2020, A&A, 641, A8
Planck Collaboration IX. 2020, A&A, 641, A9
Planck Collaboration X. 2020, A&A, 641, A10
Planck Collaboration XI. 2020, A&A, 641, A11
Planck Collaboration XII. 2020, A&A, 641, A12
Qiu, T., Gao, X., & Saridakis, E. N. 2013, Phys. Rev. D, 88, 043525
Ratra, B. 1992, ApJ, 391, L1
Rigopoulos, G. I., Shellard, E. P. S., & van Tent, B. J. W. 2006, Phys. Rev. D, 73, 083522
Rigopoulos, G. I., Shellard, E. P. S., & van Tent, B. J. W. 2007, Phys. Rev. D, 76, 083512
Riotto, A., & Sloth, M. S. 2011, Phys. Rev. D, 83, 041301
Sachs, R. K., & Wolfe, A. M. 1967, ApJ, 147, 73
Salem, M. P. 2005, Phys. Rev. D, 72, 123516
Sasaki, M., Valiviita, J., & Wands, D. 2006, Phys. Rev. D, 74, 103003
Sefusatti, E., Liguori, M., Yadav, A. P. S., Jackson, M. G., & Pajer, E. 2009, JCAP, 12, 022
Senatore, L., & Zaldarriaga, M. 2012a, JCAP, 1208, 001
Senatore, L., & Zaldarriaga, M. 2012b, JHEP, 1204, 024
Senatore, L., Tassev, S., & Zaldarriaga, M. 2009, JCAP, 9, 38
Senatore, L., Smith, K. M., & Zaldarriaga, M. 2010, JCAP, 1, 28
Senatore, L., Silverstein, E., & Zaldarriaga, M. 2014, JCAP, 2014, 016
Shandera, S., Dalal, N., & Huterer, D. 2011, JCAP, 3, 017
Shiraishi, M. 2012, JCAP, 6, 15
Shiraishi, M., Komatsu, E., Peloso, M., & Barnaby, N. 2013a, JCAP, 5, 2
Shiraishi, M., Ricciardone, A., & Saga, S. 2013b, JCAP, 11, 51
Shiraishi, M., Liguori, M., & Fergusson, J. R. 2014, JCAP, 5, 8
Shiraishi, M., Liguori, M., & Fergusson, J. R. 2015, JCAP, 1, 7
Shiraishi, M., Liguori, M., Fergusson, J. R., & Shellard, E. P. S. 2019, JCAP, 06, 046
Silverstein, E., & Tong, D. 2004, Phys. Rev. D, 70, 103505
Sitwell, M., & Sigurdson, K. 2014, Phys. Rev. D, 89, 123509
Smith, K. M., Senatore, L., & Zaldarriaga, M. 2009, JCAP, 9, 6
Smith, K. M., Loverde, M., & Zaldarriaga, M. 2011, Phys. Rev. Lett., 107, 191301
Smith, K. M., Senatore, L., & Zaldarriaga, M. 2015, ArXiv e-prints [arXiv:1502.00635]
Su, S.-C., Lim, E. A., & Shellard, E. P. S. 2012, ArXiv e-prints [arXiv:1212.6968]

Toffolatti, L., Argueso Gomez, F., de Zotti, G., et al. 1998, MNRAS, 297, 117
Torrado, J., Byrnes, C. T., Hardwick, R. J., Vennin, V., & Wands, D. 2018, Phys. Rev. D, 98, 063525
Tzavara, E., & van Tent, B. 2011, JCAP, 6, 26
Verde, L., Wang, L., Heavens, A. F., & Kamionkowski, M. 2000, MNRAS, 313, 141
Vernizzi, F., & Wands, D. 2006, JCAP, 5, 19
Wang, L., & Kamionkowski, M. 2000, Phys. Rev. D, 61, 063504
Weinberg, S. 2008, Phys. Rev. D, 77, 123541
Yadav, A. P. S., Komatsu, E., & Wandelt, B. D. 2007, ApJ, 664, 680
Yadav, A. P. S., Komatsu, E., Wandelt, B. D., et al. 2008, ApJ, 678, 578
Yadav, A. P. S., & Wandelt, B. D. 2010, Adv. Astron., 2010
Zaldarriaga, M. 2004, Phys. Rev. D, 69, 043508

[1] AIM, CEA, CNRS, Université Paris-Saclay, Université Paris-Diderot, Sorbonne Paris Cité, 91191 Gif-sur-Yvette, France
[2] APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/lrfu, Observatoire de Paris, Sorbonne Paris Cité, 10 rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France
[3] African Institute for Mathematical Sciences, 6-8 Melrose Road, Muizenberg, Cape Town, South Africa
[4] Astrophysics Group, Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, UK
[5] Astrophysics & Cosmology Research Unit, School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Private Bag X54001, Durban 4000, South Africa
[6] CITA, University of Toronto, 60 St. George St., Toronto, ON M5S 3H8, Canada
[7] CNRS, IRAP, 9 Av. colonel Roche, BP 44346, 31028 Toulouse cedex 4, France
[8] California Institute of Technology, Pasadena, CA, USA
[9] Centre for Theoretical Cosmology, DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK
[10] Computational Cosmology Center, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[11] Département de Physique Théorique, Université de Genève, 24 quai E. Ansermet, 1211 Genève 4, Switzerland
[12] Département de Physique, École normale supérieure, PSL Research University, CNRS, 24 rue Lhomond, 75005 Paris, France
[13] Departamento de Astrofísica, Universidad de La Laguna (ULL), 38206 La Laguna, Tenerife, Spain
[14] Departamento de Física, Universidad de Oviedo, C/ Federico García Lorca, 18, Oviedo, Spain
[15] Department of Astrophysics/IMAPP, Radboud University, PO Box 9010, 6500 GL Nijmegen, The Netherlands
[16] Department of General Education, National Institute of Technology, Kagawa College, 355 Chokushi-cho, Takamatsu, Kagawa 761-8058, Japan
[17] Department of Mathematics, University of Stellenbosch, Stellenbosch 7602, South Africa
[18] Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia, Canada
[19] Department of Physics & Astronomy, University of the Western Cape, Cape Town 7535, South Africa
[20] Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK
[21] Department of Physics, Gustaf Hällströmin katu 2a, University of Helsinki, Helsinki, Finland
[22] Department of Physics, Princeton University, Princeton, NJ, USA
[23] Department of Physics, University of California, Santa Barbara, CA, USA
[24] Dipartimento di Fisica e Astronomia G. Galilei, Università degli Studi di Padova, Via Marzolo 8, 35131 Padova, Italy
[25] Dipartimento di Fisica e Scienze della Terra, Università di Ferrara, Via Saragat 1, 44122 Ferrara, Italy
[26] Dipartimento di Fisica, Università La Sapienza, P.le A. Moro 2, Roma, Italy
[27] Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria, 16, Milano, Italy
[28] Dipartimento di Fisica, Università degli Studi di Trieste, Via A. Valerio 2, Trieste, Italy
[29] Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy
[30] European Space Agency, ESAC, Planck Science Office, Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, Madrid, Spain
[31] European Space Agency, ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
[32] Gran Sasso Science Institute, INFN, viale F. Crispi 7, 67100 L'Aquila, Italy
[33] HEP Division, Argonne National Laboratory, Lemont, IL 60439, USA
[34] Haverford College Astronomy Department, 370 Lancaster Avenue, Haverford, PA, USA
[35] Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland
[36] INAF – OAS Bologna, Istituto Nazionale di Astrofisica – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Area della Ricerca del CNR, Via Gobetti 101, 40129 Bologna, Italy
[37] INAF – Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, Padova, Italy
[38] INAF – Osservatorio Astronomico di Trieste, Via G.B. Tiepolo 11, Trieste, Italy
[39] INAF, Istituto di Radioastronomia, Via Piero Gobetti 101, 40129 Bologna, Italy
[40] INAF/IASF Milano, Via E. Bassini 15, Milano, Italy
[41] INFN – CNAF, viale Berti Pichat 6/2, 40127 Bologna, Italy
[42] INFN, Sezione di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italy
[43] INFN, Sezione di Ferrara, Via Saragat 1, 44122 Ferrara, Italy
[44] INFN, Sezione di Milano, Via Celoria 16, Milano, Italy
[45] INFN, Sezione di Roma 1, Università di Roma Sapienza, Piazzale Aldo Moro 2, 00185 Roma, Italy
[46] INFN, Sezione di Roma 2, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy
[47] Imperial College London, Astrophysics group, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK
[48] Institut d'Astrophysique Spatiale, CNRS, Univ. Paris-Sud, Université Paris-Saclay, Bât. 121, 91405 Orsay cedex, France
[49] Institut d'Astrophysique de Paris, CNRS (UMR7095), 98bis boulevard Arago, 75014 Paris, France
[50] Institute Lorentz, Leiden University, PO Box 9506, 2300 RA Leiden, The Netherlands
[51] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[52] Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo, Norway
[53] Instituto de Astrofísica de Canarias, C/Vía Láctea s/n, La Laguna, Tenerife, Spain
[54] Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências da Universidade de Lisboa, Campo Grande 1749-016, Lisboa, Portugal
[55] Instituto de Física de Cantabria (CSIC-Universidad de Cantabria), Avda. de los Castros s/n, Santander, Spain
[56] Istituto Nazionale di Fisica Nucleare, Sezione di Padova, Via Marzolo 8, 35131 Padova, Italy
[57] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, USA
[58] Jodrell Bank Centre for Astrophysics, Alan Turing Building, School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester M13 9PL, UK
[59] Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[60] Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), UTIAS, The University of Tokyo, Chiba 277-8583, Japan

[61] Laboratoire d'Océanographie Physique et Spatiale (LOPS), Univ. Brest, CNRS, Ifremer, IRD, Brest, France
[62] Laboratoire de Physique Subatomique et Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, 53 rue des Martyrs, 38026 Grenoble Cedex, France
[63] Laboratoire de Physique Théorique, Université Paris-Sud 11 & CNRS, Bâtiment 210, 91405 Orsay, France
[64] Low Temperature Laboratory, Department of Applied Physics, Aalto University, Espoo 00076, Finland
[65] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
[66] Mullard Space Science Laboratory, University College London, Surrey RH5 6NT, UK
[67] NAOC-UKZN Computational Astrophysics Centre (NUCAC), University of KwaZulu-Natal, Durban 4000, South Africa
[68] National Centre for Nuclear Research, ul. L. Pasteura 7, 02-093 Warsaw, Poland
[69] Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada
[70] Purple Mountain Observatory, No. 8 Yuan Hua Road, 210034 Nanjing, PR China
[71] SISSA, Astrophysics Sector, Via Bonomea 265, 34136 Trieste, Italy
[72] San Diego Supercomputer Center, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
[73] School of Chemistry and Physics, University of KwaZulu-Natal, Westville Campus, Private Bag X54001, Durban 4000, South Africa
[74] School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff CF24 3AA, UK
[75] School of Physics and Astronomy, Sun Yat-sen University, 2 Daxue Rd, Tangjia, Zhuhai, PR China
[76] School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK
[77] School of Physics, Indian Institute of Science Education and Research Thiruvananthapuram, Maruthamala PO, Vithura, Thiruvananthapuram 695551, Kerala, India
[78] School of Physics, The University of New South Wales, Sydney, NSW 2052, Australia
[79] Simon Fraser University, Department of Physics, 8888 University Drive, Burnaby, BC, Canada
[80] Sorbonne Université, Observatoire de Paris, Université PSL, École normale supérieure, CNRS, LERMA, 75005 Paris, France
[81] Sorbonne Université, UMR7095, Institut d'Astrophysique de Paris, 98bis boulevard Arago, 75014 Paris, France
[82] Space Research Institute (IKI), Russian Academy of Sciences, Profsoyuznaya Str, 84/32, Moscow 117997, Russia
[83] Space Science Data Center – Agenzia Spaziale Italiana, Via del Politecnico snc, 00133 Roma, Italy
[84] Space Sciences Laboratory, University of California, Berkeley, CA, USA
[85] The Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, AlbaNova, 106 91 Stockholm, Sweden
[86] Université de Toulouse, UPS-OMP, IRAP, 31028 Toulouse Cedex 4, France
[87] Van Swinderen Institute for Particle Physics and Gravity, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
[88] Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland