Planck 2015 results

XVI. Isotropy and statistics of the CMB


(Affiliations can be found after the references)

Received 5 June 2015 / Accepted 7 September 2015

ABSTRACT

We test the statistical isotropy and Gaussianity of the cosmic microwave background (CMB) anisotropies using observations made by the Planck satellite. Our results are based mainly on the full Planck mission for temperature, but also include some polarization measurements. In particular, we consider the CMB anisotropy maps derived from the multi-frequency Planck data by several component-separation methods. For the temperature anisotropies, we find excellent agreement between these sky maps over both a very large fraction of the sky and a broad range of angular scales, establishing that potential foreground residuals do not affect our studies. Tests of skewness, kurtosis, multi-normality, N-point functions, and Minkowski functional indicators indicate consistency with Gaussianity, while a power deficit across large angular scales is manifested in several ways, for example low map variance. The results of a peak statistics analysis are consistent with the expectations of a Gaussian random field. The “Cold Spot” is detected with several methods, including map kurtosis, peak statistics, and mean temperature profile. We thoroughly probe the large-scale dipole power asymmetry, detecting it with several independent tests, and address the subject of a posteriori correction. Tests of directionality suggest the presence of angular clustering from large to small scales, but at a significance that is dependent on the details of the approach. We perform the first examination of polarization data, finding the morphology of stacked peaks to be consistent with the expectations of statistically isotropic simulations. Where they overlap, these results are consistent with the Planck 2013 analysis based on the nominal mission data and provide our most thorough view of the statistics of the CMB fluctuations to date.

Key words. cosmology: observations – cosmic background radiation – polarization – methods: data analysis – methods: statistical

* Corresponding author: A. J. Banday, e-mail: anthony.banday@irap.omp.eu

A16, page 1 of 62

Article published by EDP Sciences
1. Introduction

This paper, one of a set associated with the 2015 release of data from the Planck\(^1\) mission (Planck Collaboration I 2016), describes a set of studies undertaken to determine the statistical properties of both the temperature and polarization anisotropies of the cosmic microwave background (CMB).

The standard cosmological model is described well by the Friedmann-Lemaître-Robertson-Walker solution of the Einstein field equations. This model is characterized by a homogeneous and isotropic background metric and a scale factor of the expanding Universe. It is hypothesized that at very early times the Universe went through a period of accelerated expansion, the so-called “cosmological inflation”, driven by a hypothetical scalar field, the “inflaton”. During inflation the Universe behaves approximately as a de Sitter space, providing the conditions by which some of its present properties can be realized and specifically relaxing the problem of initial conditions. In particular, the seeds that gave rise to the present large-scale matter distribution via gravitational instability originated as quantum fluctuations of the inflaton about its vacuum state. These fluctuations in the inflaton produce energy density perturbations that are distributed as a statistically homogeneous and isotropic Gaussian random field. Linear theory relates those perturbations to the temperature and polarization anisotropies of the CMB, implying a distribution for the anisotropies very close to that of a statistically isotropic Gaussian random field.

The aim of this paper is to use the full mission Planck data to test the Gaussianity and isotropy of the CMB as measured in both intensity and, in a more limited capacity, polarization. Testing these fundamental properties is crucial for the validation of the standard cosmological scenario, and has profound implications for our understanding of the physical nature of the Universe and the initial conditions of structure formation. Moreover, the confirmation of the statistically isotropic and Gaussian nature of the CMB is essential for justifying the corresponding assumptions usually made when estimating the CMB power spectra and other quantities to be obtained from the Planck data. Indeed, the isotropy and Gaussianity of the CMB anisotropies are implicitly assumed in critical science papers from the 2015 release, in particular those describing the likelihood and the derivation of cosmological parameter constraints (Planck Collaboration XI 2016; Planck Collaboration XIII 2016). Conversely, if the detection of significant deviations from these assumptions cannot be traced to known systematic effects or foreground residuals, the presence of which should be diagnosed by the statistical tests set forth in this paper, this would necessitate a major revision of the current methodological approaches adopted in deriving the mission’s many science results.

Well-understood physical processes due to the integrated Sachs-Wolfe (ISW) effect (Planck Collaboration XVII 2014; Planck Collaboration XXI 2016) and gravitational lensing (Planck Collaboration XIX 2014; Planck Collaboration XV 2016) lead to secondary anisotropies that exhibit marked deviations from Gaussianity. In addition, Doppler boosting, due to our motion with respect to the CMB rest frame, induces both a dipolar modulation of the temperature anisotropies and an aberration that corresponds to a change in the apparent arrival directions of the CMB photons (Challinor & van Leeuwen 2002). Both of these effects are aligned with the CMB dipole, and were detected at a statistically significant level on small angular scales in Planck Collaboration XXVII (2014). Beyond these, Planck Collaboration XXIII (2014, hereafter PCIS13) established that the Planck 2013 data set showed little evidence for non-Gaussianity, with the exception of a number of CMB temperature anisotropy anomalies on large angular scales that confirmed earlier claims based on WMAP data. Moreover, given that the broader frequency coverage of the Planck instruments allowed improved component separation methods to be applied in the derivation of foreground-cleaned CMB maps, it was generally considered that the case for anomalous features in the CMB had been strengthened. Hence, such anomalies have attracted considerable attention in the community, since they could be the visible traces of fundamental physical processes occurring in the early Universe.

However, the literature also supports an ongoing debate about the significance of these anomalies. The central issue in this discussion is connected with the role of a posteriori choices – whether interesting features in the data bias the choice of statistical tests, or if arbitrary choices in the subsequent data analysis enhance the significance of the features. Indeed, the WMAP team (Bennett et al. 2011) base their rejection of the presence of anomalies in the CMB on such arguments. Of course, one should attempt to correct for any choices that were made in the process of detecting an anomaly. However, in the absence of an alternative model for comparison to the standard Gaussian, statistically isotropic one adopted to quantify significance, this is often simply not possible. In this work, whilst it is recognized that care must be taken in the assessment of significance, we proceed on the basis that allowing a posteriori reasoning permits us to challenge the limits of our existing knowledge (Pontzen & Peiris 2010). That is, by focusing on specific properties of the observed data that are shown to be empirically interesting, we may open up new paths to a better theoretical understanding of the Universe. We will clearly describe the methodology applied to the data, and attempt to study possible links among the anomalies in order to search for a physical interpretation.

The analysis of polarization data introduces a new opportunity to explore the statistical properties of the CMB sky, including the possibility of improvement of the significance of detection of large-scale anomalies. However, this cannot be fully included in the current data assessment, since the component-separation products in polarization are high-pass filtered to remove large angular scales (Planck Collaboration IX 2016), owing to the persistence of significant systematic artefacts originating in the High Frequency Instrument (HFI) data (Planck Collaboration VII 2016; Planck Collaboration VIII 2016). In addition, limitations of the simulations with which the data are to be compared (Planck Collaboration XII 2016), in particular a significant mismatch in noise properties, limit the extent to which any polarization results can be included. Therefore, we only present a stacking analysis of the polarized data, although this is a significant extension of previous approaches found in the literature.

With future Planck data releases, it will be important to determine in more detail whether there are any peculiarities in the CMB polarization, and if so, whether they are related to existing features in the CMB temperature field. Conversely, the absence of any corresponding features in polarization might imply that

---

\(^1\) 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).

A16, page 2 of 62
the temperature anomalies (if they are not simply flukes) could be due to a secondary effect such as the ISW effect, or alternative scenarios in which the anomalies arise from physical processes that do not correlate with the temperature, e.g., texture or def-ect models. Either one of these possible outcomes could yield a breakthrough in understanding the nature of the CMB anomalies. Of course, there also remains the possibility that anomalies may be found in the polarization data that are unrelated to existing features in the temperature measurements.

Following the approach established in Planck Collaboration XXIII (2014), throughout this paper we quantify the significance of a test statistic in terms of the p-value. This is the probability of obtaining a test statistic at least as extreme as the observed one, under the assumption that the null hypothesis (i.e., primordial Gaussianity and isotropy of the CMB) is true. In some tests, where it is clearly justified to only use a one-tailed probability, the p-value is replaced by the corresponding upper-or lower-tail probability.

This paper covers all relevant aspects related to the phenomenological study of the statistical isotropy and Gaussian nature of the CMB measured by the Planck satellite. Specific theoretically-motivated model constraints on isotropy or non-Gaussianity, as might arise from non-standard inflationary models, the geometry and topology of the Universe, and primordial magnetic fields are provided in the companion papers (Planck Collaboration XVII 2016; Planck Collaboration XX 2016; Planck Collaboration XVIII 2016; Planck Collaboration XIX 2016). The paper is organized as follows. Section 2 summarizes the Planck full mission data used for the analyses, and important limitations of the polarization maps that are studied. Section 3 describes the characteristics of the simulations that constitute our reference set of Gaussian sky maps representative of the null hypothesis. In Sect. 4 the null hypothesis is tested with a number of standard tests that probe different aspects of non-Gaussianity. Several important anomalous features of the CMB sky, originally detected with the WMAP data and subsequently confirmed in PCIS13, are re-assessed in Sect. 5. Aspects of the CMB fluctuations specifically related to dipolar asymmetry are examined in Sect. 6. The sensitivity of the results for a number of statistical tests to the sky fraction is examined in Sect. 7. Section 8 presents tests of the statistical nature of the polarization signal observed by Planck using a local analysis of stacked patches of the sky. Finally, Sect. 9 provides the main conclusions of the paper.

2. Data description

In this paper, we use data from the Planck-2015 full mission data release. This contains approximately 29 months of data for the HFI and 50 months for the Low Frequency Instrument (LFI). The release includes sky maps at nine frequencies in intensity (seven in polarization), with corresponding “half-mission” maps that are generated by splitting the full-mission data sets in various ways. The maps are provided in HEALPix format (Görski et al. 2005), with a pixel size defined by the $N_{\text{side}}$ parameter. This set of maps allows a variety of consistency checks to be made, together with estimates of the instrumental noise contributions to our analyses and limits on time-varying systematic artefacts. Full details are provided in a series of companion papers (Planck Collaboration II 2016; Planck Collaboration III 2016; Planck Collaboration IV 2016; Planck Collaboration V 2016; Planck Collaboration VI 2016; Planck Collaboration VII 2016; Planck Collaboration VIII 2016).

Our main results are based on estimates of the CMB generated by four distinct component-separation algorithms – Commander, NILC, SEVEM, and SMICA – as described in Planck Collaboration IX (2016). These effectively combine the raw Planck frequency maps in such a way as to minimize foreground residuals from diffuse Galactic emission. Note that the additional information in the full mission data set allows us to improve the reconstruction noise levels by roughly a factor of 2 (in temperature) as compared to the Planck-2013 nominal mission data release. The CMB intensity maps were derived using all channels, from 30 to 857 GHz, and provided at a common angular resolution of 5° FWHM and $N_{\text{side}} = 2048$. The intensity maps are only partially corrected for the second order temperature quadrupole (Kamionkowski & Knox 2003). Therefore, where appropriate, the component-separated maps should be corrected for the residual contribution (Notari & Quartin 2015), specifically as described in Planck Collaboration IX (2016). The polarization solutions include all channels sensitive to polarization, from 30 to 353 GHz, at a resolution of 10° FWHM and $N_{\text{side}} = 1024$. Possible residual emission is then mitigated in our analyses by the use of sky-coverage masks, provided for both intensity and polarization.

Since in some cases it is important to study the frequency dependence of the cosmological signal, either to establish its primordial origin or to test for the frequency dependence associated with specific effects such as Doppler boosting (see Sect. 6.4), we also consider the foreground-cleaned versions of the 100, 143, and 217 GHz sky maps generated by the SEVEM algorithm (Planck Collaboration IX 2016), hereafter referred to as SEVEM – 100, SEVEM – 143, and SEVEM – 217, respectively.

For the present release, a post-processing high-pass-filtering has been applied to the CMB polarization maps in order to mitigate residual large-scale systematic errors in the HFI channels (Planck Collaboration VII 2016). The filter results in the elimination of structure in the maps on angular scales larger than about 10°, and a weighted suppression of power down to scales of 5°, below which the maps remain unprocessed.

Lower-resolution versions of these data sets are also used in the analyses presented in this paper. The downgrading procedure for maps is to decompose them into spherical harmonics on the full sky at the input HEALPix resolution. The spherical harmonic coefficients, $a_{\ell m}$, are then convolved to the new resolution using

$$ a_{\ell m}^\text{out} = \frac{\int P_{\ell m}^\text{in} \, dV_{\ell m}^\text{in}}{\int P_{\ell m}^\text{out} \, dV_{\ell m}^\text{out}} a_{\ell m}, $$

where $b_{\ell m}$ is the beam transfer function, $p_{\ell m}$ is the HEALPix pixel window function, and the “in” and “out” superscripts denote the input and output resolutions. They are then synthesized into a map directly at the output HEALPix resolution. Masks are downgraded in a similar way. The binary mask at the starting resolution is first downgraded like a temperature map. The smooth downgraded mask is then thresholded by setting pixels where the value is less than 0.9 to zero and all others to unity in order to make a binary mask. Table 1 lists the $N_{\text{side}}$ and FWHM values defining the resolution of these maps, together with the different masks and their sky coverages that accompany the signal maps. In general, we make use of standardized masks that are the union of those associated with the individual component-separation methods.

As recommended in Planck Collaboration IX (2016), the mask UT78 is adopted for all high-resolution analyses of temperature data. UT76 is an extended version of this mask more suitable for some non-Gaussianity studies. The mask preferred
Table 1. Standardized data sets used in this paper.

<table>
<thead>
<tr>
<th>$N_{side}$</th>
<th>FWHM [arcmin]</th>
<th>Mask</th>
<th>Unmasked pixels [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2048 . . .</td>
<td>5 UT78</td>
<td></td>
<td>77.6</td>
</tr>
<tr>
<td>2048 . . .</td>
<td>5 UT76</td>
<td></td>
<td>76.1</td>
</tr>
<tr>
<td>1024 . .</td>
<td>10 UT10276</td>
<td></td>
<td>75.6</td>
</tr>
<tr>
<td>512 .</td>
<td>20 UT5174</td>
<td></td>
<td>73.7</td>
</tr>
<tr>
<td>256 .</td>
<td>40 UT2567</td>
<td></td>
<td>72.5</td>
</tr>
<tr>
<td>128 .</td>
<td>80 UT12878</td>
<td></td>
<td>69.7</td>
</tr>
<tr>
<td>64 .</td>
<td>160 UT6467</td>
<td></td>
<td>67.0</td>
</tr>
<tr>
<td>32 .</td>
<td>320 UT3264</td>
<td></td>
<td>63.8</td>
</tr>
<tr>
<td>16 .</td>
<td>640 UT1658</td>
<td></td>
<td>58.4</td>
</tr>
<tr>
<td>1024 .</td>
<td>10 UPB77</td>
<td></td>
<td>77.4</td>
</tr>
</tbody>
</table>

**Notes.** The resolutions of the sky maps used are defined in terms of the $N_{side}$ parameter and corresponding FWHM of the Gaussian beam with which they are convolved. The corresponding common masks and the fraction of unmasked pixels used for analysis are also specified.

The results presented in this paper are derived using the extensive full focal plane (FFP8) simulations described in Planck Collaboration XII (2016). Of most importance are the Monte Carlo (MC) simulations that provide the reference set of Gaussian sky maps used for the null tests employed here. They also form the basis of any debiasing in the analysis of the real data as required by certain statistical methods. The simulations include both CMB signal and instrumental noise realizations that capture important characteristics of the Planck scanning strategy, telescope, detector responses, and data reduction pipeline over the full mission period. In particular, the signal realizations include FEBecOp (Mitra et al. 2011) beam convolution at each of the Planck frequencies, and are propagated through the various component-separation pipelines using the same weights as derived from the Planck full mission data analysis.

The FFP8 fiducial CMB power spectrum has been adopted from our best estimate of the cosmological parameters from the first Planck data release (Planck Collaboration I 2014). This corresponds to a cosmology with baryon density given by $\omega_b = \Omega_b h^2 = 0.0222$, cold dark matter (CDM) density $\omega_c = \Omega_c h^2 = 0.1203$, neutrino energy density $\omega_\nu = \Omega_\nu h^2 = 0.00064$, density parameter for the cosmological constant $\Omega_\Lambda = 0.6823$, Hubble parameter $H_0 = 100h$ km s$^{-1}$ Mpc$^{-1}$ with $h = 0.6712$, spectral index of the power spectrum of the primordial curvature perturbation $n_s = 0.96$, and amplitude of the primordial power spectrum (at $k = 0.05$ Mpc$^{-1}$) $A_s = 2.09 \times 10^{-9}$, and with the Thomson optical depth through reionization defined to be $\tau = 0.065$. Each realization of the CMB sky is generated including lensing, Rayleigh scattering, and Doppler boosting effects, the latter two of which are frequency-dependent. Unfortunately, the aberration contribution to the Doppler boost was erroneously omitted from the simulations, but, with possible exceptions described in Sect. 6, this does not lead to any significant impact on the results in this paper. A second order temperature quadrupole (Kamionkowski & Knox 2003) is added to each simulation with an amplitude corresponding to the residual uncorrected contribution present in the observed data, as described in Planck Collaboration XII (2016).

However, the Planck maps were effectively renormalized by approximately 2% to 3% in power in the time between the generation of the FFP8 simulations and the final maps. As discussed in Planck Collaboration XII (2016), correction for this calibration effect should have no significant impact on cosmological parameters. As recommended, in this paper the CMB component of the simulations is simply rescaled by a factor of 1.0134 before analysis.

Of somewhat more importance is an observed noise mismatch between the simulations and the data. Whilst this has essentially no impact on studies of temperature anisotropy, it imposes important limitations on the statistical studies of polarization sky maps that can be included here. Conversely, analyses based on 1-point statistics, such as the variance, and the $N$-point correlation functions have played important roles in establishing the nature of this mismatch, which seems to be scale-dependent with an amplitude around 20% at lower resolutions but falling to a few per cent at higher resolution. As a consequence, this paper only includes results from a stacking analysis of the polarized data, in which the stacking of the data themselves necessarily acts to lower the effect of the noise mismatch. Polarization studies that do not rely on auto-statistics can still yield interesting new results, as found in Planck Collaboration XIII (2016); Planck Collaboration XVII (2016); Planck Collaboration XVIII (2016).

4. Tests of non-Gaussianity

There is no unique signature of non-Gaussianity, but the application of a variety of tests over a range of angular scales allows us to probe the data for departures from theoretically motivated Gaussian statistics. One of the more important tests in the context of inflationary cosmology is related to the analysis of the bispectrum. This is discussed thoroughly in Planck Collaboration XVII (2016), and is therefore not discussed further in this paper. In this section, we present the results from a variety of statistical tools. Unless otherwise specified, the analyses are applied to all four component separation products (Commander, NILC, SEVEM, and SMICA) at a given resolution with the accompanying common mask, and significance levels are determined by comparison with the corresponding results derived from the FFP8 simulations, with typically 1000 being used for this purpose. Establishing the consistency of the results derived from the different component-separation techniques is essential in order to be able to make robust claims about the statistical nature of the observed temperature fluctuations, and potential deviations from Gaussianity.

4.1. One-dimensional moments

In this section we consider simple tests of Gaussianity based on the variance, skewness, and kurtosis of the CMB temperature
The results are in good agreement with PCIS13; the skewness, or kurtosis than the observed map, for these analyses. The results, also presented in Table 2, are similar to those found for the Planck 2015 component-separated CMB maps at \( N_{\text{side}} = 16 \) and 32, then compared with the equivalent quantities derived from the corresponding FFP8 simulations. For those cases in which the covariance matrix \( C_{ij} \) is fully defined by the Planck angular power spectrum \( C_l \):
analysis are presented in Table 3 and indicate that the data are consistent with Gaussianity. We note that the lower-tail probabilities for the N-pdf decrease when the resolution of the data is increased from $N_{\text{side}} = 16$ to 32. However, this behaviour is consistent with that seen for simulations, and should not be considered to be significant.

### 4.3. N-point correlation functions

In this section, we present tests of the non-Gaussianity of the Planck 2015 temperature CMB maps using real-space N-point correlation functions. While harmonic-space methods are often preferred over real-space methods for studying primordial fluctuations, real-space methods have an advantage with respect to systematic errors and foregrounds, since such effects are usually localized in real space. It is therefore important to analyse the data in both spaces in order to highlight different features.

An N-point correlation function is defined as the average product of N temperatures, measured in a fixed relative orientation on the sky,

$$C_N(\theta_1, \ldots, \theta_{2N-3}) = \langle T(\hat{n}_1) \cdots T(\hat{n}_N) \rangle,$$

(5)

where the unit vectors $\hat{n}_1, \ldots, \hat{n}_N$ span an N-point polygon. Under the assumption of statistical isotropy, these functions depend only on the shape and size of the N-point polygon, and not on its particular position or orientation on the sky. Hence, the smallest number of parameters that uniquely determines the shape and size of the N-point polygon is $2N - 3$.

The correlation functions are estimated by simple product averages over all sets of N pixels fulfilling the geometric requirements set by $\theta_1, \ldots, \theta_{2N-3}$ characterizing the shape and size of the polygon,

$$\hat{C}_N(\theta_1, \ldots, \theta_{2N-3}) = \frac{\sum (w_1 \cdots w_N) (T_1 \cdots T_N)}{\sum w_1 \cdots w_N},$$

(6)

Pixel weights $w_1, \ldots, w_N$ can be introduced in order to reduce noise or mask boundary effects. Here they represent masking by being set to 1 for included pixels and to 0 for excluded pixels.

The shapes of the polygons selected for the analysis are the pseudo-collapsed and equilateral configurations for the 3-point function, and the rhombic configuration for the 4-point function, composed of two equilateral triangles that share a common side. We use the same definition of pseudo-collapsed as in Eriksen et al. (2005), i.e., an isosceles triangle where the length of the baseline falls within the second bin of the separation angles. The length of the longer edge of the triangle, $\theta$, parameterizes its size. Analogously, in the case of the equilateral triangle and rhombus, the size of the polygon is parameterized by the length of the edge, $\theta$. Note that these functions are chosen for ease of implementation, not because they are better suited for testing Gaussianity than other configurations. For a Gaussian field, Wick’s theorem (Wick 1950) means that the ensemble average of the 4-point function may be written in terms of the 2-point function. In the following, all results refer to the connected 4-point function, i.e., are corrected for this Gaussian contribution.

We use a simple $\chi^2$ statistic to quantify the agreement between the observed data and simulations, defined by

$$\chi^2 = \sum_{i,j} \left( \hat{C}_N(\theta_i) - \langle C_N(\theta_i) \rangle \right) M_{ij}^{-1} \left( \hat{C}_N(\theta_j) - \langle C_N(\theta_j) \rangle \right),$$

(7)

Here, $\hat{C}_N(\theta_i)$ is the N-point correlation function for the bin with separation angle $\theta_i$, $\langle C_N(\theta_i) \rangle$ is the corresponding average from the MC simulation ensemble, and $N_{\text{bins}}$ is the number of bins used for the analysis. If $\hat{C}_N(\theta_i)$ is the i-th simulated N-point correlation function and $N_{\text{bins}}$ is the number of simulations, then the covariance matrix $M_{ij}$ is given by

$$M_{ij} = \frac{1}{N_{\text{bins}} - 1} \sum_{k=1}^{N_{\text{bins}}} \left( \hat{C}_N^{(k)}(\theta_i) - \langle C_N(\theta_i) \rangle \right) \left( \hat{C}_N^{(k)}(\theta_j) - \langle C_N(\theta_j) \rangle \right),$$

(8)

where $N_{\text{bins}} = N_{\text{sim}} - 1$. Following Hartlap et al. (2007), we then correct for bias in the inverse covariance matrix by multiplying it by the factor $(N_{\text{bins}} - N_{\text{sim}} - 1)/N_{\text{sim}}$. Below, we quote the significance level in terms of the fraction of simulations with a larger $\chi^2$ value than the observed map.

We analyse the CMB estimates at a resolution of $N_{\text{side}} = 64$, this being constrained by computational limitations. The results are presented in Fig. 2, where we compare the N-point functions for the data and the mean values estimated from 1000 MC simulations. The probabilities of obtaining values of the $\chi^2$ statistic for the Planck fiducial ΛCDM model at least as large as the observed values are given in Table 4.

It is worth noting that the values of the N-point functions for different angular separations are strongly correlated, and for this reason the figures show only one profile of each function in multi-dimensional space. Since the estimated probabilities take into account the correlations, they provide more reliable information on the goodness of fit between the data and a given model than simple inspection of the figures.

The results exhibit excellent consistency between the CMB maps estimated using the different component-separation methods. No statistically significant deviations of the CMB maps from Gaussianity are found. Indeed, the slight preference for super-Gaussianity of the equilateral 3-point and 4-point functions observed for the 2013 data is now less pronounced. That may be caused by differences between the masks used for the analysis. Interestingly, the 2-point function shows clear evidence of a lack of structure for large separation angles. Such behaviour was originally noted for the WMAP first-year data by Bennett et al. (2003), and has subsequently been discussed at length in the literature (Efstathiou 2004; Copi et al. 2007, 2015). We will return to this issue in Sect. 5.2.

### Table 3. Lower-tail probabilities for the N-pdf $\chi^2$ statistics derived from the Planck 2015 component-separated maps at $N_{\text{side}} = 16$ and 32.

<table>
<thead>
<tr>
<th>$N_{\text{side}}$</th>
<th>Comm. NILC SEVEN SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 . . . . . . . .</td>
<td>24.7 26.2 25.4 24.5</td>
</tr>
<tr>
<td>32 . . . . . . . .</td>
<td>11.9 20.8 10.6 10.8</td>
</tr>
</tbody>
</table>

### Table 4. Probabilities of obtaining values for the $\chi^2$ statistic of the N-point functions for the Planck fiducial ΛCDM model at least as large as the observed values of the statistic for the Planck 2015 temperature CMB maps with resolution parameter $N_{\text{side}} = 64$, estimated using the Commander, NILC, SEVEN, and SMICA methods.

<table>
<thead>
<tr>
<th>Function</th>
<th>Comm. NILC SEVEN SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-p. . . . . . .</td>
<td>97.2 98.9 97.4 98.1</td>
</tr>
<tr>
<td>Pseudo-coll. 3-p.</td>
<td>92.1 94.7 91.8 92.2</td>
</tr>
<tr>
<td>Equil. 3-p.</td>
<td>74.0 80.4 75.8 79.0</td>
</tr>
<tr>
<td>Rhombic 4-p.</td>
<td>64.6 70.9 65.6 65.9</td>
</tr>
</tbody>
</table>
4.4. Minkowski functionals

The Minkowski functionals (hereafter MFs) describe the morphology of fields in any dimension and have long been used as estimators of non-Gaussianity and anisotropy in the CMB (see e.g., Mecke et al. 1994; Schmalzing & Buchert 1997; Schmalzing & Gorski 1998; Komatsu et al. 2003; Eriksen et al. 2004b; Curto et al. 2007; De Troia et al. 2007; Spergel et al. 2007; Curto et al. 2008; Hikage et al. 2008; Komatsu et al. 2009; Planck Collaboration XXIII 2014). They are additive for disjoint regions of the sky and invariant under rotations and translations. In the literature, the contours are traditionally defined by a threshold \( \nu \), usually given in units of the sky standard deviation (\( \sigma_0 \)).

We compute MFs for the regions colder and hotter than a given threshold \( \nu \). Thus, the three MFs, namely the area \( V_0(\nu) = A(\nu) \), the perimeter \( V_1(\nu) = C(\nu) \), and the genus \( V_2(\nu) = G(\nu) \), are defined respectively as

\[
V_0(\nu) = A(\nu) = \frac{N_c}{N_{\text{pix}}},
\]

\[
V_1(\nu) = C(\nu) = \frac{1}{4A_{\text{tot}}} \sum S_i,
\]

\[
V_2(\nu) = G(\nu) = \frac{1}{2\pi A_{\text{tot}}} (N_{\text{hot}} - N_{\text{cold}}),
\]

where \( N_c \) is the number of pixels where \( \Delta T / \sigma_0 > \nu \), \( N_{\text{pix}} \) is the total number of available pixels, \( A_{\text{tot}} \) is the total area of the available sky, \( N_{\text{hot}} \) is the number of compact hot spots, \( N_{\text{cold}} \) is the number of compact cold spots, and \( S_i \) is the contour length of each hot spot.

For a Gaussian random field in pixel space, the MFs can be written in terms of two functions: \( A_k \), which depends only on the power spectrum, and \( v_k \), which is a function only of the threshold \( \nu \) (see, e.g., Vanmarcke 1983; Pogosyan et al. 2009; Gay et al. 2012; Matsubara 2010; Fantaye et al. 2015). The analytical expressions are

\[
V_k(\nu) = A_k v_k(\nu),
\]
find no significant deviations from Gaussianity for any of the presented in Table 5 for each component separation technique and et al. 2007, 2008). The corresponding covariance matrix is com-
each component-separation method by combining the 39 statis-
tical MFs – area, contour length, and genus – are evaluated over all of the resolutions considered. A χ2 in these statistics for the threshold range
−3.5 and 3.5 in σ units, except for θ = 640′ where 13 thresholds between −3.0 and 3.0 in σ units were more appropriate. Table 6 indicates that no significant deviation from Gaussianity is found.

Third, we tested MFs on needlet components. The needlet components of the CMB field as defined by Marinucci et al. (2008) and Baldi et al. (2009) are given by:

\[
β_{m} = \sum_{ℓ=0}^{B+1} b_{m}^{ℓ} T(ℓ) .
\]

Here, \(T(ℓ)\) denotes the component at multipole \(ℓ\) of the CMB map \(T(\hat{n})\), i.e.,

\[
T(\hat{n}) = \sum_{ℓ} T(ℓ) .
\]

Then we consider the four normalized functionalities described above. For every scale we used 26 thresholds ranging between −3.5 and 3.5 in σ units, except for θ = 640′ where 13 thresholds between −3.0 and 3.0 in σ units were more appropriate. Table 6 indicates that no significant deviation from Gaussianity is found.

<table>
<thead>
<tr>
<th>(N_{\text{side}})</th>
<th>Probability [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024 . . . . . .</td>
<td>91.4 90.7 95.5 95.8</td>
</tr>
<tr>
<td>512 . . . . . .</td>
<td>95.4 90.9 62.6 92.6</td>
</tr>
<tr>
<td>256 . . . . . .</td>
<td>55.8 34.5 55.9 55.9</td>
</tr>
<tr>
<td>128 . . . . . .</td>
<td>43.6 56.4 19.9 19.2</td>
</tr>
<tr>
<td>64 . . . . . .</td>
<td>59.3 37.8 22.7 80.0</td>
</tr>
<tr>
<td>32 . . . . . .</td>
<td>62.0 16.2 29.9 67.0</td>
</tr>
<tr>
<td>16 . . . . . .</td>
<td>43.4 45.8 47.7 31.0</td>
</tr>
</tbody>
</table>

with

\[
v_{2}(ν) = \exp(-ν^{2}/2)H_{k-1}(ν) , \quad k \leq 2 ,
\]

\[
v_{3}(ν) = \frac{e^{-ν^{2}}}{\text{erfc}(ν/\sqrt{2})} ,
\]

and

\[
H_{k}(ν) = e^{ν^{2}/2} \left( -\frac{d}{dν} \right)^{k} e^{-ν^{2}/2} .
\]

The amplitude \(A_{k}\) depends only on the shape of the power spectrum \(C_{ℓ}\) through the rms of the field \(σ_{0}\) and its first derivative \(σ_{1}\):

\[
A_{k} = \left( \frac{σ_{1}}{2σ_{0}} \right) k , \quad k \leq 2 ,
\]

\[
A_{3} = \frac{2}{π} \left( \frac{σ_{1}}{2σ_{0}} \right) ,
\]

where \(σ_{k} \equiv π^{k}/Γ(k/2 + 1)\).

Since this factorization is still valid in the weakly non-Gaussian case, we can use the normalized MFs, \(v_{k}\), to focus on deviations from Gaussianity, with a reduced sensitivity to cosmic variance.

Apart from the characterization of the MFs using full-resolution temperature sky maps, we also consider results at different angular scales. In this paper, two different approaches are considered to study these degrees of freedom: in real space via a standard Gaussian smoothing and degradation of the maps; and, for the first time, in harmonic space by using needlets. Such a complete investigation provides an insight regarding the harmonic and spatial nature of possible non-Gaussian features detected with the MFs.

First, we apply scale-dependent analyses in real space by considering the sky maps at different resolutions. The three classical MFs – area, contour length, and genus – are evaluated over the threshold range \(3 \leq ν \leq 3\) in σ units, with a step of 0.5. This provides a total of 39 different statistics. The values of these statistics for the Planck data are all within the 95% confidence region when compared with Gaussian simulations for all of the resolutions considered. A χ2 value is computed for each component-separation method by combining the 39 statistics and taking into account their correlations (see e.g., Curto et al. 2007, 2008). The corresponding covariance matrix is computed using 1000 simulations. The \(p\)-value of this χ2 test is presented in Table 5 for each component separation technique and for map resolutions between \(N_{\text{side}} = 1024\) and \(N_{\text{side}} = 16\). We find no significant deviations from Gaussianity for any of the resolutions considered.

<table>
<thead>
<tr>
<th>(N_{\text{side}})</th>
<th>Probability [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2048 . . . . . .</td>
<td>97.2 77.7 99.0 93.0</td>
</tr>
<tr>
<td>1024 . . . . . .</td>
<td>93.1 98.0 90.2 92.6</td>
</tr>
<tr>
<td>512 . . . . . .</td>
<td>53.7 36.7 30.4 77.6</td>
</tr>
<tr>
<td>256 . . . . . .</td>
<td>89.0 85.9 96.8 58.1</td>
</tr>
<tr>
<td>128 . . . . . .</td>
<td>93.0 63.5 94.1 37.1</td>
</tr>
<tr>
<td>64 . . . . . .</td>
<td>37.1 70.4 54.1 62.5</td>
</tr>
<tr>
<td>32 . . . . . .</td>
<td>28.9 77.4 75.5 46.7</td>
</tr>
<tr>
<td>16 . . . . . .</td>
<td>33.1 39.4 44.1 38.8</td>
</tr>
</tbody>
</table>

Then we consider the four normalized functionalities described above. For every scale we used 26 thresholds ranging between −3.5 and 3.5 in σ units, except for θ = 640′ where 13 thresholds between −3.0 and 3.0 in σ units were more appropriate. Table 6 indicates that no significant deviation from Gaussianity is found.

<table>
<thead>
<tr>
<th>(N_{\text{side}})</th>
<th>Probability [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024 . . . . . .</td>
<td>91.4 90.7 95.5 95.8</td>
</tr>
<tr>
<td>512 . . . . . .</td>
<td>95.4 90.9 62.6 92.6</td>
</tr>
<tr>
<td>256 . . . . . .</td>
<td>55.8 34.5 55.9 55.9</td>
</tr>
<tr>
<td>128 . . . . . .</td>
<td>43.6 56.4 19.9 19.2</td>
</tr>
<tr>
<td>64 . . . . . .</td>
<td>59.3 37.8 22.7 80.0</td>
</tr>
<tr>
<td>32 . . . . . .</td>
<td>62.0 16.2 29.9 67.0</td>
</tr>
<tr>
<td>16 . . . . . .</td>
<td>43.4 45.8 47.7 31.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(N_{\text{side}})</th>
<th>Probability [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2048 . . . . . .</td>
<td>97.2 77.7 99.0 93.0</td>
</tr>
<tr>
<td>1024 . . . . . .</td>
<td>93.1 98.0 90.2 92.6</td>
</tr>
<tr>
<td>512 . . . . . .</td>
<td>53.7 36.7 30.4 77.6</td>
</tr>
<tr>
<td>256 . . . . . .</td>
<td>89.0 85.9 96.8 58.1</td>
</tr>
<tr>
<td>128 . . . . . .</td>
<td>93.0 63.5 94.1 37.1</td>
</tr>
<tr>
<td>64 . . . . . .</td>
<td>37.1 70.4 54.1 62.5</td>
</tr>
<tr>
<td>32 . . . . . .</td>
<td>28.9 77.4 75.5 46.7</td>
</tr>
<tr>
<td>16 . . . . . .</td>
<td>33.1 39.4 44.1 38.8</td>
</tr>
</tbody>
</table>
Fig. 3. Needlet space MFs for Planck 2015 data using the four component-separated maps, Commander (red), NILC (orange), SEVEM (green), and SMICA (blue); the grey regions, from dark to light, correspond, respectively, to 1, 2, and 3σ confidence regions estimated from the 1000 FFP8 simulations processed by the Commander method. The columns from left to right correspond to the needlet parameters $j = 4, 6, 8$, respectively; the $j$th needlet parameter has compact support over multipole ranges $[2^{j-1}, 2^{j+1}]$. The $\ell_c = 2$ value indicates the central multipole of the corresponding needlet map. Note that to have the same range at all the needlet scales, the vertical axis has been multiplied by a factor that takes into account the steady decrease of the variance of the MFs as a function of scale.

Table 7. Probability $P(\chi^2 > \chi^2_{\text{Planck}})$ as a function of needlet scale.

<table>
<thead>
<tr>
<th>Needlet scale ($\ell$ range)</th>
<th>Commander</th>
<th>NILC</th>
<th>SEVEM</th>
<th>SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 (4,16)</td>
<td>32.1</td>
<td>36.1</td>
<td>40.4</td>
<td>39.8</td>
</tr>
<tr>
<td>4 (8,32)</td>
<td>84.0</td>
<td>57.9</td>
<td>79.4</td>
<td>59.4</td>
</tr>
<tr>
<td>5 (16,64)</td>
<td>23.8</td>
<td>11.2</td>
<td>29.1</td>
<td>43.8</td>
</tr>
<tr>
<td>6 (32,128)</td>
<td>14.8</td>
<td>38.9</td>
<td>33.5</td>
<td>34.1</td>
</tr>
<tr>
<td>7 (64,256)</td>
<td>11.9</td>
<td>7.5</td>
<td>15.4</td>
<td>1.1</td>
</tr>
<tr>
<td>8 (128,512)</td>
<td>46.1</td>
<td>55.2</td>
<td>67.7</td>
<td>52.2</td>
</tr>
</tbody>
</table>
these figures represent the data, while the histogram shows the results for the 1000 FFP8 simulations. We also show in Table 7 the $p$-values for the four component-separation methods, as well as all needle scales we considered. Despite the relatively small $p$-values for some scales, the Planck temperature maps show no significant deviation from the Gaussian simulations up to $\ell_{\text{max}} = 512$, which corresponds to the maximum multipole of our highest-frequency needlelet map.

4.5. Multiscale analyses

Multiscale data analysis is a powerful approach for probing the fundamental hypotheses of the isotropy and Gaussianity of the CMB. The exploration of different scales (in an almost independent manner) not only helps to test the specific predictions of a given scenario for the origin and evolution of the fluctuations, but also is an important check on the impact of systematic errors or other contaminants on the cosmological signal.

There are several ways of performing a multiscale analysis, the simplest being to smooth/upgrade the CMB map to different resolutions. However, in this section, we will focus on image processing techniques related to the application of wavelets and more general band-pass filtering kernels to the original CMB fluctuations. The advantage of wavelet-like analyses over scale degradation is clear: they allow the exploration of characteristics of the data that are related to specific angular scales. Wavelets have already been extensively used in the study of the Gaussianity and isotropy of the CMB (e.g., McEwen et al. 2007; Vielva 2007). Indeed, a wavelet-based (needlelet) analysis of the Planck 2015 data has already been presented in Sect. 4.4.

We recall that in the 2013 analysis, some of the applied estimators deviated from the null hypothesis. In particular, it was determined that the cold area of the spherical Mexican hat wavelet (SMHW, Martínez-González et al. 2002) coefficients at scales of around $5^\circ$ yielded a $p$-value of 0.3%. In addition, we also found an excess in the kurtosis of the wavelet coefficients on the same scales. Previous analyses (for a review, see Vielva 2010) have suggested that the “Cold Spot” (see Sect. 5.7) was the major contributor to these statistical outliers.

In what follows, we will consider the application of the SMHW, together with matched filters for a 2D-Gaussian profile (GAUSS), and for generalized spherical Savitzky-Golay kernels (SSG, Savitzky & Golay 1964, see Appendix A).

The application of a filter $\psi(R, p)$ to a signal on the sky $S(p)$ can be written as

$$\omega_f (R, p) = \sum_{\ell=0}^{\ell_{\text{max}}} \sum_{m=\ell-\ell}^{\ell-\ell} s_{lm} W^p_{\ell} (R) Y^p_{\ell m} (p),$$

where $p$ represents a given position/pixel, $R$ parameterizes a characteristic scale for the filter (e.g., a wavelet scale), $W^p_{\ell} (R)$ is the window function associated with the filter $\psi(R, p)$, $\ell_{\text{max}}$ is the maximum multipole allowed by the corresponding HEALPix pixelization, and $Y_{\ell m} (p)$ is the spherical harmonic basis. Here, $s_{lm}$, the spherical harmonic coefficients of the analysed map, are given by

$$s_{lm} = \int d\Omega Y^*_{\ell m} (p) S(p),$$

where $d\Omega = d\ell^2 \sin \theta d\theta d\phi$ and the asterisk denotes complex conjugation. Note that the filtered map (or the wavelet coefficient map, if $\psi(R, p)$ is a continous wavelet) conserves the statistical properties of the original map, since the convolution is a linear operation. In particular, if $S(p)$ is a Gaussian and statistically isotropic random signal, $\omega_f (R, p)$ is also Gaussian and statistically isotropic.

In the present work, the signal $S(p)$ corresponds to a temperature map $T(p)$. Several statistics can then be computed from the derived filtered map as a function of the filter scale, in particular, the first moments (the dispersion $\sigma_p$, the skewness $\psi_p$, and the kurtosis $K_p$), the total area above/below a given threshold, and the peak distribution. These statistics are compared to the corresponding results determined from the FFP8 simulations to establish the degree of compatibility with the null hypothesis.

4.5.1. First-order moments of the multiscale maps

For the three filters considered (SMHW, GAUSS, and SSG84) the variance, skewness, and kurtosis are computed at 18 scales, $R(\text{arcmin}) = [2, 4, 7, 14, 25, 50, 75, 100, 150, 200, 250, 300, 400, 500, 600, 750, 900, 1050]$. These scales are chosen to be consistent with previous analyses. They cover a wide angular range, and are selected so that the intervals between them increase with scale. Notice that, for a given scale, the three filters do not cover exactly the same multipole range, since that depends on the specific filter definition. This can be seen in

3 The digits 8 and 4 denote the order of the spherical Savitzky-Golay kernel and the smoothing weight, described in Appendix A.
diffuse Galactic emission (i.e., the auxiliary mask does not mask point sources).

The following figures represent the upper-tail probability (UTP) for a given statistic, i.e., the fraction of simulations that yield a value equal to or greater than that obtained for the data. In fact, as explained in PCIS13, if a given UTP is larger than 0.5, a new quantity is defined as \( \text{mUTP} = 1 - \text{UTP} \). Therefore, \( \text{mUTP} \) is constrained to lie between \( 1/N \) and 0.5, where \( N \) is the number of simulations used for each statistic.

Figure 6 presents the comparison of the CMB temperature maps with the corresponding simulations for the SMHW, GAUSS, and SSG84 filters. The full mission Planck data confirm the results already obtained with the 2013 release for temperature. In particular, for the SMHW, we find (i) an excess of kurtosis \((\approx 0.8\%)\) at scales of around 300’; (ii) that the dispersion of the wavelet coefficients at these scales and at around 700’ is relatively low \((\approx 1\%)\); and (iii) that the dispersion of the wavelet coefficients at scales below 5’ is significantly high \((\approx 0.1\%)\).

The excess of kurtosis has been previously associated with the Cold Spot (e.g., Vielva et al. 2004), and the low value of the standard deviation of the coefficients on large scales could be related to the low variance discussed in Sect. 5.1. Regarding the large dispersion of the coefficients on the smallest scales, this can be understood either by the presence of residual foreground contributions (extragalactic point sources) or by incomplete characterization of the true instrumental noise properties by the FFP8 simulations. We explore these possibilities with two additional tests undertaken with the SMHW.

Figure 7 shows the significance of the statistics derived from the SEVEN-100, SEVEN-143, and SEVEN-217 maps. The three cleaned maps yield very consistent values of the mUTP for the standard deviation, skewness, and kurtosis of the wavelet coefficients, with only small differences seen at small scales. This frequency-independence of the results argues against the foreground residuals hypothesis. Figure 8 presents the same statistics as applied to an estimator of the noise properties of the CMB maps. This is derived from the half-difference of the half-ring data sets, which provides the best estimate of the noise properties of the full mission data set. However, since there is still a known mismatch in noise properties, any conclusions will be more qualitative than quantitative. Nevertheless, the noise study reveals that, at the smallest scales, there are some discrepancies with the FFP8 simulations, and in particular the estimated dispersion of the SMHW noise coefficients is higher than predicted.

4.5.2. The area above/below a threshold

In the context of the study of the Cold Spot, the area above/below a given threshold, as a function of the SMHW wavelet scale, has been demonstrated to provide a useful and robust statistic (e.g., Cruz et al. 2005), since it is rather independent of any masking required. Our previous analysis (PCIS13) confirmed that the CMB temperature fluctuations exhibit an anomalously large cold area on scales of around 10’, which can be mostly associated with the Cold Spot. Here, we extend the analysis by including results derived using the GAUSS and SSG84 filters.

At a given scale \( R \) and threshold \( \nu \), the cold (\( A^{-\nu}_R \)) and hot (\( A^{+\nu}_R \)) areas of a filtered map are defined as

\[
A^{-\nu}_R \equiv \#(|\omega_k(R, \nu) < -\nu|), \quad (25)
\]

\[
A^{+\nu}_R \equiv \#(|\omega_k(R, \nu) < +\nu|), \quad (26)
\]

where the operator \# represents the number of pixels \( p \) in which the condition defined between the braces is satisfied.
Fig. 6. Modified upper tail probabilities (mUTP) obtained from the analyses of the filter coefficients as a function of the filter scale $R$ for the Commander (red), NILC (orange), SEVEM (green), and SMICA (blue) sky maps. From left to right, the panels correspond to standard deviation, skewness, and kurtosis results, when determined using the SMHW (top), GAUSS (middle), and SSG84 (bottom) filters. The squares represent UTP values above 0.5, whereas circles represent UTP values below 0.5.

Fig. 7. Modified upper tail probabilities (mUTP) obtained from the analyses of the SMHW coefficients as a function of the wavelet scale $R$ for the SEVEM-100 (blue), SEVEM-143 (yellow), SEVEM-217 (magenta), and SEVEM (green) maps. From left to right, the panels correspond to the standard deviation, skewness, and kurtosis.

Table 8 summarizes the results for the hot and cold areas determined for the four CMB temperature maps analysed with the common mask (and its associated exclusion masks). The results are similar to those obtained in 2013, with some small differences on those scales related to the Cold Spot (between 200' and 400'). Specifically, the cold area is slightly less significant for smaller values of $R$, whereas the anomalous behaviour remains for larger filter scales. The three filters are in...
reasonable agreement, but, as expected from Fig. 6, the SMHW yields higher significance levels than the SSG84 and GAUSS filters. However, it is worth recalling that, for a given scale, the three filters are not probing exactly the same multipole range and therefore some differences should be expected.

In Fig. 9 we plot the areas for thresholds $\nu > 3.0$, where the threshold is defined in units of $\sigma_R$, as determined from the SEVEM temperature map. The results for Commander, NILC, and SMICA are in good agreement with these. The panels refer to SMHW scales of $R = 200', 250', 300',$ and $400'$. The most extreme value (in terms of $\sigma_R$) for each area is indicated.

The coldest area corresponds to the Cold Spot with the minimum value of the wavelet coefficient at the position $(209^\circ, -57^\circ)$ in Galactic coordinates. The hottest area has already been identified in the WMAP data (e.g., Vielva et al. 2007). The results are insensitive to the choice of CMB temperature map that is adopted. It is clear that the southern Galactic hemisphere yields more anomalous signatures than the northern one. These results confirm the importance of the Cold Spot as the most extreme feature in the analysed sky. More insights about its nature are provided in Sect. 5.7.

4.5.3. Peak statistics

The statistical properties of local extrema (both minima and maxima, which we refer to collectively as ‘peaks’) provide an alternative approach to search for evidence of non-Gaussianity in the data. Such peaks, defined as pixels whose amplitudes are higher or lower than the corresponding values for all of their nearest neighbours, trace topological properties of the data. Peak locations and amplitudes, and various derived quantities, such as their correlation functions, have previously been used to characterize the WMAP sky maps by Larson & Wandelt (2004, 2005) and Hou et al. (2009).

The statistical properties of peaks for a statistically isotropic Gaussian random field were derived by Bond & Efstathiou (1987). The integrated number density of peaks, $n_{pk}$ (composed of maxima and minima with corresponding densities $n_{\text{max}}$ and $n_{\text{min}}$), with amplitudes $x$ above a certain threshold $\nu = x/\sigma$ is given by

$$\frac{n_{\text{max}} + n_{\text{min}}}{n_{pk}} \left( \frac{x}{\sigma} > \nu \right) = \sqrt{\frac{3}{2\pi}} \gamma^2 \exp \left( -\frac{\nu^2}{2} \right) + \frac{1}{2} \operatorname{erfc} \left( \frac{\nu}{\sqrt{2 - \frac{4}{3} \gamma^2}} \right),$$

where $\sigma$ is the rms fluctuation amplitude measured on the sky, and $\gamma$ is the spectral shape parameter of the underlying field. Uncharacteristically cold and hot spots are then manifested as extreme outliers in the peak values, and can constitute evidence for non-Gaussianity or deviation from isotropy.

### Table 8. Modified upper tail probability (mUTP) for the cold (top) and hot (bottom) areas.

<table>
<thead>
<tr>
<th>Area</th>
<th>Scale [arcmin]</th>
<th>Comm.</th>
<th>NILC</th>
<th>SEVEM</th>
<th>SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cold</td>
<td>200</td>
<td>3.8</td>
<td>5.1</td>
<td>3.7</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>1.4</td>
<td>2.4</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.4</td>
<td>1.5</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>2.0</td>
<td>2.6</td>
<td>1.7</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>4.2</td>
<td>5.0</td>
<td>4.1</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>GAUSS</td>
<td>200</td>
<td>1.7</td>
<td>2.7</td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>1.6</td>
<td>1.8</td>
<td>1.2</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>Cold</td>
<td>200</td>
<td>2.9</td>
<td>3.5</td>
<td>2.8</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>5.7</td>
<td>6.4</td>
<td>5.6</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>SSG84</td>
<td>200</td>
<td>9.4</td>
<td>11.0</td>
<td>9.4</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>12.3</td>
<td>13.4</td>
<td>10.8</td>
<td>12.3</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>1.4</td>
<td>2.6</td>
<td>1.4</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>0.9</td>
<td>1.9</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Cold</td>
<td>200</td>
<td>1.1</td>
<td>1.8</td>
<td>1.0</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>4.8</td>
<td>5.1</td>
<td>4.5</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

**Notes.** Results are given for the $\nu > 4\sigma_R$ threshold of the SMHW, GAUSS, and SSG84 coefficients. The four most significant scales related to the Cold Spot feature are shown. An ellipsis (…) indicates that no area above that threshold was found in the data.
order to avoid potential contamination by boundary effects, the mask is extended by rejecting pixels with an effective convolution weight that differs from unity by more than 12%. Peaks are extracted from the filtered map (removing any that are adjacent to masked pixels), their positions and values are recorded for further analysis, and their cumulative density function (CDF) is constructed by sorting peak values. Table 9 presents peak counts for the component-separated sky maps for several different kernels and representative filtering scales, together with the number of peaks that are common to all maps. There is excellent agreement between the various CMB estimates. All statistical inference is then performed by comparison of the peak distributions derived from the data with equivalently processed simulations. As an internal consistency check, the properties of the FFP8 simulations are found to be in agreement with the predictions of Eq. (27).

Figure 10 presents the distributions of peaks for the SMICA CMB map filtered with two representative kernels on scales of 40' and 800' FWHM. The lower panels show the empirical peak CDFs as a function of peak value $x$, defined for a set of $n$ peaks at values $\{X_i\}$ as

$$ F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{X_i \leq x}, \quad I_{X_i \leq x} = \begin{cases} 1, & \text{if } X_i \leq x \\ 0, & \text{otherwise} \end{cases} $$

For plotting purposes alone, the horizontal axis is scaled in units of $\sigma$ defined by Eq. (27) and derived from the underlying median CDF, $F(x)$, of the simulations. The upper panels show the difference between the observed and median simulated CDF values, $\sqrt{n} [F_n(x) - F(x)]$, with the grey bands representing the 68.3%, 95.4%, and 99.7% regions of the simulated CDF distributions. The maximal value of this difference defines a Kolmogorov-Smirnov (KS) deviation estimator:

$$ K_n \equiv \sqrt{n} \sup_x |F_n(x) - F(x)|. $$

This forms the basis of a standard KS test of consistency between the two distributions. Although the KS deviation has a known limiting distribution, we also derive its CDF directly from the simulations.

<table>
<thead>
<tr>
<th>Filter scale [arcmin]</th>
<th>SMHW</th>
<th>NILC</th>
<th>SEVEM</th>
<th>SMICA</th>
<th>Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>176/187</td>
<td>170/178</td>
<td>173/182</td>
<td>169/182</td>
<td>161/169</td>
</tr>
<tr>
<td>250</td>
<td>105/105</td>
<td>104/103</td>
<td>107/123</td>
<td>105/107</td>
<td>97/99</td>
</tr>
<tr>
<td>300</td>
<td>70/70</td>
<td>71/70</td>
<td>70/70</td>
<td>72/68</td>
<td>71/66</td>
</tr>
<tr>
<td>400</td>
<td>43/32</td>
<td>46/32</td>
<td>44/31</td>
<td>43/33</td>
<td>37/30</td>
</tr>
<tr>
<td>200</td>
<td>152/170</td>
<td>152/166</td>
<td>157/179</td>
<td>156/165</td>
<td>142/155</td>
</tr>
<tr>
<td>250</td>
<td>102/93</td>
<td>94/95</td>
<td>108/99</td>
<td>99/101</td>
<td>92/85</td>
</tr>
<tr>
<td>300</td>
<td>60/63</td>
<td>57/62</td>
<td>63/64</td>
<td>56/62</td>
<td>50/53</td>
</tr>
<tr>
<td>400</td>
<td>33/28</td>
<td>29/29</td>
<td>31/33</td>
<td>29/28</td>
<td>24/27</td>
</tr>
<tr>
<td>200</td>
<td>180/187</td>
<td>178/183</td>
<td>180/185</td>
<td>183/183</td>
<td>167/175</td>
</tr>
<tr>
<td>250</td>
<td>131/119</td>
<td>118/114</td>
<td>122/123</td>
<td>121/110</td>
<td>109/103</td>
</tr>
<tr>
<td>300</td>
<td>68/69</td>
<td>73/68</td>
<td>73/73</td>
<td>73/68</td>
<td>56/61</td>
</tr>
<tr>
<td>400</td>
<td>29/35</td>
<td>29/36</td>
<td>29/32</td>
<td>30/38</td>
<td>27/27</td>
</tr>
</tbody>
</table>

Table 9. Peak counts in maps filtered to different scales.
The temperature peak distributions in Fig. 10 are consistent with Gaussian peak statistics, apart from a single anomalously cold peak on scales around 800′ FWHM. This corresponds to the previously reported Cold Spot. Although this exercise confirms that the Cold Spot is a rare cold feature, as already noted by Cruz et al. (2005) and confirmed in this paper, the most peculiar characteristic of the Cold Spot is not its coldness, but rather its size. A more detailed analysis of its nature is presented in Sect. 5.7.

The probability that the observed sky exceeds the value of the KS deviation for the adopted fiducial cosmology can be determined by counting the number of simulations with $K_{\omega,\text{sky}} < K_{\omega,\text{sky}}^\text{obs}$. The $p$-values for the KS test comparing the CDF of the observed sky with the median peak CDF derived from simulations for several different kernels and representative scales are summarized in Table 10. The similarly derived $p$-values for the total peak counts are summarized in Table 11. Most of the results indicate that the two distributions are highly consistent, with the exception of results for the SSG84 filter on scales of about 500′ FWHM. This deviation appears to be related to a hemispherical asymmetry in the peak CDFs, and will be discussed further in Sect. 5.6.

One can also test whether the observed values of the parameters, $\sigma$ and $\gamma$ as defined in Eq. (27), are consistent with the simulation ensemble, under the assumption that the peak distributions in the Planck data are described by a Gaussian peak CDF. Figure 11 demonstrates the consistency of the best-fit values of these parameters, corresponding to the peak distributions in Fig. 10, with equivalent values derived from the simulations.

Inspired by the analysis of the WMAP first-year data in Larson & Wandelt (2004) which found fewer extreme peaks than expected, we additionally evaluate whether the distributions of maxima and minima are separately consistent with simulations. The mean of all maxima, and the negative of the mean of all minima, are calculated for the filtered map, and the observed

Fig. 10. Cumulative density function of the peak distribution for the SMICA CMB temperature map. The top row shows the peak CDF for maps filtered with a GAUSS kernel of 40′ FWHM. The bottom row shows the corresponding peak CDF for an SSG84 kernel of 800′ FWHM. The spectral shape parameter $\gamma$ (see Eq. (27)) is the best-fit value for the simulated ensemble, as indicated by the cyan circle in Fig. 11. Similar results are obtained for the other component-separation methods.

Fig. 11. Distribution of best-fit Gaussian peak CDF spectral shape parameters, $\sigma$ and $\gamma$ (as defined in Eq. (27)), recovered from 1000 simulations, as indicated by the black dots and the smoothed density map and compared to those derived for the observed sky (shown by the red star). The blue contours enclose 68% and 95% of the parameter distribution, and the cyan circle represents the best-fit parameters for the median peak CDF determined from simulations. The upper panel shows the peak CDF parameters for the SMICA map filtered with a GAUSS kernel of 40′ FWHM. The lower panel shows the corresponding peak CDF for an SSG84 kernel of 800′ FWHM. Similar results are obtained for the other component-separation methods.
Corresponding peak CDF for an SSG84 upper panel. The observed minima are compared to the simulated distributions in Fig. 12. The error bars represent the sampling uncertainty due to the finite number of realizations, and are determined using a bootstrap method. As the filters overlap substantially, different points are highly correlated. The Planck CMB maps are consistent with the expectations of a statistically isotropic Gaussian model. The most significant deviation, found at an effective filter bandwidth given by $\ell = 20$, is attributable to a single region on the sky – the Cold Spot.

**Table 11.** Modified upper tail probability (mUTP) for the total peak count, comparing the data with the peak count CDF derived from simulations.

<table>
<thead>
<tr>
<th>Filter scale [arcmin]</th>
<th>Comm. NILC SEVEM SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMHW</td>
<td></td>
</tr>
<tr>
<td>200 000</td>
<td>6.1 36.9 16.2 27.2</td>
</tr>
<tr>
<td>250 000</td>
<td>32.9 47.5 1.0 25.6</td>
</tr>
<tr>
<td>300 000</td>
<td>48.8 51.7 44.7 44.3</td>
</tr>
<tr>
<td>400 000</td>
<td>33.8 16.2 34.6 26.4</td>
</tr>
<tr>
<td>GAUSS</td>
<td></td>
</tr>
<tr>
<td>200 000</td>
<td>7.1 11.2 0.7 8.7</td>
</tr>
<tr>
<td>250 000</td>
<td>18.2 11.2 2.1 8.1</td>
</tr>
<tr>
<td>300 000</td>
<td>29.0 12.8 48.2 10.0</td>
</tr>
<tr>
<td>400 000</td>
<td>11.9 3.0 26.6 2.8</td>
</tr>
<tr>
<td>SSG84</td>
<td></td>
</tr>
<tr>
<td>200 000</td>
<td>0.2 3.0 1.0 1.7</td>
</tr>
<tr>
<td>250 000</td>
<td>0.1 1.7 0.1 2.1</td>
</tr>
<tr>
<td>300 000</td>
<td>9.3 22.6 50.7 12.2</td>
</tr>
<tr>
<td>400 000</td>
<td>0.3 0.4 0.1 2.3</td>
</tr>
</tbody>
</table>

**Table 10.** Modified upper tail probability (mUTP) for the KS test, comparing the data with the median peak CDF derived from simulations.

<table>
<thead>
<tr>
<th>Filter scale [arcmin]</th>
<th>Probability [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMHW</td>
<td></td>
</tr>
<tr>
<td>200 000</td>
<td>22.0 42.8 45.9 40.5</td>
</tr>
<tr>
<td>250 000</td>
<td>11.3 17.6 3.1 11.4</td>
</tr>
<tr>
<td>300 000</td>
<td>49.4 38.5 38.4 32.1</td>
</tr>
<tr>
<td>400 000</td>
<td>32.6 24.7 35.3 24.7</td>
</tr>
<tr>
<td>GAUSS</td>
<td></td>
</tr>
<tr>
<td>200 000</td>
<td>41.3 46.6 14.4 47.2</td>
</tr>
<tr>
<td>250 000</td>
<td>43.7 34.8 7.6 48.4</td>
</tr>
<tr>
<td>300 000</td>
<td>46.3 9.9 28.0 7.7</td>
</tr>
<tr>
<td>400 000</td>
<td>30.7 5.6 35.8 6.6</td>
</tr>
<tr>
<td>SSG84</td>
<td></td>
</tr>
<tr>
<td>200 000</td>
<td>37.1 36.7 24.0 37.5</td>
</tr>
<tr>
<td>250 000</td>
<td>0.5 1.7 0.8 5.4</td>
</tr>
<tr>
<td>300 000</td>
<td>17.5 12.2 0.3 9.3</td>
</tr>
<tr>
<td>400 000</td>
<td>47.4 44.6 47.5 47.8</td>
</tr>
</tbody>
</table>

**Fig. 12.** Cumulative density function of the mean amplitude of all extrema, maxima (red) and minima (blue), derived from simulations, compared to the equivalent values observed for the SMICA CMB temperature map. The upper panel shows the peak mean amplitudes for maps filtered with a GAUSS kernel of 400 FWHM. The lower panel shows the corresponding peak CDF for an SSG84 kernel of 800 FWHM. Similar results are obtained for the other component separation methods. Since the filter kernel normalization is free, and the pre-whitened map to which the filter is applied is dimensionless, the plots are essentially in arbitrary units.

The probability that the coldest peak seen on the sky is consistent with the adopted fiducial cosmology is evaluated as a function of both filter shape and size by counting the number of simulations with $x_{\text{coldest}} < x_{\text{coldest}}(\text{sky})$. The results obtained for the SMHW filter are summarized in Fig. 13. Consistent behaviour is seen when the GAUSS and SSG84 filters are applied. The error bars represent the sampling uncertainty due to the finite number of realizations, and are determined using a bootstrap method. As the filters overlap substantially, different points are highly correlated. The Planck CMB maps are consistent with the expectations of a statistically isotropic Gaussian model. The most significant deviation, found at an effective filter bandwidth given by $\ell = 20$, is attributable to a single region on the sky – the Cold Spot.

**4.5.4. Peak locations as a function of scale**

The application of a filter kernel of variable size to a map extends it into what can be considered a “multiscale space”, such that features on different scales are represented by a one-parameter family of smoothed maps. This technique is often used for feature detection and mathematical morphology analysis. Here, we introduce a morphological description of temperature maps based on the peak connectedness graph in multiscale space, and apply this technique to a statistical analysis of the Planck CMB data. Like most morphological analyses, it is equally applicable to intrinsically non-Gaussian maps, but here we focus on the Gaussian random field statistics and attempt to understand what features of the CMB temperature map are responsible for the Cold Spot.

To construct a multiscale representation, we trace the location of the peaks in the smoothed, whitened CMB map as the smoothing scale is varied. As the smoothing scale increases, peaks merge and the total peak count decreases. Linking closest peak neighbours in position-scale space, from the finest to the coarsest resolution, produces an acyclic graph that encapsulates the peak “merger tree” history as the scale is varied. A summary of all the peak positions and CDF ranks for the SSG84 filter kernel on scales ranging from 1200 to 12000 FWHM is shown in Fig. 14. The peaks are represented by discs of varying size (reflecting the filter scale) and colour (reflecting the peak temperature rank), with peaks at all scales projected onto a single map. The lower panel shows the peak linkage graph on the coarser scales; for the statistical analysis 81 filter scales are used, log-spaced from 1200 to 12000. Peaks of the same type (i.e., maxima to maxima and minima to minima) are linked to...
Fig. 13. Fraction of the Gaussian random field realizations in which the coldest peak is as cold as or colder than that observed, as a function of SMHW filter scale for Commander (red), NILC (orange), SEVEM (green), and SMICA (blue).

Fig. 14. Peak positions and CDF rank summarized for all filtering scales. The three sky-view panels in the top row show a Lambert projection of the north pole, the usual full sky Mollweide projection, and a Lambert projection of the south pole. The lower panel shows the peak heights (in percentile of the peak distribution on the horizontal axis) as a function of filter scale (on the vertical axis, in logarithmic scale), truncated to larger scales for clarity. Circles represent peaks (nodes of the graph) coloured according to their percentile level, and scaled according to kernel size. Black lines represent edges connecting peaks at different scales (according to a minimal distance measure). The components connected to the coldest and hottest peaks at any scale are highlighted by thicker edges, and are navy blue and dark red in colour. Note that there are thick lines that do not touch the 0 and 1 percentiles in the plot view. Those edges are connected to extreme percentile values, but at scales smaller than those shown in the plot. The Cold Spot is represented by the connected nodes that have the smallest percentiles except for the coarsest scale in the plot view.

the closest peak on the coarser scale according to a distance measure, $d^2 + d^2$, where $d^2$ is the metric on the unit sphere, and $df^2$ is the difference of peak temperature ranks (but only if that distance is within a predetermined fraction of the filter scale FWHM).

The resulting peak linkage graph is then analysed for connectedness. The simplest quantifiable measure is the node-degree distribution, shown in Fig. 15 for SMICA. The node-degree distribution is highly peaked at 2; this population corresponds to a single peak being traced across multiple scales. Pre-whitening effectively decorrelates the Gaussian map across different scales, so that the resulting node distribution shows a sizeable population of degree 0 and 1 nodes. When compared to the linkage graphs derived from the simulation ensemble, the
node-degree distribution of the peak linkage graph derived from Planck CMB data is consistent, with a slight excess in node counts of degrees 5 and 6.

5. Anomalies in the microwave sky

The previous section established the lack of evidence for significant non-Gaussianity in the Planck data. Here we consider several important anomalies that were originally detected in the WMAP sky maps, and later confirmed in the analyses described in PCIS13. Many of these are connected to evidence for a violation of isotropy, or to a preferred direction, in the CMB. Tests that involve dipolar power asymmetry, either directly or via measures of directionality, are collected together in Sect. 6. In this section we consider only those anomalies not directly related to dipolar power asymmetry.

The microwave sky is intrinsically statistically anisotropic due to our motion with respect to the CMB rest frame. The resulting Doppler boosting effect, introduced in Sect. 1, was detected in the 2013 Planck data (Planck Collaboration XXVII 2014). For completeness, Appendix B repeats the analysis with the Planck full mission data set, though based only on the full velocity estimator ($\hat{\beta}$), which is the sum of the modulation and the aberration contributions. However, since the effects of Doppler boosting are now included in the simulations used for that analysis, this constitutes a consistency check for this release. More importantly, since both the data and simulations now include the cosmological parameter fits are relatively insensitive to these low-order $\ell$-modes, and are instead largely dominated by scales corresponding to $\ell > 50$. Therefore the variance of the map appears to be anomalous, since there is a dearth of large-angular-scale power compared to the model.

Before presenting our results, we return to the issue of a posteriori correction, which particle physicists refer to as correcting for the “look-elsewhere effect” (LEE). Since there are many tests that can be performed on the data to look for a violation of statistical isotropy, we expect some to indicate detections at, for example, roughly $3\sigma$ levels, since even a statistically isotropic CMB sky is a realization of an underlying statistical process corresponding to many independent random variables. However, in the absence of an existing theoretical framework (i.e., a physical model) to predict such anomalies, it is difficult to interpret their significance. It is then necessary, and equally challenging, to address the question of how often such detections would be found for statistically isotropic Gaussian skies. Unfortunately, it is not always clear how to answer this question.

There will always be a degree of subjectivity when deciding exactly how to assess the significance of these types of features in the data. As an example, one could argue that the large-scale dipole modulation signal we see is coming specifically from super-Hubble modes, in which case performing an LEE correction for dipole modulation that could have been seen on small scales ($\ell \geq 100$) would not make sense. Models for such a super-Hubble modulation exist and an example was examined in Planck Collaboration XX (2016), the conclusion being that the model could only explain part of the dipole modulation and that the allowed part was perfectly consistent with cosmic variance.

In this paper, we adopt a pragmatic approach. When there is a clear mechanism for doing so, we attempt to correct for the “multiplicity of tests”, or the possible ways in which an anomalous signal might have been detected but was not, as a consequence of any a posteriori (data-driven) choices made in searching for it. In such cases, a strong dependence of the significance on the correction would indicate that we should be cautious about the uncorrected result. When such an obvious correction is not possible, we clearly describe the methodology applied to the data and its limitations. With this approach, we also recognize that any statistical assessment is partially subjective, including those that purport to correct for the LEE.

Although many of the observed effects described in this and the next section may elude theoretical prediction today, we continue to highlight them since there is a real possibility that the significance of one or more might increase at a later date, perhaps when polarization data are included in the analysis, and lead to new insights into early Universe physics. Alternatively, such observations may directly motivate the construction of models that can make predictions for features that can be sought in new data sets. This is particularly the case for anomalies on the largest angular scales, which may have a specific connection to inflation.

5.1. Variance, skewness, kurtosis

Previous analyses of the WMAP data (Monteserin et al. 2008; Cruz et al. 2011; Gruppuso et al. 2013) have reported that the variance of the CMB sky is lower than that of simulations based on the $\Lambda$CDM model. PCIS13 confirmed this, and proposed a possible explanation of the apparent incompatibility of the observed variance with a fiducial cosmological model that has been determined from the same data set. Specifically, whilst the map-based variance is dominated by contributions from large angular scales on the sky, the cosmological parameter fits are relatively insensitive to these low-order $\ell$-modes, and are instead largely dominated by scales corresponding to $\ell > 50$. Therefore the variance of the map appears to be anomalous, since there is a dearth of large-angular-scale power compared to the model.

In Sect. 4.1, we again confirmed the presence of low variance in the data. Here, we extend the analysis to investigate which angular scales are responsible for the low variance by applying the
unit variance estimator to lower resolution component-separated maps, specifically those from $N_{\text{side}} = 1024$ to $N_{\text{side}} = 16$, with the corresponding common masks, and then comparing the results with those determined from 1000 MC simulations. The results are shown in Fig. 16.

All of the component-separation methods that we consider yield very consistent results which indicate an increasingly anomalous low variance at lower resolutions, with the lower-tail probability reaching a minimum value of 0.5% at $N_{\text{side}} = 16$. We then consider the impact of a possible look-elsewhere effect by evaluating the minimum tail probability of each simulation irrespective of the $N_{\text{side}}$ resolution at which it occurs. By comparing the distribution of these values with that of the data, we infer that the probability is slightly weakened to a value of about 1%. These results are compatible with a lack of power on large angular scales. Since the variance estimator is heavily weighted towards low-$\ell$ modes, this has an increasing impact on the observed variance when going from high to low resolution sky maps. Conversely, the skewness and kurtosis are consistent with the simulations, although there is some indication of a weak scale-dependence (albeit at low significance).

We also investigate the stability of the results at $N_{\text{side}} = 16$ with respect to the possible presence of residual foregrounds by considering two additional masks obtained by extending the edge of the $N_{\text{side}} = 16$ common mask by 5° and 9°, reducing the usable sky fraction from 58% to 48% and 40%, respectively. We then re-apply the unit variance estimator to the low resolution component-separated maps with these masks and determine the variance, skewness, and kurtosis values (see Table 12).

The results from 48% of the sky reveal that only 1 simulation in 1000 is found to be more anomalous (i.e., exhibit lower variance) than the observed map. In addition, both the skewness and kurtosis become more compatible with the ΛCDM model. With the more aggressive mask, the lower-tail probability slightly increases again. However, given the limited number of pixels involved in the analysis, this shift may be related to the effects of sample variance.

Overall, our results may be explained by the presence of a low-variance anomaly in the primordial CMB signal – the stability of the low-variance significance argues against foreground contamination being responsible for the lack of observed power. This is reinforced by the decrease in variance when regions close to the common mask borders, where foreground residuals are most likely to be observed, are omitted from the analysis.

### 5.2. N-point correlation function anomalies

#### 5.2.1. Lack of large-angle correlations

We first reassess the lack of correlation seen in the 2-point angular correlation function at large angular separations as reported in Sect. 4.3, and previously noted for both WMAP and the 2013 Planck temperature maps (Bennett et al. 2003; Copi et al. 2015). We attempt to quantify this lack of structure using the statistic proposed by Spergel et al. (2003):

$$S(x) = \int_{-1}^{x} \left[ \hat{C}_2(\theta) \right]^2 d(\cos \theta) ,$$

(30)

where $\hat{C}_2(\theta)$ is our estimate of the 2-point correlation function. Generally, the upper limit on the integral has been taken to correspond to a separation angle of $60^\circ$, possibly (as noted by Copi et al. 2009) motivated by the COBE-DMR 4-year results (Hinshaw et al. 1996). Inspection of the top panel of Fig. 2 suggests that the Planck 2-point function lies close to zero between $80^\circ$ and $170^\circ$, but for consistency with previous work we compute the statistic $S_{1/2}$, for $x \equiv \cos 60^\circ = \frac{1}{2}$. The results are presented in Table 13. We find that the data indeed show a lack of correlations on large angular scales, with a significance consistent with that found by Copi et al. (2015) (although note that the sense of the $p$-values differs between the papers).

Possible criticisms of the $S_{1/2}$ statistic include that it has been designed a posteriori to test for a lack of large-angle correlations, and that it does not account for the high degree of correlation between bins at different angular scales. We can address these concerns, at least in part, by considering a modified version of the commonly used and well understood $\chi^2$ statistic used in previous studies. In order to test the same hypothesis as the $S_{1/2}$ statistic – that there are no correlations on scales larger than some angular cut-off – we do not subtract an averaged 2-point...
The served 2-point function both to zero and to the predictions of full separation angle range. This is even more pronounced for an analysis in the ΛCDM model. Moreover, for θ > 60° the results suggest that the separation angle that delineates the interesting region of behaviour of the correlation function. We therefore consider the generalized statistic S(θ) and compute its value for all values of θ, both for the data and for the simulations. Then, for each value of θ, we determine the number of simulations with a higher value of S(θ), and hence infer the most significant value of the statistic and the separation angle that it corresponds to. However, since such an analysis is sensitive to the LEE, we define a global statistic to evaluate the true significance of the result. Specifically, we repeat the procedure for each simulation and search for the largest probability irrespective of the value of θ which it occurs. The fraction of these probabilities higher than the maximum probability found for the data defines a global p-value. As seen in Table 13, this corresponds to values of order 98% for all of the CMB estimates.

The previous analyses essentially test how consistent the observed 2-point correlation function data is with a lack of correlations on large angular scales, in particular for separation angles θ > 60°. A conventional χ² statistic allows us to test the consistency of this quantity with the predictions of the ΛCDM model. In this case, the statistic is defined as in Eq. (7), except that we constrain the computations to those bins that correspond to the intervals defined by θ < 60° and θ > 60°. The results of these studies are shown in Table 14.

The analysis for θ < 60° indicates that the observed 2-point function is a good match to the mean 2-point function predicted by the ΛCDM model. Moreover, for θ > 60° the results suggest that the problem is that the fit of the data to the model is too good, and this is even more pronounced for an analysis in the full separation angle range.

Overall, the tests indicate an unusually good fit of the observed 2-point function both to zero and to the predictions of the ΛCDM model for angles above 60°. This problem may be related to the fact that the theoretical variance for the best-fit model is larger than the observed value at large scales, so that the simulations based on this model that have been used in all of the statistical tests may overestimate the variance of the 2-point function.

### Table 13. Probabilities of obtaining values for the S₁/₂ and χ² statistics for the Planck fiducial ΛCDM model at least as large as the observed values of the statistic for the Planck 2015 temperature CMB maps with resolution parameter Nside = 64, estimated using the Commander, NILC, SEVEM, and SMICA maps.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Comm.</th>
<th>NILC</th>
<th>SEVEM</th>
<th>SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>S₁/₂</td>
<td>99.5</td>
<td>99.6</td>
<td>99.5</td>
<td>99.6</td>
</tr>
<tr>
<td>S(θ) (global)</td>
<td>97.7</td>
<td>97.8</td>
<td>97.8</td>
<td>97.9</td>
</tr>
<tr>
<td>χ²(θ &gt; 60°)</td>
<td>98.1</td>
<td>98.8</td>
<td>98.1</td>
<td>98.4</td>
</tr>
</tbody>
</table>

### Table 14. Probabilities of obtaining values for the χ² statistic for the Planck fiducial ΛCDM model at least as large as the observed values of the statistic for the Planck 2015 temperature CMB maps with resolution parameter Nside = 64, estimated using the Commander, NILC, SEVEM, and SMICA maps.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Comm.</th>
<th>NILC</th>
<th>SEVEM</th>
<th>SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>χ²(θ &lt; 60°)</td>
<td>91.5</td>
<td>93.3</td>
<td>91.6</td>
<td>91.7</td>
</tr>
<tr>
<td>χ²(θ &gt; 60°)</td>
<td>96.8</td>
<td>98.3</td>
<td>96.9</td>
<td>98.1</td>
</tr>
</tbody>
</table>

### 5.2.2. Hemispherical asymmetry

We now turn to a reassessment of the asymmetry between the real-space N-point correlation functions computed on hemispheres reported previously for the WMAP (Eriksen et al. 2005) and Planck 2013 temperature maps (PCIS13). We initially focus on the analysis on the hemispheres determined in the ecliptic coordinate frame for which the largest asymmetry was observed. However, we also carry out the corresponding calculations in other relevant reference frames, such as those defined by the Doppler boost (DB, see Sect. 6.4, Appendix B, and Planck Collaboration XXVII 2014) and the dipole modulation (DM, see Sect. 6.2) directions. We use the same configurations of the N-point functions as described in Sect. 4.3. However, here the functions are not averaged over the full sky and depend on a choice of specific direction, so they constitute tools for studying statistical isotropy rather than non-Gaussianity (Ferreira & Magueijo 1997).

As in Sect. 4.3, we analyse the CMB estimates at a resolution of Nside = 64 and quantify their agreement with the fiducial cosmological model using a χ² statistic. The results determined from the Planck 2015 temperature data for the ecliptic hemispheres are shown in Fig. 17. If we consider that the χ² statistic itself can act as a measure of fluctuation level, then asymmetry between the two measured hemispheres can be quantified by the ratio of the corresponding χ² values. The probabilities of obtaining values of the χ² statistic or ratio for the Planck fiducial ΛCDM model at least as large as the observed values are given in Table 15. Since we do not have any predictions concerning the behaviour of a given hemisphere, in the case of the χ² ratios we provide the complementary probabilities of the 2-tailed statistic.

The significance levels of the 3- and 4-point functions in the northern hemisphere are nominally very high, exceeding 99.9% for the pseudo-collapsed 3-point function. However, proper interpretation requires that one recognize that the analysis is affected by posteriori choices for the smoothing scale and reference frame defining the hemispheres. This typically leads to an overestimation of the significance of the results. Accounting for such effects requires the repetition of the analysis for all possible reference directions and also for data at other resolutions. Unfortunately, because of computational limitations, such an extended analysis is not possible for these higher-order statistics. Nevertheless, the observed properties of the Planck data are consistent with a clear lack of fluctuations in a direction towards the north ecliptic pole. However, the χ²-ratio statistic indicates...
a slightly smaller significance level for the asymmetry, not exceeding 99% for any of the N-point functions.

The results for the N-point correlation functions determined in the DB and DM reference frames for the SMICA map are shown in Fig. 18 and the probabilities are presented in Table 16. Note that the positive hemisphere for the ecliptic reference frame corresponds to the southern hemisphere in the previous table. Whilst the largest asymmetry is seen in ecliptic coordinates, a substantial asymmetry is present also for the DM direction. This can be explained by the fact that the DM direction is more closely aligned with the south ecliptic pole (with a separation of around 47°) than the DB direction is. For the DB direction we do not find any significant asymmetry. The equivalent results for Commander, NILC, and SEVEM are consistent with those shown here.

In conclusion, the correlation functions for the Planck 2015 temperature data are consistent with the results presented in PCIS13. Specifically, we observe that the northern hemisphere correlation functions are relatively featureless (both the 3- and 4-point functions lie very close to zero), whereas the southern hemisphere functions exhibit a level of structure consistent with Gaussian simulations.

5.3. Constraints on quadrupolar modulation

The most natural extension of the class of statistically anisotropic models that we have considered previously involves the quadrupolar modulation of an initially statistically isotropic CMB sky map. No detection of a corresponding quadrupolar power asymmetry is currently claimed. An initial BipoSH analysis of the WMAP 7-year data (Bennett et al. 2011) found evidence of corresponding non-zero spectra, $A_{l0}^2$ and $A_{l2}^2$, in ecliptic coordinates. However, Hanson et al. (2010) demonstrated that the signal could be attributed to an incomplete treatment of beam asymmetries in the data, and this was subsequently confirmed in Bennett et al. (2013). The corresponding analysis of the Planck 2013 data indicated consistency with statistical isotropy (Planck Collaboration XXIII 2014).

<table>
<thead>
<tr>
<th>Hemisphere</th>
<th>2-point function</th>
<th>3-point function</th>
<th>4-point function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern</td>
<td>89.7 90.6 89.8 88.0</td>
<td>&gt;99.9 &gt;99.9 &gt;99.9 &gt;99.7</td>
<td>99.7 99.7 99.7 99.6</td>
</tr>
<tr>
<td>Southern</td>
<td>80.5 82.7 82.9 77.6</td>
<td>35.1 34.9 35.8 31.4</td>
<td>38.6 38.7 38.6 38.6</td>
</tr>
<tr>
<td>Northern</td>
<td>98.8 98.5 98.5 98.4</td>
<td>98.6 98.6 98.8 98.4</td>
<td>97.3 97.1 97.2 97.0</td>
</tr>
<tr>
<td>Southern</td>
<td>86.6 86.7 86.6 86.7</td>
<td>45.7 45.7 47.8 42.6</td>
<td>22.8 22.5 23.2 20.1</td>
</tr>
</tbody>
</table>

| Table 15. Probabilities of obtaining values for the $\chi^2$ statistic and ratio of $\chi^2$ of the N-point functions shown in Fig. 17 for the Planck fiducial ΛCDM model at least as large as the observed values of the statistic for the Planck 2015 CMB maps estimated on northern and southern ecliptic hemispheres. |
|-----------------|-----------------|-----------------|
| Probability [%] | Comm. NILC SEVEM SMICA |
| Northern 2-point function | 89.7 90.6 89.8 88.0 |
| Southern 2-point function | 80.5 82.7 82.9 77.6 |
| Northern 3-point function | >99.9 >99.9 >99.9 >99.7 |
| Southern 3-point function | 35.1 34.9 35.8 31.4 |
| Northern 4-point function | 98.8 98.5 98.5 98.4 |
| Southern 4-point function | 98.6 98.6 98.8 98.4 |
| Northern 5-point function | 97.3 97.1 97.2 97.0 |
| Southern 5-point function | 97.0 57.9 92.5 |

Table 16. Probabilities of obtaining values for the $\chi^2$ statistic and ratio of $\chi^2$ of the N-point functions shown in Fig. 18 for the Planck fiducial ΛCDM model at least as large as the observed values of the statistic for the SMICA map on hemispheres defined by the ecliptic (first column), Doppler boost (DB, second column), and dipolar modulation (DM, third column) reference frames.

<table>
<thead>
<tr>
<th>Hemisphere</th>
<th>2-point function</th>
<th>Ecl.</th>
<th>DB</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern</td>
<td>88.0 86.9 61.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern</td>
<td>77.6 91.1 59.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern</td>
<td>99.7 64.1 95.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern</td>
<td>31.4 79.3 48.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern</td>
<td>98.4 23.3 78.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern</td>
<td>42.6 95.0 78.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern</td>
<td>86.7 67.7 88.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern</td>
<td>99.6 46.4 97.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern</td>
<td>20.1 86.3 23.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern</td>
<td>97.0 57.9 92.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here, we proceed further and consider the quadrupolar modulation of the primordial power spectrum as suggested by Ackerman et al. (2007):

$$P(k) = P(k) \left[ 1 + \sum_{M} g_{2M} Y_{2M}(k) \right].$$

(32)

Given such a spectrum, the CMB temperature field is expected to exhibit a correlation between $a_{\ell m}$ and $a_{\ell' m'}$ with $M = 0, 2$. Therefore, the BipoSH coefficients $A_{l0}^2$ and $A_{l2}^2$ are sensitive to $g_{2M}$. In the limit of weak anisotropy, Kim & Komatsu (2013) proposed an optimal estimator for $g_{2M}$:

$$\hat{g}_{2M} = \frac{1}{2} \sum_{M'} \left( F^{-1} \right)_{MM} \sum_{\ell m} \sum_{\ell' m'} \frac{\partial C_{\ell m, \ell' m'}(k)}{\partial C_{2M, 0}} \frac{C_{2M, 0}}{g_{2M=0}},$$

(33)

where $a$ is the CMB data vector in harmonic space and $C$ is its covariance matrix, and

$$F_{MM'} = \frac{1}{2} \sum_{\ell m} \sum_{\ell' m'} \left( C^{-1} \right)_{\ell m} \frac{\partial C_{\ell m, \ell' m'}}{\partial C_{2M, 0}} \left( C^{-1} \right)_{\ell' m'} \frac{\partial C_{2M, 0}}{\partial C_{2M', 0}} \frac{g_{2M=0}}{g_{2M=0}}.$$
Fig. 17. Difference of the $N$-point correlation functions determined from the $N_{\text{side}} = 64$ Planck CMB 2015 temperature estimates and the corresponding means estimated from 1000 simulations. Results are shown for the 2-point, pseudo-collapsed 3-point (upper left and right panels, respectively), equilateral 3-point, and connected rhombic 4-point functions (lower left and right panels, respectively). Correlation functions are shown for the analysis performed on northern (blue) and southern (red) hemispheres determined in the ecliptic coordinate frame. The solid, dashed, dot-dashed, and dotted lines correspond to the Commander, NILC, SEVEM, and SMICA maps, respectively. Note that the lines lie on top of each other. The shaded dark and light grey regions indicate, for reference, the 68% and 95% confidence regions, respectively, determined from the SMICA simulations.

Table 17. Constraints on the quadrupolar modulation, determined from the Commander, NILC, SEVEM, and SMICA foreground-cleaned maps.

<table>
<thead>
<tr>
<th>Method</th>
<th>$q_{2M} \times 10^7$</th>
<th>$q_2 \times 10^2$</th>
<th>Data</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commander</td>
<td>1.31 ± 1.22</td>
<td>$0.43 \pm 0.86$ + $i(-0.01 \pm 0.68)$</td>
<td>(1.08 ± 0.89) + $i(-0.38 \pm 0.86)$</td>
<td>0.97</td>
</tr>
<tr>
<td>NILC</td>
<td>0.88 ± 1.21</td>
<td>$0.37 \pm 0.85$ + $i(0.33 \pm 0.67)$</td>
<td>(0.87 ± 0.88) + $i(-0.26 \pm 0.86)$</td>
<td>0.77</td>
</tr>
<tr>
<td>SEVEM</td>
<td>0.85 ± 1.22</td>
<td>$0.35 \pm 0.85$ + $i(0.34 \pm 0.67)$</td>
<td>(1.00 ± 0.88) + $i(-0.25 \pm 0.86)$</td>
<td>0.81</td>
</tr>
<tr>
<td>SMICA</td>
<td>1.10 ± 1.10</td>
<td>$0.46 \pm 0.81$ + $i(0.26 \pm 0.64)$</td>
<td>(0.93 ± 0.83) + $i(-0.26 \pm 0.82)$</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Notes. The first three columns correspond to the five independent parts of the quadrupolar modulation, which we have chosen to present using a complex notation for $q_{2M}$. The quoted error bars are at the 68% confidence level. The quadrupolar modulation amplitude is given in the fourth column, while the mean and standard deviation of $q_2$, estimated from simulations, are provided in the fifth column.

Table 17 presents results from an analysis of the Planck data using the extended common mask, UTA76, and limiting the range of multipoles to $2 \leq \ell \leq 1200$. When including data at higher $\ell$-values, simulations show evidence for large statistical uncertainties in the recovered $q_{2M}$ values that are a consequence of the many holes in the mask related to point...
**Fig. 18.** Difference of the $N$-point correlation functions determined from the $N_{\text{side}} = 64$ Planck SMICA CMB 2015 temperature estimates and the corresponding means estimated from 1000 simulations. Results are shown for the 2-point, pseudo-collapsed 3-point (upper left and right panels, respectively), equilateral 3-point, and connected rhombic 4-point functions (lower left and right panels, respectively). Correlation functions are shown for the analysis performed on negative (blue) and positive (red) hemispheres determined in the ecliptic (solid lines), Doppler boost (DB, dashed lines), and dipole modulation (DM, dot-dashed lines) coordinate frames. The shaded dark and light grey regions indicate the 68% and 95% confidence regions, respectively.

sources. Therefore, imposing this limit $\ell \leq 1200$ does not significantly affect the constraining power of the analysis. We then estimate the amplitude of the quadrupolar modulation using the relation $g_2 = \left( \frac{1}{5} \sum_M |g_2 M|^2 \right)^{1/2}$. Due to the nature of the estimator, which is necessarily positive, the estimation is biased. For an unbiased assessment, we estimate the mean and standard deviation of $g_2$ from simulations. We find no evidence for quadrupolar modulation of the primordial power spectrum. However, the derived limits allow us to impose tight constraints on statistically anisotropic inflationary models, such as those including vector fields during inflation. A companion paper, Planck Collaboration XX (2016), contains a more complete discussion on the theoretical implications of this constraint.

### 5.4. Point-parity asymmetry

The CMB anisotropy field defined on the sky, $T(\hat{n})$, may be divided into symmetric, $T^+(\hat{n})$, and antisymmetric, $T^- (\hat{n})$, functions with respect to the centre of the sphere, as previously described in PCIS13. These functions have even and odd parity, and thus correspond to spherical harmonics with even and odd $\ell$-modes, respectively. On the very large scales corresponding to the Sachs-Wolfe plateau of the temperature power spectrum ($2 \leq \ell \leq 30$), the Universe should be parity neutral with no particular parity preference exhibited by the CMB fluctuations. However, an odd point-parity preference has previously been observed in the WMAP data releases (Land & Magueijo 2005a,b; Kim & Naselsky 2010a,b; Gruppuso et al. 2011) and the Planck 2013 results. Here, we investigate the parity asymmetry in the 2015 temperature maps at $N_{\text{side}} = 32$. We consider the following estimator:

$$R_{TT}(\ell_{\text{max}}) = \frac{D_{TT}^+(\ell_{\text{max}})}{D_{TT}^- (\ell_{\text{max}})},$$

(36)

where $D_+(\ell_{\text{max}})$ and $D_-(\ell_{\text{max}})$ are given by

$$D_{TT}^{+, -} = \frac{1}{N_{\text{tot}}} \sum_{\ell = 2}^{\ell_{\text{max}}} \frac{\ell (\ell + 1)}{2\pi} C_{\ell}^{TT},$$

(37)

A16, page 23 of 62
$\ell^{\text{max}}$ is the total number of even (+) or odd (−) multipoles included in the sum up to $\ell_{\text{max}}$, and $D_{TT}^{\ell}$ is the temperature angular power spectrum computed using a quadratic maximum likelihood (QML) estimator (Gruppuso et al. 2011). The $\ell(\ell + 1)/(2\pi)$ factor in Eq. (37) effectively flattens the spectrum across the $\ell$-range of the Sachs-Wolfe plateau (up to $\ell = 50$) in a $\Lambda$CDM model.

Figure 19 presents the ratio, $R_{TT}(\ell_{\text{max}})$, for the 2015 component-separated maps, together with the distribution determined from the SMICA MC simulations which serves as a reference for the expected behaviour of the statistic in a parity-neutral Universe. The distributions for the other CMB maps are very similar. The four component-separation products are in good agreement, indicating an odd-parity preference at very large scales for the multipole range considered in this test.

Figure 20 shows the lower-tail probability for the data as compared to simulations as a function of $\ell_{\text{max}}$. The results are in good agreement with those in PCIS13. The cleaned CMB maps yield generally consistent profiles which signify an anomalous odd-parity preference in the multipole region $\ell_{\text{max}} = 20–30$. The minimum in the lower-tail probability occurs at $\ell = 28$ corresponding to a value of 0.2% for NILC, SEVEM, and SMICA, and 0.3% for Commander4.

As a first attempt to quantify any a posteriori effects in the significance levels, we consider how many MC simulations appear in the lower tail of the MC distribution with a probability equal to, or lower than, 0.2%, for at least one $\ell_{\text{max}}$ value over a specific range. For $\ell_{\text{max}}$ in the range 3–50, the total number of simulated maps with this property is less than 20 over 1000 MC maps, implying that, even considering the LEE, an odd-parity preference is observed with a lower-tail probability of less than 2%.

5.5. Mirror-parity asymmetry

For the Planck 2013 data release, we studied the properties of the temperature data at a resolution of $N_{\text{side}} = 16$ under reflection with respect to a plane to search for mirror symmetries. Such a symmetry might be connected to non-trivial topologies (Starobinsky 1993; Stevens et al. 1993; de Oliveira-Costa et al. 1996). In Planck Collaboration XXIII (2014), we reported evidence for an antisymmetry plane, with a perpendicularly directed vector given by $(l, b) = (264^\circ, -17^\circ)$. However, the probability of the results was slightly dependent on the method of foreground cleaning, with a $p$-value ranging from 0.5% for Commander-Ruler to 8.9% for SMICA. The same direction was also found in the WMAP 7-year data (Finelli et al. 2012), and is close to that determined for the dipole modulation in the Planck 2013 data release (PCIS13), suggesting possible connections between the two directional anomalies.

We now proceed to reanalyse the status of mirror symmetries using the Planck 2015 full mission temperature data at both $N_{\text{side}} = 16$ and $N_{\text{side}} = 32$. In order to avoid possible bias introduced by the use of the Galactic mask5 the results are derived from the full-sky Commander, NILC, and SMICA maps described in Sect. 2. For SEVEM, a customized map is first produced by inpainting about 3% of the map along the Galactic plane using a diffusive inpainting technique. This is then smoothed to the appropriate lower resolutions for further analysis. Following Finelli et al. (2012), we consider the estimators in the pixel domain given by:

$$S^\pm(\hat{n}_i) = N_{\text{pix}}^{-1} \sum_{j=1}^{N_{\text{pix}}} \left[ 1 - \frac{3}{4} \left( \frac{\delta T}{T}(\hat{n}_i) \pm \frac{\delta T}{T}(\hat{n}_j) \right) \right]^2,$$

(38)

where the sum is over all $N_{\text{pix}}$ HEALPix pixels, $(\delta T/T)(\hat{n}_i)$ is the CMB temperature anisotropy measured at the pixel defined by the unit vector $\hat{n}_i$, and $\hat{n}_k$ is the opposite direction with respect to the plane defined by $\hat{n}_i$, i.e.,

$$\hat{n}_k = \hat{n}_i - 2(\hat{n}_i \cdot \hat{n}_i)\hat{n}_i.$$

(39)

In the case where we would like to test the probability of finding a Universe with either odd or even parity preference, the probability would be higher by a factor of about two.
Table 18. Lower-tail probability for the $S^+$ statistics of the component-separated maps at $N_{\text{side}} = 16$ and $N_{\text{side}} = 32$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Probability [%]</th>
<th>Direction $(l, b)$ ['']</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{side}} = 16$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>min$(S^+)$ . . . . 2.9 (264.4, −17.0)</td>
<td></td>
</tr>
<tr>
<td>NILC</td>
<td>min$(S^-)$ . . . . 12.0 (260.4, 48.1)</td>
<td></td>
</tr>
<tr>
<td>$N_{\text{side}} = 32$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>min$(S^+)$ . . . . 1.9 (264.4, −15.7)</td>
<td></td>
</tr>
<tr>
<td>NILC</td>
<td>min$(S^-)$ . . . . 10.0 (265.3, 46.2)</td>
<td></td>
</tr>
</tbody>
</table>

Note that we expect $S^+$ to be small if the points on opposite sides of the mirror are negatives of each other, and $S^-$ to be small when they are the same.

We compute these quantities for each of the 3072 (12288) directions defined at resolution $N_{\text{side}} = 16$ (32), and allow the $j$ and $k$ indices to run over all of the pixels of the low-resolution full-sky maps. We perform the same analysis on 1000 FFP8 simulations and store the minimum value of $S^+$ for each of these to compute probabilities. The results are summarized in Table 18 and Fig. 21.

We confirm that the full mission Planck temperature data at $N_{\text{side}} = 16$ exhibits the most anomalous mirror antisymmetry in the direction $(l, b) = (264^\circ, -17^\circ)$, consistent with the result from the 2013 nominal mission data, with a probability which ranges from 1.6% for SEVEM to 2.9% for Commander. This is within 40' of the preferred direction identified by the dipole modulation analysis in Sect. 6.2. The corresponding results at $N_{\text{side}} = 32$ yield approximately the same direction, $(l, b) = (264^\circ, -16^\circ)$, with a slightly increased probability, ranging from 0.8% for SEVEM to 1.9% for Commander.

We also note that the CMB pattern exhibits a mirror symmetry in the direction $(l, b) = (260^\circ, 48^\circ)$, consistent with that found in the WMAP 7-year data (Finelli et al. 2012), and close to that identified by the solar dipole (Planck Collaboration VIII 2016). However, the significance of the symmetry pattern is less than in the antisymmetric case.

This extension of the analysis to higher resolution than in our previous work shows that the antisymmetry property does not seem to be confined to the largest angular scales, although we have not attempted to correct for any a posteriori choices made in the analysis. The detailed connection of this antisymmetry property to the low-variance and hemispherical asymmetry observed on these scales remains an open issue.

5.6. Local peak statistics

Local extrema or peaks, as introduced in Sect. 4.5.3, can be employed to search for localized anomalies on the CMB sky by examining how their statistical properties vary in patches as a function of location.

Initially, we consider a further test for asymmetry by examining the differences in the peak distribution when divided according to orientation with respect to a previously specified asymmetry direction. In particular, we select the peaks both in a disc of radius 70' centred on $(l, b) = (225^\circ, -18^\circ)$ (the positive direction of the dipole defined in Sect. 6.2 for SMICA) and in the corresponding antipodal disc, then construct the empirical peak height CDFs to be compared with the full-sky median FFP8 distribution, as shown in Fig. 22. For maps filtered with a 40' FWHM GAUSS filter the distribution of the peaks for the positive-direction disc is in general agreement with the full sky result, while that for the negative-direction is marginally different. Moreover, this pattern of behaviour is seen over a number of filtering scales, both for the KS deviation from the median full-sky simulated CDFs, and the spread of extremal values when comparing positive and negative regions. We also find that the properties of the negative disc affect the $p$-value results for a full sky KS test on data filtered with an SSG84 filter of 500' FWHM, as seen in Sect. 4.5.3.

We can then extend the analysis for the 40' GAUSS-filtered data by considering the variation in the peak statistical properties.
the $p$-value for certain disc locations drops to 0.1% (i.e., the sky counts exceed anything seen in simulations). However, one needs to account for the a posteriori selection of significant regions in the determination of the true significance. It should also be noted that regional variations of the UTP are seen at similar levels when inspecting the peak-count statistics maps derived for randomly selected realizations of the simulations. Moreover, the significance of such peak-counting anomalies is degraded with larger disc diameters, and becomes insignificant for counts on the full sky. Thus, no significant anomalies can be claimed for the peak-count statistics of the Planck data.

A powerful non-parametric test of statistical isotropy is provided by the two-sample KS-deviation between the full sky empirical peak height CDF $F_{\theta}(x)$ (see Eq. (28)) and an empirical peak height CDF $F_{\theta'}(x)$ derived from a subsample of the distribution, again defined by the peaks within discs of $30^\circ$ diameter as defined above. The two-sample KS-deviation

\begin{equation}
K_{\text{UTP}} \equiv \sqrt{\frac{mn'}{n+n'}} \sup_x |F_{\theta'}(x) - F_{\theta}(x)|
\end{equation}

for a partial sky region shares samples between the two CDFs, and can be calculated extremely efficiently using rank statistics according to

\begin{equation}
K_{\text{UTP}} \equiv \sqrt{\frac{mn'}{n+n'}} \max_i \left| \frac{r'(i) - 1}{n'} - \frac{r(i) - 1}{n} \right|, \tag{41}
\end{equation}

where $r$ and $r'$ denote the ranks of a value with index $i$ in the full set of $n$ and restricted set of $n'$ samples, respectively. Maps of the upper tail probability are then determined by comparison with the equivalent quantities computed from simulations; $-\log_{10}(\text{UTP})$ maps are shown in Fig. 24. The majority of the selected locations are consistent with the full-sky distribution, thus indicating the statistical isotropy of the Planck maps. The most prominent feature in each of the local KS-deviation maps appears south of the Galactic centre and may be associated with a cold region crossing the Galactic plane. However, as with the peak counts, it cannot be interpreted as statistically anomalous.

5.7. The Cold Spot

Since its discovery in the WMAP first-year data (Vielva et al. 2004), the Cold Spot, centred at Galactic coordinates $(l, b) = (210^\circ, -57^\circ)$ has been one of the most extensively studied large-scale CMB anomalies. In the 2013 release (Planck Collaboration XXIII 2014), Planck confirmed the apparently anomalous nature of this feature in temperature, in terms of the area of the SMHWW coefficients on angular scales of $\approx 10^\circ$ on the sky; the 2015 release has also confirmed this feature (see Sects. 4.5.2 and 4.5.3). The CMB temperature anisotropies around the Cold Spot as observed by Planck are shown in the top panel of Fig. 25. The peak merger tree within the Cold Spot region is presented in the lower panel of the figure and provides a multiscale view of its structure (see Sect. 4.5.4 for details).

The robustness of the detection of the anomalies discussed in this paper is a non-trivial issue. For the particular case of the Cold Spot, this has been reviewed by Vielva (2010), and addressed in detail by Cruz et al. (2006), paying specific attention to the impact of a posteriori choices. In particular, the latter study focused on the original test that indicated the presence of this feature on the sky, confirming a significance between 1% and 2%. An alternative analysis of the significance based on two statistical tests with different levels of conservativeness was made.
by McEwen et al. (2005), providing values of 0.1% and 4.7%, respectively. The statistical significance of the Cold Spot was questioned by Zhang & Huterer (2010) who found a low significance after performing a study based on different kernels. As discussed in more detail by Vielva (2010), this result can also be interpreted as evidence that not all kernels are necessarily suitable for the detection of arbitrary non-Gaussian features.

The possibility that the Cold Spot arises from instrumental systematics (Vielva et al. 2004) or foreground residuals (Liu & Zhang 2005; Cruz et al. 2006) has been largely rejected. However, several non-standard physical mechanisms have been proposed as possible explanations. These include the gravitational effect produced by a collapsing cosmic texture (Cruz et al. 2007), the linear and nonlinear ISW effect caused by a void in the large-scale structure (e.g., Tomita 2005; Inoue & Silk 2006; Rudnick et al. 2007; Tomita & Inoue 2008; Finelli et al. 2016), a cosmic bubble collision within the eternal inflation framework (Czech et al. 2010; Feeney et al. 2011; McEwen et al. 2012), and a localized version of the inhomogeneous reheating scenario within the inflationary paradigm (Bueno Sanchez 2014).

Since the other scenarios lack additional evidence, the void hypothesis would seem to be the most plausible, depending on the sizes, density contrasts, and profiles assumed in the computations, some of which are not in agreement with either observation (Cruz et al. 2008) or current N-body studies (Cai et al. 2010; Watson et al. 2014). However, Szapudi et al. (2015) have recently detected a large void in the WISE-2MASS galaxy catalogue aligned with the Cold Spot, with an estimated radius of around 200 $h^{-1}$ Mpc, an averaged density contrast of $\delta \approx -0.1$, and centred on a redshift of $z \approx 0.15$. Large voids with similar characteristics are not unusual in the standard ΛCDM model (Nadathur et al. 2014). In fact, N-body simulations predict about 20 such voids in the local Universe ($z < 0.5$). However, Zibin (2014) and Nadathur et al. (2014) indicate that the expected signal due to the linear and nonlinear ISW effects caused by this structure is not large enough to explain the temperature decrement associated with the Cold Spot.

The new Planck data release allows us to further explore the statistical nature of the Cold Spot. Two previous studies (Zhao 2013; Gurzadyan et al. 2014) have claimed inconsistencies of the internal properties of the Cold Spot with the Gaussian hypothesis, which we re-address here. In particular, we consider the small-scale fluctuations within a disc-like region of radius $\approx 25'$. Several statistical quantities are computed from the full-resolution temperature maps within the Cold Spot region. This is divided into a central disc of diameter 1° surrounded by a set of 13 concentric annuli with central radii spaced in steps of about 2°, thus allowing us to build angular profiles for the mean, variance, skewness, and kurtosis. These are then compared to specialized CMB realizations, generated as follows. A set of Gaussian CMB skies is simulated using the FFP8 reference spectrum, and convolved with a Gaussian beam of 5' FWHM. As for the FFP8 simulations themselves, these maps are rescaled, as discussed previously. Only those that contain a spot as extreme as the Cold Spot at a scale $R = 300'$ in SMHW space are retained, and these are rotated such that each simulated cold spot is relocated to the actual position of the Cold Spot (this ensures that the noise properties are identical for both data and simulations). This selection criterion corresponds to the characteristic that originally indicated the presence of the Cold Spot in the observed sky. As a final step, for each remaining CMB simulation a noise realization is added, consistent with each component-separation method.

The results are presented in Fig. 26. Focusing on the profile of the mean value, it is apparent that the largest deviations from the simulations appear on scales around 15', which corresponds to a hot ring structure, as seen in Fig. 25 and previously discussed in Cayón et al. (2005) and Nadathur et al. (2014). Notice that on the smallest scales the mean profile is also somewhat deviant with respect to the simulations, but this may be connected to selection bias, since we are considering CMB

**Fig. 23.** Map of $-\log_{10}$ (UTP) for peak counts in the Planck 40' GAUSS-filtered temperature data, where each pixel encodes the probability determined for a 30' diameter disc centred on it.

**Fig. 24.** Map of $-\log_{10}$ (UTP) for the two-sample KS-deviation where each pixel encodes the probability determined for a 30' diameter disc centred on it, as computed from the Planck 40' GAUSS-filtered temperature data.
set in latitude [deg]

In conclusion, it appears that only the mean temperature profile of the Cold Spot should be considered anomalous when compared to CMB cold spots that are as statistically extreme. All other measures of its internal structure are consistent with expectations.

As a final remark, we note that the high-pass filtering currently applied to the Planck CMB polarization maps severely limits the possibility of conducting targeted analyses to discriminate between different possible origins of the Cold Spot. For example, no polarization signal would be expected in those models producing secondary anisotropies due to a gravitational effect, whereas a specific pattern might be expected in a bubble collision scenario (Czech et al. 2010). Appropriate tests will be pursued in future work, once the large-scale polarization data are available.

6. Dipole modulation and directionality

In this section, we examine isotropy violation related to dipolar asymmetry, various forms of which have been noted since the early WMAP releases (Eriksen et al. 2004a). We perform a non-exhaustive series of tests in an attempt to narrow down the nature of the asymmetry (on the assumption that it is not simply a statistical fluke). First, we will briefly describe some similarities and differences between the tests that are important for making a proper comparison of the results.

All the tests in this section share in common the fitting of a dipole. This is done either by fitting for a dipole explicitly in a map of power on the sky (Sects. 6.1 and 6.5), by employing Bayesian techniques in pixel space for a specific model (Sect. 6.2), or by measuring the coupling of \( \ell \) to \( \ell \pm 1 \) modes in the CMB covariance matrix (Sects. 6.3, 6.4, and 6.6). The differences arise from how the fitted dipoles are combined, which determines the specific form of asymmetry that the test is sensitive to.
The tests can be divided into two categories, amplitude-based and direction-based. Sections 6.1 to 6.4 are all sensitive to the amplitude of a dipole modulation. Specifically, Sect. 6.1 looks for dipole modulation in the pixel-to-pixel variance of the data, while Sects. 6.2–6.4 all search for dipole modulation of the angular power spectrum. The distinction between these two approaches is mainly one of ℓ weighting.

Sections 6.5 and 6.6 both examine aspects of directionality in the data, where the directions are extracted from dipole fits but combined in different ways. Section 6.5 fits for dipoles in band power (with similar results for variance) and only uses the direction information, while Sect. 6.6 weights each dipole equally across all scales and uses the amplitude information as well.

The differences between the approaches of these sections should be kept in mind when comparing their results. For example, although Sects. 6.5 and 6.6 both look for a directional signal in the data, they are optimized for different forms of deviations from statistical isotropy. It is therefore unsurprising that they arrive at different results. However, the signal found in Sect. 6.5, if not simply a statistical fluke, is constrained by the results of Sect. 6.6.

Regarding the impact on the dipolar modulation results of the lack of the aberration contribution to the simulations, we note the following. In general the analyses are either sensitive only to large angular scales, or only claim possible detections on such scales, where the effect of aberration will be negligible and hence the conclusions are unlikely to change. A possible exception is in relation to the results of Sect. 6.5, where claims are made about effects extending out to ℓ_{max} = 1500. It is plausible that the effects of aberration could start to become important on these scales.

### 6.1. Variance asymmetry

The study of power asymmetry via the local variance of the CMB fluctuations was first performed by Akrami et al. (2014) for the Planck 2013 and WMAP 9-year temperature data. The approach was motivated by its conceptual and implementational simplicity, its directly intuitive interpretation, and by virtue of being defined in pixel space, a useful complementarity to other mostly harmonic-based methods. The statistic was computed over patches of different sizes and positions on the sky, and compared with the values obtained from statistically isotropic simulations. It was found that none of the 1000 available simulations had a larger variance asymmetry than that estimated from the data. This suggested the presence of asymmetry at a statistical significance of at least 3.3σ, with a preferred direction (ℓ, b) ≈ (212°, −13°) in good agreement with other studies. In this section, we revisit the variance asymmetry and report the results of the analysis for the Planck 2015 temperature maps at full resolution, N_{side} = 2048.

The analysis proceeds as follows. We consider a set of discs of various sizes centred on the pixels of a HEALPix map defined by a specific N_{side} value. For each sky map, we first remove the monopole and dipole components from the masked sky and then compute the variance of the fluctuations on a given disc using only the unmasked pixels. This yields a local-variance map at the HEALPix resolution of interest. We also estimate the expected average and variance of the variance on each disc from the simulations and then subtract the resulting average variance map from both the observed and simulated local-variance maps. Finally, we define the amplitude and direction of the asymmetry by fitting a dipole to each of the local-variance maps, where each pixel is weighted by the inverse of the variance of the variances computed from the simulations at that pixel. At all stages, we use only the discs for which more than 10% of the area is unmasked, although our results are robust against the choice of this value.

The computed local-variance amplitudes are then used to compare the data with statistically isotropic simulations. Note that we use only the dipole amplitudes of the local-variance maps to measure the significance of the asymmetry; the amplitudes of higher multipoles were shown by Akrami et al. (2014) to be consistent with statistically isotropic simulations and we therefore do not consider them in the present paper.

In Akrami et al. (2014), the sensitivity to the disc size was assessed using both statistically isotropic and anisotropic simulations. The free parameters, i.e., the number and size of the discs, were then fixed by these simulations. It was found that for 3072 patches centred on the set of pixels defined at N_{side} = 16, the simulated asymmetry signals were not detected when either very small (r_{disc} < 4°) or very large (r_{disc} > 16°) discs were used.

The former effect is due to a combination of the low number of pixels per disc and an insufficient number of discs to cover the entire sky when N_{side} = 16 reference grids are used. However, it has recently been shown by Adhikari (2015) that using a larger number of small discs (by increasing N_{side} to 32, 64, 128, and 256, depending on the disc size) in order to cover the entire sky allows the local-variance method to detect the large-scale anomalous asymmetry as well as the Doppler boost signal from the Planck 2013 data, at a significance of >3.3σ. Fantaye (2014) has demonstrated that the Doppler boost signal can be...
detected at a similar level of significance using needlet bandpass filtering of the data, even with large discs, when simulations are deboosted. Here, in contrast to the 2013 analysis, we use maps which contain Doppler boosting, for both simulations and data, and therefore we do not detect any Doppler boost signal when using a large number of small discs.

The low observed significance levels when large discs are used is due to the cosmic variance associated with the largest-scale modes. Motivated by the analysis of Fantaye (2014), and in order to address this issue, we also perform analyses using a Butterworth high-pass filter,

\[ H(\ell) = \frac{(\ell/\ell_0)^4}{1 + (\ell/\ell_0)^4}, \]

centred at multipoles \( \ell_0 = 5, 10, 15, 20, \) and 30. In addition, the filtering of low multipoles allows us to establish the contribution of such modes to any detected asymmetry.

Table 20. \( p \)-values for the variance asymmetry measured by 8° discs for the four component-separated temperature maps and different high-pass filter scales.

<table>
<thead>
<tr>
<th>( \ell_0 )</th>
<th>Comm.</th>
<th>NILC</th>
<th>SEVEM</th>
<th>SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfiltered</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>5</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>10</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>15</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>20</td>
<td>0.4</td>
<td>&lt;0.1</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>30</td>
<td>1.8</td>
<td>0.8</td>
<td>1.8</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Notes. The values represent the fraction of simulations with local-variance dipole amplitudes larger than those inferred from the data.

Here, based on the analysis of Akrami et al. (2014), we restrict our analysis to those disc sizes for which 3072 discs, corresponding to an \( N_{\text{side}} = 16 \) map, cover the entire sky, i.e., to the range 4°–90°. Consistent results can be obtained by choosing other values of \( N_{\text{side}} \) for a given disc size provided that the entire sky is covered by the discs. Here, for simplicity, we work with the same \( N_{\text{side}} = 16 \) for all disc sizes.

Our results for the measured amplitude of the variance asymmetry, compared to the values from the simulations, as well as the corresponding dipole directions, are shown in Fig. 27. The \( p \)-values are given for different disc sizes and in terms of the number of simulations with local-variance dipole amplitudes greater than the ones measured from the data. Note that since the discs with different sizes used in our analysis are correlated, the significance levels are also correlated. For this reason we choose to show the \( p \)-values as a function of disc size instead of combining them into a single number. Moreover, it should be noted that the significance values we present here do not incorporate any corrections to account for the choice of parameters adopted during method calibration, specifically the dipole amplitudes and directions for the anisotropic simulations that were used to fix the range of disc sizes and number of patches.

It can be seen from the upper panel of Fig. 27 that for the unfiltered map the significance of the power asymmetry drops quickly when we increase the disc size to radii greater than 16°. This is no longer the case, however, when the lowest multipoles are filtered out. For example, when the filter scale is set to \( \ell_0 = 5 \), i.e., when the very low multipoles which are affected most by cosmic variance are suppressed, the variance asymmetry is detected at the 3\( \sigma \) level for all disc sizes, as shown in Fig. 27. Table 20 presents the \( p \)-values of the variance asymmetry using 8° discs and for various values of \( \ell_0 \). Our results show that variance asymmetry is detected with a remarkable significance for all disc sizes when very low multipoles are filtered out. In addition, the variance asymmetry amplitude slowly decreases with increasing \( \ell_0 \), as seen in the upper panel of Fig. 28. For \( \ell_0 \geq 20 \), the dipole amplitude becomes too small and we find no significant variance asymmetry. It is interesting to note, however, that the dipole directions found for large \( \ell_0 \) are closely aligned with those found for \( \ell_0 < 20 \).

The lower panel of Fig. 27 shows the dipole directions we find using different disc sizes and different filter scales for SMICA. The dipole directions for the Commander, NILC, and SEVEM component-separated maps are very similar to those shown. The asymmetry directions found here are consistent with those determined by other analyses in this paper.
Table 21. Local-variance dipole amplitudes and directions.

<table>
<thead>
<tr>
<th>ℓ₀</th>
<th>A°</th>
<th>Direction (l, b) [']</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfiltered . .</td>
<td>0.052 ± 0.016</td>
<td>(210, −26)</td>
</tr>
<tr>
<td>5 . . . . . . . .</td>
<td>0.046 ± 0.014</td>
<td>(208, −24)</td>
</tr>
<tr>
<td>10 . . . . . . .</td>
<td>0.040 ± 0.014</td>
<td>(199, −16)</td>
</tr>
<tr>
<td>15 . . . . . . .</td>
<td>0.038 ± 0.012</td>
<td>(206, −16)</td>
</tr>
<tr>
<td>20 . . . . . . .</td>
<td>0.028 ± 0.010</td>
<td>(202, −18)</td>
</tr>
<tr>
<td>30 . . . . . . .</td>
<td>0.025 ± 0.010</td>
<td>(199, −19)</td>
</tr>
</tbody>
</table>

Notes. All values quoted here are for 8° discs. This table is for the Commander component-separation method, but the results are similar for the other methods. (a) A = 2(A_{\text{Planck}} − A_{\text{FFP8}}), where A_{\text{Planck}} and A_{\text{FFP8}} are the local-variance dipole amplitudes of the data and the FFP8 simulations, respectively. The quoted errors are the dispersion of the simulation amplitudes. Assuming a pure dipole modulation model, A to first order would correspond to the modulation amplitude.

6.2. Dipole modulation: pixel-based likelihood

In PClIS13 we presented an analysis of the apparent anisotropic distribution of large-scale power in the Planck 2013 temperature data within the parametric framework defined by Gordon (2007) and Hoftuft et al. (2009), who introduced an explicit dipole modulation field to model potential hemispherical power asymmetry. The following is a direct update of that analysis using the Planck 2015 CMB data at N_{\text{side}} = 32, retaining the 2013 common mask to explicitly test for consistency with the earlier study. All results are found to be in excellent agreement. In the following, we therefore only consider a smoothing scale of 5° FWHM as a representative example. This is the highest angular resolution accessible for an N_{\text{side}} = 32 map.

Recall first the basic data model adopted in the dipole modulation approach: rather than assuming the CMB sky to be a statistically isotropic Gaussian field, we allow for an additional dipole modulation, resulting in a data model of the form \( d = BM_0 + n \), where \( M_0 = (1 + \alpha \hat{p} \cdot \mathbf{n}) \delta \) is an offset dipole field modifying an intrinsically isotropic signal \( \mathbf{n} \) with a dipole amplitude \( \alpha \) pointing towards some preferred direction \( \hat{p} \). \( B \) denotes convolution with an instrumental beam, and \( n \) denotes instrumental noise. Additionally, we model the power spectrum of the underlying statistically isotropic field in terms of a two-parameter amplitude—tilt model of the form \( C(q, n) = q \left( \ell / 30^\circ \right)^{2} C_{\Lambda CD M}^\ell \), where \( C_{\Lambda CD M}^\ell \) is the best-fit Planck 2015 ΛCDM spectrum (Planck Collaboration XI 2016). The two parameters \( q \) and \( n \) can accommodate a deficit in power at low \( \ell \) as compared to the best-fit cosmology that would otherwise create a tension with the underlying statistically isotropic model and result in the analysis measuring a combination of both asymmetry and power mismatch.

In the absence of any dipole modulation, \( \alpha = 0 \), the total data covariance matrix is given by \( C = BSBM + N \), where \( S_{\text{iso}} \) is the standard statistically isotropic CMB covariance matrix given by the power spectrum, \( C_{\nu} \), \( N \) is the noise covariance matrix, and the corresponding likelihood is given by the usual expression for a multivariate Gaussian distribution. With dipole modulation, this generalizes straightforwardly to \( C = BSM_{\text{iso}} M^T B^T + N \), with the likelihood given by

\[
L(\alpha, \hat{p}, q, n) \propto \exp \left[ -\frac{1}{2} d^{T} (BSBM^{T} B^{T} + N)^{-1} d \right].
\]

Figure 29 and Table 22 summarize this five-dimensional likelihood in terms of marginal parameters for each of the four
Fig. 29. Top: marginal posterior distributions on the dipole modulation amplitude $\alpha$, as derived from Planck 2015 temperature observations at a smoothing scale of 5\' FWHM for Commander (red), NILC (orange), SEVEM (green), and SMICA (blue). The plot corresponds directly to Fig. 32 of Planck Collaboration XXIII (2014). The Commander, SEVEM, and SMICA posterior distributions coincide almost perfectly both internally and with the corresponding marginal SMICA 2013 posterior, shown as a dashed black line. Bottom: corresponding marginal two-dimensional constraints on the low-\ell power spectrum amplitude $q$ and tilt $n$, defined relative to the best-fit Planck 2015 CMB sky model.

Table 22. Summary of dipole modulation results at a smoothing scale of 5\' for all Planck 2015 CMB temperature solutions, as derived by the brute-force likelihood given by Eq. (43).

<table>
<thead>
<tr>
<th>Method</th>
<th>2013</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dipole modulation amplitude, $\alpha$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>0.078 ± 0.021</td>
<td>0.066 ± 0.021</td>
</tr>
<tr>
<td>NILC</td>
<td>0.069 ± 0.021</td>
<td>0.061 ± 0.022</td>
</tr>
<tr>
<td>SEVEM</td>
<td>0.066 ± 0.021</td>
<td>0.065 ± 0.021</td>
</tr>
<tr>
<td>SMICA</td>
<td>0.065 ± 0.021</td>
<td>0.066 ± 0.021</td>
</tr>
<tr>
<td>Dipole modulation direction, $(\ell, b)$ [']</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>(227, −15) ± 19</td>
<td>(230, −16) ± 24</td>
</tr>
<tr>
<td>NILC</td>
<td>(226, −16) ± 22</td>
<td>(228, −19) ± 29</td>
</tr>
<tr>
<td>SEVEM</td>
<td>(227, −16) ± 24</td>
<td>(226, −17) ± 25</td>
</tr>
<tr>
<td>SMICA</td>
<td>(226, −17) ± 24</td>
<td>(225, −18) ± 24</td>
</tr>
<tr>
<td>Power spectrum amplitude, $q$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>0.961 ± 0.025</td>
<td></td>
</tr>
<tr>
<td>NILC</td>
<td>0.954 ± 0.024</td>
<td></td>
</tr>
<tr>
<td>SEVEM</td>
<td>0.966 ± 0.025</td>
<td></td>
</tr>
<tr>
<td>SMICA</td>
<td>0.960 ± 0.025</td>
<td></td>
</tr>
<tr>
<td>Power spectrum tilt, $n$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>0.082 ± 0.043</td>
<td></td>
</tr>
<tr>
<td>NILC</td>
<td>0.077 ± 0.043</td>
<td></td>
</tr>
<tr>
<td>SEVEM</td>
<td>0.077 ± 0.043</td>
<td></td>
</tr>
<tr>
<td>SMICA</td>
<td>0.081 ± 0.043</td>
<td></td>
</tr>
</tbody>
</table>

6.3. Dipole modulation: QML analysis

In this section we use the QML estimator introduced in Moss et al. (2011) and described in Appendix C to assess the level of dipole modulation in our estimates of the CMB sky at $N_{\text{side}} = 2048$. The specific implementation is essentially identical to that used in Hanson & Lewis (2009), Planck Collaboration XVII (2014), and Planck Collaboration XXVII (2014), and exploits the fact that dipole modulation of any cosmological parameter is equivalent to coupling of $\ell$ to $\ell \pm 1$ modes in the CMB covariance matrix to leading order (see Appendix C). Planck Collaboration XX (2016) presents an alternate analysis for a specific isocurvature model.

Since we are interested in dipole modulation there are three independent estimators. For our particular approach, these are a real-valued $m = 0$ and a complex-valued $m = 1$ estimator, and take the form

$$
\hat{X}_0 = \frac{6}{\tilde{f}_{10}} \sum_{\ell \in \Lambda_0} \delta C_{\ell \ell+1} A_{\ell m} \left( T^*_{\ell m} T^{\ell+1 m} - \left( T^*_{\ell m} T^*_{\ell+1 m} \right) \right),
$$

(44)

$$
\hat{X}_1 = \frac{6}{\tilde{f}_{11}} \sum_{\ell \in \Lambda_1} \delta C_{\ell \ell+1} B_{\ell m} \left( T^*_{\ell m} T^{\ell+1 m+1} - \left( T^*_{\ell m} T^*_{\ell+1 m+1} \right) \right),
$$

(45)

Here $T_{\ell m}$ are $C$-inverse filtered data and $F_{\ell} \equiv \left< T_{\ell m} T^{\ell m}_* \right>$. We adopt the inverse-variance filter from Planck Collaboration XVII (2014), where the approximate filter functions are also specified. We define $\delta C_{\ell \ell+1} \equiv dC_{\ell m}/dX + dC_{\ell+1 m}/dX$, where $X$ is the parameter modulated, and $A_{\ell m}$ and $B_{\ell m}$ are numerical coefficients (details can be found in Appendix C). The factor $f_{1m}$ corrects the normalization for errors introduced by masking:

$$
f_{1m} = \int d\Omega Y^*_{\ell m}(\Omega) M(\Omega),
$$

(46)

where $M(\Omega)$ is the mask. Finally, we correct the direction for the effects of inhomogeneous noise which is not accounted for in the Planck CMB maps, as evaluated over the common mask using the multi-dimensional grid-based Snake algorithm (Mikkelsen et al. 2013). All results correspond to a smoothing scale of 5\' FWHM, the highest resolution supported by an $N_{\text{side}} = 32$ HEALPix grid, but, as in 2013, we consider all smoothing scales between 5\' and 10\' FWHM, reaching similar conclusions in each case: the dipole modulation results derived from the Planck 2015 temperature maps are essentially identical to the 2013 results, with improved internal consistency between the four CMB maps due to better mitigation of systematic errors. The best-fit dipole modulation amplitude at 5\' FWHM is 6–7\% whilst the low-\ell power spectrum has an approximately 3–5\% lower amplitude compared to the best-fit CDM prediction. These results are fully consistent with expectations given that the Planck 2013 sky maps were already cosmic-variance-limited on these angular scales, and the 2015 maps differ from the 2013 maps at the level of only a few microkelvin (Planck Collaboration IX 2016).
filtering process, by weighting the $X_n$ by the inverse of the variance derived from filtered and mean-field corrected simulations.

The physics is readily accessible in this estimator: the $\ell$-dependence in modulation determined by the parameter $X$ is expressed in the $\delta C_\ell X_\ell$ factor, and the relevant scales appear directly in the limits of the sum. We consider the estimator over the range $\ell_{\min} = 2 \leq \ell \leq \ell_{\max}$. The modulation amplitude and direction are then given by

$$\hat{A} = \sqrt{\hat{X}_0^2 + 2|\hat{X}_1|^2},$$

$$\hat{\theta} = \cos^{-1} \left( \frac{\hat{X}_0}{\hat{A}} \right),$$

$$\hat{\phi} = -\tan^{-1} \left( \frac{\text{Im}[\hat{X}_1]}{\text{Re}[\hat{X}_1]} \right).$$

It is worth re-emphasizing that the quantities $\hat{A}$, $\hat{\theta}$, and $\hat{\phi}$ are all dependent on the $\ell$ range considered.

As a consequence of the central limit theorem, for sufficiently large $\ell_{\max}$ the $\hat{X}_n$s are Gaussian-distributed with mean zero, so that the amplitude parameter has a Maxwell-Boltzmann distribution. We fit to this distribution for $\ell_{\max} \gtrsim 10$ when computing the $p$-value, so as not to be influenced by Poisson noise in the tails of the empirical distribution (and we have determined that this is a good fit to the simulations by applying a KS test). For the case of scalar amplitude modulation (i.e., $X = A_1$), and $\ell_{\min} = 2$, the cosmic-variance-limited expectation for the modulation amplitude from statistically isotropic skies is

$$\left( \frac{\Delta A_1}{A_1} \right) \approx \frac{48}{\pi(\ell_{\max} + 4)(\ell_{\max} - 1)}.$$  

This is the cosmic variance for a scale-invariant dipole modulation, and gives a more explicit expression than the $\ell_{\max}^3$ scaling discussed in Hanson & Lewis (2009).

The top panel of Fig. 30 presents results for the $p$-value of the fitted modulation amplitude as a function of $\ell_{\max}$. Note that there are several peaks, at $\ell \approx 40$ and $\ell \approx 67$ (the focus of most attention in the literature), and $\ell \approx 240$. The latter peak, while not previously emphasized, is also present in the WMAP results (see Fig. 15 in Bennett et al. 2011). It is also interesting to note that a modulation amplitude is observed at $\ell_{\max} \approx 800$ that is somewhat lower than what one would typically expect for a statistically isotropic sky. However, the significance is not at the level of the excess dipole modulation at low $\ell$ and will not be discussed further. The dip at $\ell_{\max} \approx 67$, with a $p$-value of 0.9–1.0%, corresponds to the well-known low-$\ell$ signal.

Table 23 presents the corresponding dipole modulation parameters, which are seen to be consistent with previous studies. Note that the mean amplitude expected for a set of statistically isotropic simulations at this $\ell_{\max}$ is 2.9% (in close agreement with the expected value due to cosmic variance, Eq. (50)).

We have therefore determined a phenomenological signature of modulation for $\ell = 2–67$ with a $p$-value of 0.9–1.0%. If such a signal had been predicted by a specific model, then we could claim a significance of about 3$\sigma$. However, in the absence of such an a priori model, we can assess how often we might find a 3$\sigma$ effect by chance, given that it could have occurred over any

---

**Table 23.** Amplitude ($A$) and direction of the low-$\ell$ dipole modulation signal determined from the QML analysis for the range $\ell \in [2, 64]$.

<table>
<thead>
<tr>
<th>Method</th>
<th>$A$</th>
<th>Direction $(l, b)$ [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commander</td>
<td>$0.063^{+0.025}_{-0.013}$</td>
<td>$(213, −26) \pm 28$</td>
</tr>
<tr>
<td>NILC</td>
<td>$0.064^{+0.027}_{-0.013}$</td>
<td>$(209, −25) \pm 28$</td>
</tr>
<tr>
<td>SEVEM</td>
<td>$0.063^{+0.026}_{-0.013}$</td>
<td>$(211, −25) \pm 28$</td>
</tr>
<tr>
<td>SMICA</td>
<td>$0.062^{+0.013}_{-0.011}$</td>
<td>$(213, −26) \pm 28$</td>
</tr>
</tbody>
</table>

**Notes.** The errors are calculated from the cosmic variance expected for statistically isotropic CMB realizations.
which was used as the search limit in Bennett et al. (2011). The probability to exceed is fortunately not very sensitive to \( \ell \), and for any reasonable choice is above 10%.

\( \ell \) range. Since we are looking for a large-scale phenomenon, we assume that the analysis should include the corresponding low-\( \ell \) modes and start at \( \ell = 2 \). In order to correct for a posteriori effects we then adopt the following scheme.

1. We calculate the modulation of each simulation on the scales \( 2 \to \ell \), where \( \ell \in [10, \ell_{\text{max}}] \). For each simulation we find the modulation that gives the smallest probability, \( \eta \) (in the same way that was done for the data).

2. With the distribution of \( \eta \)s given by the simulations we then compare this to the data. That is, we calculate the probability that one would find oneself in a Hubble patch with a modulation amplitude up to \( \ell \in [10, \ell_{\text{max}}] \) that is as significant as (or more significant than) the modulation in the real data.

If \( \ell_{\text{max}} = 132 \) (as chosen by Bennett et al. 2011), the probability of achieving a modulation as large as the Commander (red), NILC (orange), SEVEM (green), and SMICA (blue) data sets, for the range \( \ell \in [10, \ell_{\text{max}}] \). The vertical line corresponds to \( \ell_{\text{max}} = 132 \) which was used as the search limit in Bennett et al. (2011). The probability grows approximately logarithmically with \( \ell_{\text{max}} \). This means that the adopted probability to exceed is fortunately not very sensitive to \( \ell_{\text{max}} \), and for any reasonable choice is above 10%.

6.4. Bipolar spherical harmonics

In the absence of the assumption of statistical isotropy, the CMB two-point correlation function \( C(\hat{n}_1, \hat{n}_2) \neq C(\hat{n}_1, -\hat{n}_2) \) can be most generally expanded in the bipolar spherical harmonic (BipoSH) basis representation as follows:

\[
C(\hat{n}_1, \hat{n}_2) = \sum_{LM \ell \ell'} A_{\ell\ell'}^{LM} (\hat{Y}_{\ell \ell'}(\hat{n}_1) \otimes \hat{Y}_{\ell' \ell'}(\hat{n}_2))_{LM}.
\]  

The BipoSH basis functions, \( \{\hat{Y}_{\ell \ell'}(\hat{n})\}_{LM} \) are tensor products of ordinary spherical harmonic functions, and the corresponding expansion coefficients are termed BipoSH coefficients (Hajian & Souradeep 2003; Hajian & Souradeep 2006). The BipoSH basis provides a complete representation of any form of statistical isotropy violation with the key advantage of separating the angular scale-dependence of the signal in spherical harmonic multipoles, \( \ell \), from the nature of the violation indexed in the bipolar multipole space by \( L \). Consequently, it is possible to simultaneously determine that such a signal is dipolar (\( L = 1 \)), quadrupolar (\( L = 2 \)), octopolar (\( L = 3 \)), and so on, in nature and that the power is restricted to specific ranges of angular scales.

The estimation of BipoSH coefficients from CMB maps is a natural generalization of the more routinely undertaken estimation of the angular power spectrum \( C_{\ell} \). To allow a direct connection to the angular power, we further introduce a set of BipoSH spectra at every bipolar harmonic moment, \( (L, M) \), labelled by a difference index \( d \), defined as follows:

\[
\hat{A}_{L,M}^{\ell+d} = \frac{\hat{\Pi}_L}{\hat{\Pi}_{L+d} C_{\ell}^{00}(\ell+d)}, \quad (0 \leq d \leq L),
\]  

where \( C_{\ell}^{00}(\ell+d) \) are the Clebsch-Gordon coefficients and for brevity the notation \( \Pi_L^{d+1} = \prod_{i=1}^{d+1} \sqrt{2\ell_i + 1} \). BipoSH spectra, clearly, are then simply a generalized set of CMB angular power spectra, with the standard CMB angular power spectrum \( \hat{C}_{\ell} \), being one of them.\(^7\) While \( A_{L,M}^{\ell+d} \) quantifies the properties of the statistically isotropic part of the CMB fluctuations, the additional BipoSH coefficients quantify the statistically anisotropic part of the CMB two-point correlation function.

Thus BipoSH provides a mathematically complete description of all possible violations of statistical isotropy in a Gaussian CMB sky map. It is then always possible to translate any specific model for such a signal into the language of BipoSH and provide a common approach for the multiple specialized tests that have been implemented previously in this paper and elsewhere. However, improving on the analysis of the 2013 Planck data, a new formalism is developed in order to reliably analyse a masked sky, as concisely described in Appendix D. Aluri et al. (2015) provides a more detailed description of the approach and includes an explicit demonstration of its validity using simulations.

\(^7\) The BipoSH spectra, as defined in Eq. (52), restrict us to working with only even-parity BipoSH coefficients \( L + d \) is even due to the vanishing of \( C_{\ell}^{00}(\ell+d) \) otherwise. While most known isotropy-violating phenomena like weak lensing, Doppler boost, non-circular beams, etc., can only produce even-parity BipoSH spectra, measurement of odd-parity BipoSH spectra can be used to test for systematic effects, or to search for the signatures of exotic effects such as the lensing of CMB photons by tensor metric perturbations.
Initially, we revisit the simple phenomenological model of dipole modulation of the CMB sky from Sect. 6.2,
\[ T(\hat{n}) = T_0(\hat{n})(1 + M(\hat{n})), \]  
(53)
where \( T(\hat{n}) \) represents the modulated CMB sky, \( T_0(\hat{n}) \) is the underlying (statistically isotropic) random CMB sky, and \( M(\hat{n}) \) is a dipolar field. The BipoSH coefficients resulting from such a modulation are given by
\[ A_{\ell M}^{(\ell+1)} = \tilde{A}_{\ell M}^{(\ell+1)} + m_{1 M} G_{\ell M}^{(\ell+1)}, \]  
(54)
\[ G_{\ell M}^{(\ell+1)} = \frac{C_\ell}{C_{\ell+1}} \sqrt{\frac{(2\ell+1)(2\ell+3)}{3}} r^{(\ell+1)0}. \]  
(55)
Here \( \tilde{A}_{\ell M}^{(\ell+1)} \) corresponds to the BipoSH coefficients of the unknown, but statistically isotropic, unmodulated CMB field, \( m_{1 M} \) are the spherical harmonic coefficients of the modulation field, and \( C_\ell \) is the best-fit CMB angular power spectrum.

The BipoSH representation further enables an estimate of the modulation field to be made over specific angular scales by windowing regions in multipole space in the sum over multipoles \( \ell \) in Eq. (55). This additional information is important for identifying the origin of the isotropy-breaking signal, which could be either cosmological or due to systematic artefacts.

We perform the analysis for the \( N_{\text{side}} = 2048 \) component separated CMB maps with an apodized version of the common mask at that resolution and reconstruct the modulation signal in independent bins of width \( \Delta \ell = 64 \) up to \( \ell_{\text{max}} = 512 \). The application of the common mask introduces a mean field bias in the BipoSH coefficients derived from the data. This bias is estimated from the FFP8 simulations and subtracted from the derived coefficients. The process of masking induces a coupling between the modulation field and the mask that results in a modification of the spectral shape of the modulation signal by the modified shape function (MSF; see Appendix D for details). Further, the covariance of the bias-subtracted BipoSH coefficients is not easy to derive analytically in this case. To overcome this problem, we consider the diagonal approximation to the covariance matrix and estimate it from simulations.

The results presented in the top panel of Fig. 32 indicate that the dipole modulation signal is most significant in the lowest multipole window \( \ell \in [2, 64] \). Note that the power in the dipole modulation field \( m_1 = [m_{11}^2 + m_{00}^2 + m_{10}^2 - m_{00}^2] / 3 \) is related to the dipole amplitude by \( A = 1.5 \sqrt{m_1 / \pi} \). The best-fit amplitude (A) and direction corresponding to the reconstructed dipole modulation field from this lowest multipole bin is quoted in Table 24 for each component-separation method. Also shown are the corresponding results for the cleaned frequency maps SEVEM-100, SEVEM-143, and SEVEM-217. As expected for signals with a cosmological origin, no evidence for frequency dependence is seen.

Since the amplitude of the dipole modulation field is consistent with zero within 2σ for all of the higher \( \ell \)-bins considered, it is plausible that the simple modulation model in Eq. (53) is inadequate to describe the features seen in the BipoSH spectra and should minimally allow for the amplitude, \( A(\ell) \), of the dipole to depend on CMB multipole, \( \ell \). Although this may appear to be a more complex model, it does not necessarily lack motivation. It is readily conceivable that physical mechanisms that cause a dipolar modulation of the random CMB sky would be
Reassuringly, in PCIS13 it was established that such a signal
strated in Planck Collaboration XXVII (2014) and Appendix B.

Table 25. Doppler boost amplitude ($|\beta|$) and direction in Galactic coordinates derived over the multipole range $\ell \in [640, 1024]$ as evaluated from a BipoSH analysis.

| Method      | $|\beta| \times 10^3$ | Direction $(\ell, b)$ ['] |
|-------------|----------------------|--------------------------|
| SEVEM-100   | $1.24 \pm 0.66$      | (277,40) ± 50            |
| SEVEM-143   | $1.35 \pm 0.56$      | (264,39) ± 39            |
| SEVEM-217   | $1.28 \pm 0.45$      | (257,42) ± 32            |

Notes. The errors are estimated from an identical analysis of a set of 1000 Doppler boosted simulations for each frequency.

would not contaminate a dipole modulation signal up to $\ell_{\text{max}} \approx 700$. We now confirm the Doppler boost signal using the BipoSH methodology.

An equivalent description of the Doppler boost in terms of BipoSH coefficients is given by

$$A_{\ell, \ell_2}^{1M} = \tilde{A}_{\ell, \ell_2}^{1M} + \beta_{1M} G_{\ell, \ell_2}^{1M},$$

$$G_{\ell, \ell_2}^{1M} = \frac{\left[ b_f (G_{\ell, \ell_2} - G_{\ell, \ell_2}^{b}) \right]}{12\pi} \bar{C}_{\ell \ell_2} G_{\ell_1, \ell_2}^{10}.$$ (56)

$$\left[ G_{\ell, \ell_2} \right]_{\ell_1} = \left[ C_{\ell_1} + C_{\ell_2} \right],$$

$$\left[ G_{\ell, \ell_2} \right]_{\ell_1}^b = \left[ C_{\ell_1} + C_{\ell_2} \right] + \left[ C_{\ell_1} - C_{\ell_2} \right] \left[ \ell_1 (\ell_1 + 1) - \ell_2 (\ell_2 + 1) \right] / 2,$$ (57)

$$\beta_{1M} = \int d^3 y_1 \hat{\beta} \cdot \hat{n} = \nu / c$$

where $\beta_{1M}$, as discussed in Appendix D, is adopted with the shape function $G_{\ell, \ell_2}^{1M}$ replaced by the corresponding Doppler boost term given in Eq. (56). Corresponding unboosted CMB simulations were also used, in particular to correct for the mean field bias. However, we use a set of Doppler-boosted simulations in order to estimate the error on the reconstructed Doppler boost vector.

Since it is expected that the low multipole modes of the $A_{\ell, \ell_2}^{1M}$ spectrum are contaminated by the dipolar signal reported previously, in order to monitor the impact of this anomalous signal on the Doppler reconstruction we implement a cumulative analysis using multipoles with a varying $\ell_{\text{min}}$ from 2 to 640 in increments of $\Delta \ell_{\text{max}} = 128$ and a fixed $\ell_{\text{max}} = 1024^8$. The recovered Doppler amplitudes from the three SEVEM frequency cleaned maps as a function of $\ell_{\text{max}}$ are shown in the top panel of Fig. 34, while the lower panel indicates the corresponding direction $\hat{\beta}$ in Galactic coordinates determined from the SEVEM–217 data. Table 25 records the best-fit amplitudes and directions for $\ell \in [640, 1024]$. 8 We fix $\ell_{\text{max}} = 1024$ since at higher $\ell$ values the mismatch between the data and simulation power spectra becomes more important and is a concern for the bias subtraction applied when reconstructing the Doppler boost signal.

scale-dependent and possibly significant only at low wavenumbers. It is also intriguing to note that, although in most cases the amplitude of the modulation dipole is seen at low significance, the directions in the first four bins, $\ell_{\text{42}} \in [2, 64]$, $\ell_{\text{46}} \in [65, 128]$, $\ell_{\text{60}} \in [129, 192]$, and $\ell_{24} \in [193, 256]$, are seen to be clustered together, as shown in the bottom panel of Fig. 32. Note that the lower significance of the modulation for the multipole bins at $\ell > 64$ results in larger errors for their respective directions than the value quoted for the $\ell \in [2, 64]$ bin recorded in Table 24.

We extend our analysis to carry out the dipole modulation reconstruction in cumulative bins up to $\ell_{\text{max}} = 512$, making cumulative increments in the multipole in steps of $\Delta \ell = 64$. The results of this analysis are summarized in Fig. 33.

As noted previously, as a consequence of our motion with respect to the CMB rest frame, the observed CMB map is expected to be statistically anisotropic, as has been demonstrated in Planck Collaboration XXVII (2014) and Appendix B. Reassuringly, in PCIS13 it was established that such a signal...
6.5. Angular clustering of the power distribution

In the Planck 2013 data release we reported a possible deviation from statistical isotropy in the multipole range $\ell = 2$–600, thus confirming earlier findings based on the WMAP data (Hansen et al. 2009; Axelsson et al. 2013). This claim of asymmetry extending to higher multipoles was made only on the basis of the alignment of preferred directions as determined from maps of the power distribution on the sky for specific multipole ranges. In particular, it was found that the directions of the dipoles fitted to such maps in the multipole range $\ell = 2$–600 were significantly more aligned than in simulations. In addition, we showed that the ratio of the power spectra in the two opposite hemispheres defined by the asymmetry axis for $\ell = 2$–600 was not statistically anomalous (as later confirmed over the extended multipole range $\ell = 2$–2000 by Quartin & Notari 2015).

Here, we test for the alignment in the Planck 2015 data set. We adopt the approach for the estimation of the dipole alignment that was described in detail in PCIS13, a brief summary of which follows.

1. Local power spectra are estimated from the data at $N_{\text{side}} = 2048$ for 12 patches of the sky corresponding to the $N_{\text{side}} = 1$ HEALPix base pixels. Only those high-resolution pixels surviving the application of the common mask are included in the analysis. As a consequence of this masking, when patches based on HEALPix pixels with $N_{\text{side}} > 1$ are used, the available sky fraction for those patches close to the Galactic plane is too small for power-spectrum estimation. For most of the analysis, we use the cross-spectra determined from half-mission data sets. Due to a mismatch between the noise level in the data and the simulated maps, the results based on auto-spectra are less reliable and also more prone to other systematic effects than the cross-spectra. We therefore do not consider such results here. The spectra are binned over various bin sizes between $\Delta \ell = 8$ and $\Delta \ell = 32$.

2. For each power spectrum multipole bin, an $N_{\text{side}} = 1$ HEALPix map with the local power distribution is constructed.

3. The best-fit dipole amplitude and direction are estimated from this map using inverse-variance weighting, where the variance is determined from the local spectra computed from the simulations. We do not compute error bars for the direction, but expect this to be accounted for in part by the use of equivalently treated simulations in the clustering analysis.

4. A measure of the alignment of the different multipole blocks is then constructed. In PCIS13, we considered the mean angle between all possible pairs of dipole directions up to a given $\ell_{\text{max}}$. Here, for greater consistency with Sect 6.6, we use the mean of the cosine of the angles, rather than of the angles themselves, between all pairs of dipoles. This effectively corresponds to the Rayleigh statistic (RS) introduced formally in Sect. 6.6, and we will refer to it as such, although it differs by ignoring all amplitude information. Clearly, smaller values of the RS correspond to less clustering.

5. The clustering as a function of $\ell_{\text{max}}$ is then assessed using $p$-values determined as follows. We first construct the RS using all multipoles up to $\ell_{\text{max}}$. The $p$-value is then given by the fraction of simulations with a higher RS than for the data for this $\ell_{\text{max}}$. A small $p$-value therefore means that there are few simulations that exhibit as strong clustering as the data. Note that the $p$-values are highly correlated as the RS is a cumulative function of $\ell_{\text{max}}$.

6. We then define two measures of significance. To achieve this, it is necessary to reduce the 1499 different $p$-values determined for $\ell_{\text{max}} \in [2, 1500]$ to a single measure of clustering. We do this in two different ways, using the mean of these $p$-values, and by finding the minimum of the $p$-values, for both the data and for each available simulation. We then determine the percentage of simulations with (i) a lower mean $p$-value and (ii) a lower minimum $p$-value than the data. Note that these two measures of significance take into account different aspects of the data. Note further that since the RS is cumulative and the $p$-values therefore correlated, different scales are weighted unequally and a detection in the mean

---

Fig. 34. Top: amplitude $|\beta|$ of the Doppler boost from the SEVEM-100, SEVEM-143, and SEVEM-217 maps for different multipole bins determined using a BipoSH analysis. The maximum multipole of each bin is fixed at $\ell_{\text{max}} = 1024$, while $\ell_{\text{min}}$ is incremented from $\ell = 2$ to $\ell = 640$ in steps of $\Delta \ell = 128$. The dashed line corresponds to the actual dipole boost amplitude, $|\beta| = 1.23 \times 10^{-3}$. Bottom: Doppler boost direction $\hat{\beta}$ measured in Galactic coordinates from SEVEM-217. The coloured circles denote $\ell_{\text{max}}$ used in the analysis, while $\ell_{\text{max}} = 1024$ is held fixed. The low-$\ell$ and WMAP-9 directions are identical to those in Fig. 35.
and/or minimum $p$-value may be difficult to interpret and to correct for the multiplicity of tests effect (LEE).

Note that the statistics defined in step 6 above correspond to two choices of what were referred to as “global statistics” in PCIS13 in order to assess the degree to which the significance of the results depends on a specific choice for $\ell_{\text{max}}$. The mean $p$-value over all available $\ell_{\text{max}}$ measures the degree to which clustering is present over large multipole ranges independently of whether the clustering is strongly focused in one given direction. Clearly the $p$-values for different $\ell_{\text{max}}$ are strongly correlated, but if the clustering is present only over a small multipole range, the RS will drop and the corresponding $p$-values will eventually rise. By comparing this value to simulations, we test not only whether the dipole alignment in the data is stronger than in statistically isotropic random simulations, but also whether it is present over larger ranges of multipoles than expected. The minimum $p$-value will give strong detections if there is a strong asymmetry over a limited multipole range or weaker clustering over larger multipole ranges when the clustering is strongly focused in a given direction.

For Commander, NILC, and SEVEM, only 500 simulations are available. However, 5000 simulations are available for SMICA, which allows a better estimate of significance to be determined when the probabilities obtained are very low. In this case, we use half of the 5000 simulations to calibrate the statistic (obtain $p$-values following step 5 above) and the remaining half to determine significance levels (compute the mean and minimum over these $p$-values as a function of $\ell_{\text{max}}$ following step 6). When using 500 simulations, it is necessary to use the same set of simulations to calibrate as well as to obtain probabilities. A related issue with these results is that this set of simulations (corresponding to the first 500 out of the 5000 available for SMICA) are observed to yield higher $p$-values for the clustering angle due to a statistical fluctuation. Another 9 sets of 500 simulations that can be obtained from partitioning the 5000 available SMICA simulations all result in lower $p$-values. As a consequence, we observe that results based on the larger number of simulations often give lower $p$-values than when only 500 simulations are used.

In Fig. 35 we show the dipole directions of the 15 lowest 100-multipole bins for the SMICA map. Here, the binning has been chosen for visualization purposes; in further analysis of the Planck data we use finer $\ell$-intervals. The preferred low-$\ell$ modulation direction determined in Sect. 6.2 is also indicated, along with the WMAP-9 result determined over the range $\ell = 2$ to 600 (Axelsson et al. 2013). The observed clustering of the dipole directions is similar to that shown in figure 27 of PCIS13. Note that differences in masking, foreground subtraction, and residual systematic effects will displace the direction of a given dipole with respect to the previous analysis. Similar behaviour is seen for all of the Planck component-separated maps.

In PCIS13, we calculated the mean angle between all possible pairs of dipole directions determined from maps of the local power in multipole bins of size $\Delta \ell = 16$. Here we test the possible bias arising from such a choice by considering bin sizes between $\Delta \ell = 8$ and $\Delta \ell = 32$ in steps of 2. The lower limit avoids significant bin-to-bin coupling in the power spectra for smaller binnings, whilst the upper limit excludes cases where there are an insufficient number of derived dipoles from which the mean angle can be calculated, this leading to poor statistics. In addition to showing results for each bin size, we also calculate the
Fig. 36. Derived \(p\)-values for the angular clustering of the power distribution as a function of \(\ell_{\text{max}}\), determined for Commander (red), NILC (orange), SEVEM (green), and SMICA (blue), based on 500 simulations. For SMICA, the \(p\)-values based on 2500 simulations are also shown (black). The \(p\)-values are based on the fraction of simulations with a higher \(R\)S, determined over the \(\ell\)-range up to the given \(\ell_{\text{max}}\), compared to the data. The results shown here have been marginalized over bin sizes in the range \(\Delta \ell = 8\) to \(\Delta \ell = 32\).

Table 26. Significance of the angular clustering of the power distribution.

<table>
<thead>
<tr>
<th>Method</th>
<th>Bin size</th>
<th>Mean (p)-value</th>
<th>% (mean)</th>
<th>Min. (p)-value</th>
<th>% (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMICA</td>
<td>8</td>
<td>261/2500</td>
<td>1.60</td>
<td>35/2500</td>
<td>16.2</td>
</tr>
<tr>
<td>SMICA</td>
<td>10</td>
<td>51/2500</td>
<td>0.08</td>
<td>3/2500</td>
<td>2.36</td>
</tr>
<tr>
<td>SMICA</td>
<td>12</td>
<td>75/2500</td>
<td>0.20</td>
<td>1/2500</td>
<td>0.96</td>
</tr>
<tr>
<td>SMICA</td>
<td>14</td>
<td>83/2500</td>
<td>0.16</td>
<td>2/2500</td>
<td>1.52</td>
</tr>
<tr>
<td>SMICA</td>
<td>16</td>
<td>78/2500</td>
<td>0.24</td>
<td>4/2500</td>
<td>2.00</td>
</tr>
<tr>
<td>SMICA</td>
<td>18</td>
<td>51/2500</td>
<td>0.04</td>
<td>1/2500</td>
<td>0.68</td>
</tr>
<tr>
<td>SMICA</td>
<td>20</td>
<td>21/2500</td>
<td>&lt;0.04</td>
<td>1/2500</td>
<td>0.76</td>
</tr>
<tr>
<td>SMICA</td>
<td>22</td>
<td>60/2500</td>
<td>0.08</td>
<td>2/2500</td>
<td>1.24</td>
</tr>
<tr>
<td>SMICA</td>
<td>24</td>
<td>34/2500</td>
<td>0.08</td>
<td>2/2500</td>
<td>1.00</td>
</tr>
<tr>
<td>SMICA</td>
<td>26</td>
<td>38/2500</td>
<td>0.08</td>
<td>1/2500</td>
<td>0.96</td>
</tr>
<tr>
<td>SMICA</td>
<td>28</td>
<td>42/2500</td>
<td>0.20</td>
<td>0/2500</td>
<td>&lt;0.52</td>
</tr>
<tr>
<td>SMICA</td>
<td>30</td>
<td>27/2500</td>
<td>0.20</td>
<td>0/2500</td>
<td>&lt;0.60</td>
</tr>
<tr>
<td>SMICA</td>
<td>32</td>
<td>21/2500</td>
<td>0.04</td>
<td>0/2500</td>
<td>&lt;0.52</td>
</tr>
<tr>
<td>SMICA</td>
<td>marg.</td>
<td>43/2500</td>
<td>&lt;0.04</td>
<td>0/2500</td>
<td>&lt;1.00</td>
</tr>
<tr>
<td>SMICA(^a)</td>
<td>marg.</td>
<td>48/2500</td>
<td>&lt;0.04</td>
<td>1/2500</td>
<td>1.70</td>
</tr>
<tr>
<td>SMICA(^b)</td>
<td>marg.</td>
<td>47/2500</td>
<td>&lt;0.04</td>
<td>1/2500</td>
<td>1.76</td>
</tr>
<tr>
<td>SMICA(^c)</td>
<td>marg.</td>
<td>50/2500</td>
<td>&lt;0.04</td>
<td>0/2500</td>
<td>&lt;0.76</td>
</tr>
<tr>
<td>SMICA(^d)</td>
<td>marg.</td>
<td>254/2500</td>
<td>1.52</td>
<td>34/2500</td>
<td>20.1</td>
</tr>
</tbody>
</table>

Notes: We indicate the actual mean/min \(p\)-value of the data, determined from Fig. 36 and written as a fraction of the number of simulations used to assess the values, together with the percentage of simulations with a lower mean/minimum \(p\)-value than the data. Unless otherwise specified, the numbers are determined from half-mission cross spectra \(C_\ell C_\ell\) for all multipoles in the range \(\ell = 2\)–1500, and for the common mask. \(^a\) Half-ring maps instead of half-mission maps. \(^b\) \(C_\ell (2\ell + 1)\) instead of \(C_\ell\). \(^c\) Restricted to multipoles \(\ell > 100\). \(^d\) Restricted to multipoles \(\ell > 200\).

In agreement with the conclusions in PCIS13, a large degree of alignment is seen at least to \(\ell_{\text{max}} \approx 600\). However, in contrast to the earlier results where the \(p\)-values started increasing systematically for \(\ell_{\text{max}} > 1000\), the current \(p\)-values remain low for \(\ell_{\text{max}} > 750\). The full component-separated maps which have higher resolution and sensitivity are used for the current analysis, instead of the single-frequency foreground-cleaned map (SEVEM-143) used in PCIS13. We note that the results for the updated SEVEM-143 map are consistent with the earlier analysis, both with and without correction for the Doppler modulation. Note also that the SMICA results with improved statistics (based on 2500 simulations) generally show lower \(p\)-values than the corresponding results based on 500 simulations.

Table 26 presents the fraction of simulations with a lower mean/minimum \(p\)-value than in the data for a number of different cases. The table shows probabilities for SMICA with different bin sizes (showing only every second bin size since these are correlated), as well as for the results marginalized over bin sizes. We also show results for the different component-separated maps, results based on half-ring cross-spectra instead of half-mission cross-spectra, and results using a different \(\ell\)-weighting scheme, specifically \((2\ell + 1)C_\ell\) instead of \((\ell + 1)C_\ell\), the former being a measure of the variance of the temperature fluctuations. The table indicates probabilities of approximately 0–2% for most of these cases, although results for the smallest bin size show much less significant results. This could be due to the strong anticorrelations between adjacent bins found for this bin size in those Galactic \(N_{\text{side}} = 1\) patches with very small available sky fraction. For the other bin sizes, these correlations are much weaker. Note that many of the significances based on minimum \(p\)-value are only upper limits. This is due to the fact that the limited number of simulations in some cases results in the lowest minimum \(p\)-value being zero. When the minimum \(p\)-value in the data is zero, we show the percentage of simulations which also have zero as the minimum \(p\)-value. Clearly this fraction is only an upper limit on the real significance.

In order to further investigate the \(\ell\)-dependence of the asymmetry, we follow two approaches from PCIS13. Firstly,
we restrict the analysis to multipoles above a minimum multipole \( \ell_{\text{min}} \). Table 26 indicates that clustering at the <1\% significance level is still found when considering only those multipoles with \( \ell_{\text{min}} \) greater than 100. However, when this limit is increased to 200, no significant clustering is found. We then calculate the RS between pairs of dipoles where one dipole is determined from an \( \ell \)-range above a certain limiting multipole \( \ell_{\text{lim}} \), and the other dipole below this limit. Figure 37 shows the RS as a function of \( \ell_{\text{max}} \) for some selected values of \( \ell_{\text{lim}} \). The \( \ell_{\text{lim}} = 300 \) curve (purple) indicates that dipole directions for \( \ell > 1000 \) are significantly aligned with dipoles for \( \ell < 300 \). Similarly, the \( \ell_{\text{lim}} = 700 \) curve (cyan) indicates that the dipole directions for \( \ell = 700–1000 \) are strongly correlated with the dipole directions for \( \ell < 700 \).

Combining these results, we note that when using only multipoles with (i) \( \ell > 200 \); or (ii) \( \ell < 200 \), no significant clustering is found. The strong clustering significance shown to persist to high multipoles in Fig. 36 must therefore be the result of clustering of the dipole directions between low and high multipoles as supported by Fig. 37. The low \( p \)-values can be explained by the alignment of dipole directions for multipoles extending all the way to \( \ell = 1500 \) correlated with directions for \( \ell < 200 \). The observed asymmetry is therefore not consistent with a model based on dipole modulation or power asymmetry located in one specific multipole range or for one given direction, but rather as a correlation of the dipole directions between \( \ell < 200 \) and \( \ell > 200 \). This correlation with lower multipoles is found to persist all the way to \( \ell_{\text{max}} = 1500 \).

An advantage of the directional analysis performed here is that it focuses on a central issue for tests of deviation from isotropy − whether there is a preferred direction. Indeed, Bunn & Scott (2000) noted that the CMB may exhibit a pattern that cannot be identified from the power spectrum, but which would indicate some non-trivial large-scale structure. Evidence for the close correlation and alignment of directions on different angular scales may present a signature of broken statistical isotropy, since in the standard model, these directions should all be independent random variables. In this context, we do not quote a specific direction for such asymmetry here since our results indicate a clustering of angles between different multipoles, but not necessarily that all multipoles are clustered about one specific direction. However, crucially we have shown that the measured clustering is driven by the correlations of directions between higher and lower multipoles.

Some of the analyses in other sections of the paper focus on dipolar modulation, a specific model for a dipolar power enhancement of the statistically isotropic CMB field towards a preferred direction of the sky, and use methods optimized for the detection of such a signal. While the results of Sect. 6.6 show no detection of the clustering of directions, there is no clear contradiction with the results presented here, since they are based on tests for \( d_{\ell m} \) correlations between different multipoles as expected in the dipolar modulation model. The clustering analysis presented here is a model-independent test for deviations from statistical isotropy which could induce very different correlation structure. It is therefore sensitive to other forms of asymmetry, such as the addition of power in one part of the sky or more general phase correlations.

### 6.6. Rayleigh statistic: QML analysis

Results from Sect. 6.5 and in PCIS13 suggest that, beyond a dipole modulation of power on large angular scales, some form of directional asymmetry continues to small scales. There are also indications from Sect. 6.5 that the directions of dipolar asymmetry are correlated between large and small angular scales. Since the nature of the asymmetry is unknown we use the RS, a generic test for directionality that makes minimal assumptions about the nature of the asymmetry. This statistic has been used both in previous CMB studies (Stannard & Coles 2005) and other areas of cosmology (Scott 1991). In our context, for a statistically isotropic sky this statistic is identical to a three-dimensional random walk. The implementation here incorporates all information pertaining to modulation, not just the direction. The approach in this section differs from that of Sect. 6.5 in the method of reconstructing power, the choice of binning, and the choice of how to weight directions in each bin. Another important difference is that Sect. 6.5 only considers the direction of dipolar asymmetry and does not take into account its amplitude.

The statistic is cumulative and thus narrowing down the specific scales from which a signal may be originating is a non-trivial task. However, it is the case that all statistics that measure this form of asymmetry (dipole modulation or large-scale clustering of power) are in some way cumulative and so we will not worry about this issue any further. Another disadvantage of this approach is that it will generally be less powerful than a test that uses a specific model for the directionality. Again, this is a distinction shared when one compares any non-parametric versus parametric statistic.

The construction of the statistic is as follows.

1. Beginning with the estimator from Eqs. (44) and (45), compute the following binned quantities for the data and simulation:

\[
\hat{X}_{0, \ell} = \frac{6}{f_{10}} \sum_{m} A_{\ell m} \left( T_{\ell m}^* T_{\ell+1 m} - \langle T_{\ell m}^* T_{\ell+1 m} \rangle \right) / \delta C_{\ell+1 m} F_{\ell+1} (\ell + 1),
\]

\[
\hat{X}_{1, \ell} = \frac{6}{f_{11}} \sum_{m} B_{\ell m} \left( T_{\ell m}^* T_{\ell+1 m+1} - \langle T_{\ell m}^* T_{\ell+1 m+1} \rangle \right) / \delta C_{\ell+1 m+1} F_{\ell+1} (\ell + 1).
\]
For each \( \ell \) this computes the coupling of \( \ell \) to \( \ell + 1 \). We emphasize that this is a very natural choice of binning the estimator, since any parameter that is dipole modulated will lead to coupling of \( \ell \) to \( \ell \pm 1 \) modes, albeit with different \( \ell \)-weightings (below we describe why this is not an important issue).

2. Construct a three-dimensional vector out of the three estimators for both the data and the simulations\(^{11} \), as defined by Eqs. (47)–(49).

3. Compute the mean amplitude from simulations and divide all vectors (data and simulations) by this amplitude. This choice ensures that each vector is treated equally, since we have no a priori reason to weight some scales more than others.

4. Add this new vector to the previous vector. If this is the first time going through this process the previous vector is the zero vector.

5. Repeat with \( \ell \to \ell + 1 \). Note that the statistics of this process are identical to a three dimensional random walk.

Given that a dipole modulation amplitude of roughly 3\( \sigma \) significance is known to exist at low \( \ell \) (before a posteriori correction), one would expect a similar level of detection of asymmetry to be determined by the RS. Indeed, we find that asymmetry is present out to \( \ell \approx 240 \). Figure 38 (top) presents the \( p \)-values derived when the RS is computed as a function of \( \ell_{\text{max}} \) from \( \ell = 2 \). The minimum \( p \)-value obtained by the data is 0.1–0.2\%, to be compared to the value of 0.9–1.0\% obtained for the dipole modulation amplitude at \( \ell_{\text{max}} = 67 \). The direction preferred by the data for \( \ell_{\text{max}} \approx 240 \) is \((l, b) = (208^\circ, -29^\circ)\), which is approximately 20\(^\circ\) away from the dipole modulation direction determined to \( \ell \approx 64 \).

We correct for a posteriori statistics using the same procedure as in Sect. 6.3. Specifically, we count how often simulations find asymmetry in the range \( 10 \leq \ell \leq \ell_{\text{max}} \) that is more significant than that found for the data. From Fig. 39 it is clear that generic asymmetry at the significance level found in our CMB sky occurs about 6\% or 8\% of the time (depending on the range of \( \ell \) one decides to search over).

While the PTE here is not very low, it is nevertheless somewhat lower than for the usual dipole modulation test. Hence, it seems worth exploring whether any of this signal comes from higher multipoles. Therefore we compute the RS starting at \( \ell_{\text{min}} = 100 \), to avoid the influence of asymmetry at lower \( \ell \). The lower panel of Fig. 38 presents the corresponding \( p \)-values as a function of \( \ell_{\text{max}} \). There is a striking similarity with the lower panel of Fig. 30. It is clear that, even in the absence of a posteriori correction, we find no significant asymmetry at larger \( \ell \). Hence most of the signal we are seeing in Fig. 38 (top) is due to the usual low-\( \ell \) asymmetry.

We would like to stress that the results here are very similar to the results of the previous section. For each of the statistics used we are simply asking whether there is significant coupling of \( \ell \) with \( \ell \pm 1 \) modes. The details of how to optimally combine these couplings for a given \( \ell \) range depends on whether we are talking about dipole modulation or directionality (or some other related test, e.g., variance asymmetry). These details will change the range of scales over which the strongest signal in the data is found.

\(^{11}\) Note that here we have not specified what \( \delta C_{\ell,1} \) is (it is fully specified by choosing a parameter \( X \) to modulate). This is because we have decided to weight each \( \ell \) equally and thus any strictly positive choice for \( \delta C_{\ell,1} \) will be equivalent, since in step 3 we force the mean length of the vectors at each \( \ell \) to be equal.

---

**Fig. 38.** Rayleigh statistic \( p \)-values determined from the QML analysis as a function of \( \ell_{\text{max}} \) for the Commander (red), NILC (orange), SEVEM (green), and SMICA (blue) data sets, with (top panel) \( \ell_{\text{max}} = 2 \) and (bottom panel) \( \ell_{\text{max}} = 100 \). The general pattern of peaks is very similar to that in Fig. 30. We emphasize that the statistic here is cumulative and as such trends in the curves can be misleading.

**7. Sensitivity of anomalies to enhanced sky coverage**

One of the critical aspects in searching for anomalous features in sky maps is to ensure that the region being investigated constitutes a fair and unbiased sample. Since many of the claimed anomalies are on large angular scales, this implies that minimal masking should be applied to the data. However, residual foregrounds then become a significant consideration. The masks applied to the four component-separated maps studied in the bulk of this paper have been defined at high resolution, and then conservatively degraded for lower resolution studies. Such a procedure inevitably reduces the sky coverage available for analysis, and can be particularly problematic if significant structures are aligned by chance with the masked regions. Indeed, the WMAP team (Bennett et al. 2011) have drawn attention to several such features in their ILC reconstruction of the CMB sky, and these are clearly also present in the Planck Commander, NILC, SEVEM, and SMICA sky maps. A large cold spot is seen
near to the Galactic centre, a significant fraction of which lies within the common mask at any resolution. However, despite its location and visual impression, the feature is neither likely to be attributable to residual foreground emission, nor is it inconsistent with the ΛCDM model (Gott et al. 2007). In addition, four elongated cold fingers stretching from near the Galactic equator to the south Galactic pole are seen, although no equivalent features are evident in the northern sky. Bennett et al. (2011) have noted that the alignment of the ℓ = 2 and ℓ = 3 multipoles (Tegmark et al. 2003) seems to be intimately connected with these large-angular-scale cool fingers and the intervening warm regions. One of the latter also corresponds to the well-known “Bianchi VIIc” main lobe originally found in Jaffe et al. (2005).

Although we would ideally pursue full sky analyses, we prefer to remain mindful of the influence of residual foregrounds, but still seek to minimize the extent of any mask applied for analysis. In this context, and specifically for large-angular-scale studies, we consider the properties of an additional estimate of the CMB sky, also generated using the Commander component separation methodology. In particular, we note that the Planck low-ℓ likelihood analysis (Planck Collaboration XI 2016) uses the temperature solution from this study, degraded to a resolution of Nside = 16. The Lkl-Commander map, as we now refer to it, is initially derived from input data sets (32 bands) at 1° FWHM resolution and Nside = 256. This includes Planck individual detector and detector set maps from 30–857 GHz, the 9-year WMAP observations between 23 and 94 GHz, and the 408 MHz sky survey (Haslam et al. 1982), whereas the Commander map described in Planck Collaboration IX (2016) includes Planck data alone. It is believed that the 32-band solution is better (on large angular scales) than the Planck-only map, because the larger number of input frequencies allows more detailed foreground modelling, and in particular the separation of the low-frequency foregrounds into synchrotron, free-free, and spinning dust components. An associated confidence mask (hereafter LklT16,94) is then defined based on a goodness-of-fit measure per pixel, corresponding to a rejection of 7.3% of the pixels on the sky. A detailed discussion of these results can be found in Planck Collaboration X (2016).

We now consider the implications of using the Lkl-Commander map for studies of several large-angular-scale anomalies observed in previous sections, in particular since the larger sky coverage permitted by this data set should constitute a better sample of the Universe. Note that, at the resolutions of interest for the following analyses, the noise level is negligible (even accounting for the WMAP contribution) and should not have significant impact on the results. The exact details of the noise contribution to simulations is therefore unimportant.

### 7.1. Variance, skewness, and kurtosis

We begin by estimating the variance, skewness, and kurtosis of the CMB. We apply the unit variance estimator to the Lkl-Commander map, and specifically to the version used in the low-ℓ likelihood analysis, which is smoothed to 440° FWHM at a resolution of Nside = 16. A corresponding low-ℓ mask is generated by a simple degrading of the mask at Nside = 256, then setting those Nside = 16 pixels with a value less than 0.5 to zero and all others to unity. The resultant low-ℓ likelihood mask rejects only 6.4% of the sky. We compare the results for both this mask (also to be referred to as LklT16,94), and the standard common mask at this resolution (UT16,58). The results are summarized in Table 27 and show that, when using the low-ℓ likelihood mask, the lower tail probability for the variance is 7.0%. This value is higher than the corresponding values for the component separated maps as shown in Table 12. In addition the skewness and kurtosis are less consistent with Gaussianity than the component separated maps. However, when using the standard common mask at Nside = 16, the lower tail probability of the variance, skewness, and kurtosis become more compatible with those derived earlier.

![Fig. 39. Probability to exceed (PTE) the p-value of the signal from the Commander (red), NILC (orange), SEVEM (green), and SMICA (blue) data at ℓ = 230–240 (which is the multipole range with the most significant deviation) when searching over a range of multipoles up to ℓmax, for the RS determined from the QML analysis. Much like the equivalent curve for dipole modulation, the PTE appears to grow approximately logarithmically with ℓmax.](image)

<table>
<thead>
<tr>
<th>Mask</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>LklT16,94</td>
<td>7.0</td>
<td>1.5</td>
<td>94.0</td>
</tr>
<tr>
<td>UT16,58</td>
<td>0.7</td>
<td>19.9</td>
<td>82.5</td>
</tr>
</tbody>
</table>

Table 27. Lower-tail probability for the variance, skewness, and kurtosis of the Lkl-Commander map.

Note that the SEVEM maps used in this section have been inpainted within 3% of the sky towards the Galactic centre using a simple dif-fracive inpainting algorithm. This prevents residual foreground contamination from propagating to neighbouring regions when downgrading the map. The other component-separated maps are not pre-processed in this way since some form of inpainting of the most contaminated regions was already implemented as part of the component separation algorithms.
some foreground contamination in the regions omitted by application of the UT$_{1658}$ mask, it is appropriate to recall that this was constructed in a conservative way, and may also mask parts of the sky where the level of residual foregrounds can be considered negligible. In addition, we investigate the cleaned frequency maps produced by the SEVEM algorithm in order to test for the presence of frequency-dependent residual foregrounds. The results of the unit variance estimator analysis are summarized in Table 28.

All of the component separated maps show an increase in the lower tail probability from about 0.5% when the UT$_{1658}$ mask is applied to roughly 7% for the LklT$_{64}$94 mask. The small variations in results for the different maps may be attributable to the presence of residual foregrounds close to the Galactic plane. However, the increased probabilities can also be explained by the presence of CMB structures with higher variance within that region which is not rejected by the less conservative mask. Indeed, since the component-separated maps are affected by different residual foregrounds, if the source of the changes in probabilities is due only to the residual foregrounds, then we would expect a larger dispersion than what is observed. We also note that when we apply the low-ℓ likelihood mask the skewness and kurtosis values are shifted towards more extreme values. This implies that the sky signal is less Gaussian for the larger sky fraction, as observed. When the three statistics are considered separately, the fraction of simulations that indicate such changes are 7.6%, 4.3%, and 13.9% for the variance, skewness, and kurtosis, respectively. Of course, such subsets of the simulations also include cases where a large shift in the statistic is observed, but the statistic would not be considered anomalous for either mask. If we also impose the requirement that the simulations have these shifts for all three quantities simultaneously, then only 2 maps from 1000 are found. Of course, such a requirement is rather strong, and at this stage we are likely to be approaching the limits of what can be said based on model-independent null tests. Indeed, in order to assess whether these results are sensitive to a posteriori choices, we repeat the analysis but successively take each simulation as the reference. Thus, for each simulation the shift in the variance, skewness, and kurtosis is computed and then we determine how many times we find a case in which two or less of the remaining simulations simultaneously show larger shifts for the three moments. We find that 48 maps from 1000 satisfy these conditions. Given this, it is difficult to draw strong conclusions about the significance or otherwise of the mask-related changes in variance.

### 7.2. N-point correlation functions

The connection between sky coverage and the observed structure of the 2-point correlation function for large angular separations has previously been discussed in the literature, in particular in connection with the $S_{1/2}$ statistic discussed in Sect. 5.2. Bennett et al. (2011) consider that the use of a Galactic mask when computing these quantities is sub-optimal, and note that a full-sky computation of the 2-point correlation function from the 7-year WMAP ILC map lies within the 95% confidence region determined by simulations of their best-fit ΛCDM model over all angular separations. However, Copi et al. (2009) suggest that the origin of the inconsistencies between the full-sky and cut-sky large-scale angular correlations remains unknown, and that the observed discrepancies may indicate that the Universe is not statistically isotropic on these scales. We therefore consider the N-point correlation functions, and related statistics, of the Lkl–Commander map to contribute to this debate.

We compare results computed for both the Lkl–Commander and Commander maps at $N_{\text{side}} = 64$ after smoothing to a FWHM of 160′. A mask is constructed for the Lkl–Commander map by degrading the LklT$_{64}$93 mask to $N_{\text{side}} = 64$ and setting all resulting pixels with a value less than 0.5 to zero, with the remainder set to unity. The LklT$_{64}$92 mask retains 92% of the sky, to be compared to the 67% usable sky coverage allowed by the UT$_{64}$7 common mask at this resolution.

The results are presented in Fig. 40 where we compare the N-point functions for the data and the mean values estimated from 1000 Commander simulations. The probabilities for obtaining values of the $\chi^2$ statistic for the $Planck$ fiducial ΛCDM model at least as large as the observed values are provided in Table 29. For the estimation of the probabilities, we use the same set of 1000 Commander simulations for both versions of the Commander data. As noted previously, the details of the simulations for such highly smoothed data is essentially unimportant. We also provide an analysis of the Lkl–Commander map using the common mask to enable a direct comparison with the analysis of the Commander map. In this latter case, the results for both maps are in excellent agreement. However, the Lkl–Commander map is more consistent with simulations when the LklT$_{64}$92
Fig. 40. N-point correlation functions determined from the $N_{side} = 64$ Planck CMB 2015 temperature estimates. Results are shown for the 2-point, pseudo-collapsed 3-point (upper left and right panels, respectively), equilateral 3-point, and connected rhombic 4-point functions (lower left and right panels, respectively). The brown three dot-dashed, purple dashed, and red dot-dashed lines correspond to the Lkl-Commander map analysed using the low-$\ell$ and common masks and the Commander map analysed using the common mask, respectively. Note that the dashed and dot-dashed lines lie on top of each other. The black solid line indicates the mean for 1000 MC simulations. The shaded dark and light grey regions indicate the 68% and 95% confidence regions, respectively, estimated using 1000 Commander simulations. See Sect. 4.3 for the definition of the separation angle $\theta$.

mask is adopted for the 2-point and pseudo-collapsed 3-point functions, but less consistent for the equilateral 3-point and rhombic 4-point function results. Nevertheless, the results are generally in agreement with expectations for a Gaussian, statistically isotropic model of the CMB fluctuations.

The increased consistency of the 2-point function with simulations when analysing a larger sky fraction is consistent with the observations in Copi et al. (2009). We therefore quantify this further by determining the statistical quantities introduced in Sect. 5.2 for the Lkl-Commander map. In particular, we reassess the lack of correlation determined previously for large angular scales. It is evident from Table 30 that the results for the $S_{1/2}$ and $\chi_0^2$ statistics are less anomalous when the low-$\ell$ mask is applied. Moreover, the global $p$-value for the $S(x)$ statistic is substantially smaller.

We also repeat the conventional $\chi^2$ analysis but constraining the computations to the two separate ranges defined by

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{1/2}$</td>
<td>97.1</td>
<td>99.5</td>
</tr>
<tr>
<td>$S(x)$ (global)</td>
<td>90.9</td>
<td>97.7</td>
</tr>
<tr>
<td>$\chi_0^2(\theta &gt; 60^\circ)$</td>
<td>95.7</td>
<td>98.1</td>
</tr>
</tbody>
</table>

$\theta < 60^\circ$ and $\theta > 60^\circ$. The results of these studies are shown in Table 31. The analysis for separation angles $\theta > 60^\circ$ indicates that the unusually good fit of the observed 2-point function to the mean 2-point function determined for the $\Lambda$CDM model is
Table 31. Probabilities for obtaining values of the $\chi^2$ statistic for the simulations at least as large as the observed values of the statistic estimated from the Lkl-Commander and Commander maps using the LklT$_{2048}$92 and UT$_{256}$67 masks, respectively.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2(\theta &lt; 60^\circ)$</td>
<td>52.9</td>
<td>91.5</td>
</tr>
<tr>
<td>$\chi^2(\theta &gt; 60^\circ)$</td>
<td>96.5</td>
<td>96.8</td>
</tr>
</tbody>
</table>

Our results do appear to indicate that computations made on larger sky fractions increase the consistency of the 2-point function of the Lkl-Commander map reported in Table 29 is related mainly to correlations in the data for separation angles smaller than $60^\circ$.

Our results do appear to indicate that computations made on larger sky fractions increase the consistency of the 2-point function of the Lkl-Commander map reported in Table 29 is related mainly to correlations in the data for separation angles smaller than $60^\circ$.

Table 32. Probabilities for obtaining values of the $\chi^2$ statistic and ratio of $\chi^2$ of the N-point functions for the Planck fiducial ΛCDM model at least as large as the observed values of the statistic on the northern and southern ecliptic hemispheres estimated from the Lkl-Commander and Commander maps using the LklT$_{2048}$92 and UT$_{256}$67 masks, respectively.

<table>
<thead>
<tr>
<th>Hemisphere</th>
<th>2-point function</th>
<th>Pseudo-collapsed 3-point function</th>
<th>Equilateral 3-point function</th>
<th>Rhombic 4-point function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern</td>
<td>98.7</td>
<td>99.4</td>
<td>99.9</td>
<td>99.8</td>
</tr>
<tr>
<td>Southern</td>
<td>97.4</td>
<td>97.4</td>
<td>97.4</td>
<td>97.4</td>
</tr>
</tbody>
</table>

Notes. The values represent the fraction of simulations with local-variance dipole amplitudes larger than those inferred from the data. No high-pass filtering has been applied to the maps.

Table 33. $p$-values for the variance asymmetry measured by different discs from the Planck 2015 Lkl-Commander and Commander temperature solutions using the LklT$_{2048}$93 and UT$_{256}$73 masks, respectively.

<table>
<thead>
<tr>
<th>p-value [%]</th>
<th>Disc radius [&quot;]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.2</td>
</tr>
<tr>
<td>6</td>
<td>0.4</td>
</tr>
<tr>
<td>8</td>
<td>0.3</td>
</tr>
<tr>
<td>10</td>
<td>0.3</td>
</tr>
<tr>
<td>12</td>
<td>0.8</td>
</tr>
<tr>
<td>14</td>
<td>0.5</td>
</tr>
<tr>
<td>16</td>
<td>0.6</td>
</tr>
<tr>
<td>18</td>
<td>0.8</td>
</tr>
<tr>
<td>20</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Table 34. Local-variance dipole directions measured by 8" discs for the Planck 2015 Lkl-Commander and Commander temperature solutions.

<table>
<thead>
<tr>
<th>Method</th>
<th>$l, b$ [&quot;]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lkl-Comm.</td>
<td>(225, –28)</td>
</tr>
<tr>
<td>Commander</td>
<td>(214, –24)</td>
</tr>
</tbody>
</table>

7.3. Dipole modulation and directionality

7.3.1. Variance asymmetry

7.3.2. Dipole modulation: pixel-based likelihood

Table 35 presents constraints on the dipole modulation model as derived from the Lkl-Commander map and the LklT$_{2048}$93 mask that includes 93% of the sky, updating the results from Sect. 6.2.
5.8% and the small 0.5% shift from the Commander

The former results were derived using the

Notes. The former results were derived using the LklT_256_93 mask, whereas the latter are those determined previously in Sect. 6.2.

Table 36. Summary of the dipole modulation results for the range \( \ell \in [2, 64] \) determined from the Planck 2015 Lkl-Commander and Commander temperature solutions, as derived by the QML estimator defined in Sect. 6.3 using the LklT_256_93 and UT78 masks, respectively.

<table>
<thead>
<tr>
<th>Method</th>
<th>A (( \mu K ))</th>
<th>Direction (( \ell, b ) [( ^\circ ])]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lkl-Comm. . .</td>
<td>0.058 ( \pm ) 0.022</td>
<td>(227, ( -28 )) ( \pm ) 26</td>
</tr>
<tr>
<td>Commander . .</td>
<td>0.063 ( \pm ) 0.025</td>
<td>(213, ( -26 )) ( \pm ) 28</td>
</tr>
</tbody>
</table>

for the Commander map. We find that all previously reported results are robust with respect to data selection and sky coverage. In particular, the best-fit dipole modulation amplitude at 5° FWHM is 5.9% in the Lkl-Commander map, and is thus stable to within about 0.3\( \sigma \) when increasing the sky fraction from 78% to 93%. Likewise, the marginal low-\( \ell \) power spectrum amplitude, \( q \), shifts upward by 0.4\( \sigma \), and the power spectrum tilt, \( n \), downward by 0.3\( \sigma \), for the same sky fraction increases.

To assess the statistical significance of these shifts, we compare with Gaussian statistics, creating two Gaussian random vectors with 78 and 93 elements, respectively, where the first 78 elements of the latter vector are identical to the first vector. From these, we compute the difference between the two means, after normalizing each so that their individual errors in the mean are unity. Repeating this simple calculation \( 10^3 \) times, we find that 48% of all Gaussian realizations observe shifts larger than 0.3\( \sigma \), and 34% observe shifts larger than 0.4\( \sigma \). Thus, the parameter differences due to the different data selection and sky fractions reported above are consistent with expectations from random Gaussian statistics.

7.3.3. Dipole modulation: QML analysis

We also repeat the QML dipole modulation analysis of Sect. 6.3 for the Lkl-Commander map and corresponding mask. Table 36 summarizes the results of the low-\( \ell \) dipole modulation for the Lkl-Commander temperature solution, compared with the Commander map.

The best-fit modulation amplitude for Lkl-Commander is 5.8% and the small 0.5% shift from the Commander best-fit amplitude corresponds to a decrease of approximately 0.4\( \sigma \). These results mirror very closely the results found above for the pixel-based likelihood approach to dipole modulation, as expected, and the observed shifts are perfectly consistent with those expected from the change in the mask.

7.3.4. Bipolar spherical harmonics

We next perform a dipole modulation analysis on the Lkl-Commander temperature map using the BipoSH formalism from Sect. 6.4. The dipole modulation amplitude inferred from the analysis is smaller than that deduced from analysing the Commander map as seen in Table 37. However, it should be noted that the probability for simulations to yield a dipole modulation amplitude equal to or greater than the amplitude inferred from data is 0.4%, which is smaller by a factor of approximately 2.4 as compared to the \( p \)-value inferred from analysis on Commander. The reduction in the dipole amplitude and the enhanced significance can both be attributed to the reduced power bias which is a result of the increased sky coverage.

7.4. Summary

Using a larger sky fraction in our analyses leads to small changes in the results related to large-angular-scale anomalies, but these are essentially consistent with expectations from random Gaussian statistics. In particular, the asymmetry in power on the sky, as parameterized by a dipole modulation model, is robust to mask changes.

8. Polarization analysis

As previously discussed in Sect. 2, large angular-scale CMB fluctuations in the Planck polarization data have been suppressed by a post-processing high-pass filter to minimize the impact of systematic artefacts. Therefore, no polarization results concerning CMB statistical anomalies on such scales are presented in this paper. In addition, a noise mismatch between simulations and data also limits our ability to study polarization more generally. Nevertheless, a local analysis of the polarization data for stacked patches of the sky can still be performed, in order to test the statistical properties of the CMB anisotropies. In this case, the stacking procedure mitigates the impact of the small-scale noise and potential systematic effects.

Traditionally, the Stokes parameters \( Q \) and \( U \) are used to describe the CMB polarization anisotropies (e.g., Zaldarriaga & Seljak 1997). Such quantities are not rotationally invariant, thus for the stacking analysis it is convenient to consider a local rotation of the Stokes parameters, resulting in quantities denoted by \( Q_r \) and \( U_r \), as described in Sect. 8.1. Additionally, several other related quantities can be defined.
The polarization amplitude \( P \equiv \sqrt{Q^2 + U^2} \) and polarization angle \( \Psi \equiv \frac{1}{2} \arctan(U/Q) \), are commonly used quantities in, for example, Galactic astrophysics. However, unbiased estimators of these quantities in the presence of anisotropic and/or correlated noise are hard to define (Plaszczynski et al. 2014). Of course, a direct comparison of the observed (noise-biased) quantity to simulations analysed in the same manner is possible, but we elect here to defer the study of this representation of the polarization signal, using maps of the polarization amplitude only to define peaks around which stacking can be applied.

The rotationally invariant quantities referred to as \( E \) and \( B \) modes are commonly used for the global analysis of CMB data. Although the \( E \)-mode maps are not analysed in detail here, they are considered qualitatively, so that it is appropriate to recall their construction. Since the quantities \( Q \equiv iU \), defined relative to the direction vectors \( \hat{n} \), transform as spin-2 variables under rotations around the \( \hat{n} \) axis, they can be expanded as

\[
(Q \pm iU)(\hat{n}) = \sum_{\ell,m} \sum_{n=\ell}^{\infty} a_{\ell m}^{(\pm 2)} Y_{\ell m}(\hat{n}),
\]

where \( a_{\ell m}^{(\pm 2)} \) are the spin-weighted spherical harmonics and \( a_{\ell m}^{(\pm 2)} \) are the corresponding harmonic coefficients. If we define

\[
a_{\ell m}^{F} = \frac{1}{2} (a_{\ell m}^{(2)} + a_{\ell m}^{(-2)}), \quad a_{\ell m}^{B} = \frac{1}{2} (a_{\ell m}^{(2)} - a_{\ell m}^{(-2)}),
\]

then the invariant quantities are given by

\[
E(\hat{n}) = \sum_{\ell,m} \sum_{n=\ell}^{\infty} a_{\ell m}^{F} Y_{\ell m}(\hat{n}), \quad B(\hat{n}) = \sum_{\ell,m} \sum_{n=\ell}^{\infty} a_{\ell m}^{B} Y_{\ell m}(\hat{n}).
\]

### 8.1. Stacking around temperature hot and cold spots

The stacking of CMB anisotropies around peaks (hot and cold spots) on the sky yields characteristic temperature and polarization patterns that contain valuable information about the physics of recombination (Komatsu et al. 2011). Statistical analysis of stacked images differs from the other tests in this paper in several respects. First, peak-related new physics may be revealed that is difficult to find in a global analysis, for example, the non-Gaussian CMB cold spots predicted by a modulated preheating model (Bond et al. 2009). Secondly, stacking is a local operation, which naturally avoids mask-induced complications. Thus stacking can be used as a transparent and intuitive method to test the robustness of anomalies found with other methods. Alternatively, it can be applied as a quality indicator of the data at the map level.

Our stacking procedure is as follows. Hot (or cold) peaks are selected in the temperature map as local extrema with negative (or positive) second derivatives, and classified relative to a given threshold \( \nu \) (in rms units of the temperature map). Since the spinorial components \( Q \) and \( U \) are expressed in a local coordinate system, we employ a configuration in which the Stokes parameters around a peak at \( \theta_0 \) can be superposed (Kamionkowski et al. 1997). In particular, we use a locally defined rotation of the Stokes parameters that is written as:

\[
Q_c(\hat{n}; \theta_0) = -Q(\hat{n}) \cos(2\phi) - U(\hat{n}) \sin(2\phi),
\]

\[
U_c(\hat{n}; \theta_0) = Q(\hat{n}) \sin(2\phi) - U(\hat{n}) \cos(2\phi),
\]

where \( \phi \) is the angle between the axis aligned along a meridian (pointing to the south by convention) in the local coordinate system centred on a peak at \( \theta_0 \) and the great circle connecting this peak to a position \( \hat{n} \). This definition decomposes the linear polarization into radial (\( Q_c > 0 \)) and tangential (\( Q_c < 0 \)) contributions around the peaks. This definition of \( Q_c \) is equivalent to the “tangential shear” used in weak lensing studies.

For visualization purposes, a flat patch around each peak is then extracted, and the average stacked image computed from the subset. A position on the sky at an angular distance \( \theta \) from the central peak is labelled with the flat-sky coordinates

\[
x = \sigma \cos \phi, \quad y = \sigma \sin \phi.
\]

Here \( \sigma = 2 \sin(\theta/2) \approx \theta \) is the effective flat-sky radius. For the angular scales of a few degrees considered in the stacking analyses the difference between \( \sigma \) and \( \theta \) is negligible. We use \( \sigma \) for analyses in the flat-sky approximation, and \( \theta \) for analyses directly on the sphere.

The stacking process tends to provide an image with azimuthal symmetry about its centre, due to the almost uncorrelated orientations of the temperature peaks. The stacked images of temperature patches around hot spots selected above the null threshold for both the Commander data and a corresponding simulation are shown in the top row of Fig. 41. The observed patterns are in excellent agreement. Stacking around cold spots yields similar patterns but with flipped sign. Given the symmetry, it is often useful to consider the radial profile obtained by averaging the stacked image over the azimuthal angle \( \phi \). Figure 42 shows such a profile determined from the stacked temperature image.

At this point, it is useful to consider the underlying physics represented by the various patterns in the stacked images. During recombination, the sound horizon extends an angle \( \theta_s = r_s/D_A \approx 0.011 (0.61^\circ) \), where \( r_s \approx 0.15 \) Gpc is the size of the sound horizon at recombination and \( D_A \approx 14 \) Gpc is the angular-diameter distance to the last scattering surface. To understand the ring patterns in the stacking image, projection effects must be taken into account. Firstly, all 3D modes with wavenumber \( k \geq \ell/D_A \) contribute to a 2D \( \ell \)-mode. More modes contribute to, and therefore enhance, the power at lower \( \ell \). For the first acoustic peak, the net effect is a \( \pi/4 \) phase shift towards lower \( \ell \), such that \( \ell_s \approx (\pi - \pi/4)/\theta_s \approx 220 \). The projected acoustic scale on the temperature map is of order \( \theta_s^{2D} = \pi/\ell_s \approx 0.014 (0.81^\circ) \). Secondly, the stacked 2D modes around peaks interfere with each other. The first dark ring appears at \( 1.22 \theta_s^{2D} = 0.017 (1.0^\circ) \). The factor 1.22 is the ratio of the first minimum of the projection kernel, the Bessel function \( J_0 \), to the first minimum of the unprojected cosine wave.

The dark ring can also be regarded as a consequence of the correlation between \( T \) and \( -\nabla^2 T \). At the temperature maxima \( -\nabla^2 T \) is positive, with an amplitude of order \( T_{\text{peak}}/(\theta_s^{2D})^2 \). Thus, the quadratic terms in the local expansion of the temperature field have a negative contribution that grows as \(-T_{\text{peak}}(\theta_s^{2D})^2\). At \( \sigma \approx \theta_s^{2D} \) the quadratic terms dominate and the \( T-(\nabla^2 T) \) correlation becomes negative. Meanwhile, the \( T-(\nabla^2 T) \) correlation tends to zero on the scale \( \sigma \approx \theta_s^{2D} \), where the temperature autocorrelation becomes weak and the local quadratic expansion starts to fail. As shown in Fig. 42, the dark ring appears at the critical point where the \( T-(\nabla^2 T) \) correlation reaches its minimum and turns back toward zero.

We have discussed the projection effects that make the projected radial acoustic scale on a stacked \( T \) image larger than \( \theta_s \). For \( Q_c \), the most striking patterns in the image have
more intuitive simple explanations, since the stacking is essentially the real-space equivalent of the temperature polarization
correlation. The projection function contains an extra $\ell^2$ factor, which enhances the high-$\ell$ power and reduces the projected radial acoustic scale, coincidentally, back to $\approx \theta_s$. The quadrupole responsible for the polarization around peaks is induced by gravity on angular scales larger than twice the size of the horizon at decoupling. In the case of an overdensity, this causes a flow of photons towards the gravitational well on these scales, inducing a quadrupolar pattern (see, e.g., Coulson et al. 1994). The spherical symmetry of the gravitational interaction causes a rotation of the quadrupole in the vicinity of the well, resulting in a radial configuration in polarization. This radial polarization pattern implies $Q_r > 0$ and an overdensity implies $T < 0$ by the Sachs-Wolfe formulae, which leads to anticorrelation on these scales. Similarly, an underdensity leads to an outward flow and induces a tangential polarization pattern, once again leading to anticorrelation on these scales. At smaller scales, the polarized contribution is dominated by the dynamics of the photon fluid. The acoustic oscillations modulate the polarization pattern, leading to the different rings in the stacked images. The most noticeable rings in the stacked $Q_r$ image are approximately at $\theta_s$ and $26\theta_s$. Thanks to the $\ell^2$ enhancement, multiple acoustic peaks in the $TE$ power spectrum may be captured and projected into ring patterns in the stacked polarization images. As photons flow towards the overdensity, they are compressed and the temperature increases, slowing the fluid descent into the well. Eventually, the radiation pressure becomes large enough to reverse the photon flow. This expansion cools the photons until they fall back towards the well. Note that the resulting inner ring was not observed in the WMAP analysis (Komatsu et al. 2011), since the resolution was too low.

Figure 41 clearly reveals all of the features described above. The two bright rings at $\theta_s = 0.011 (0.6')$ and $26\theta_s = 0.021 (1.2')$ are the predicted patterns associated with the first $C_2^{LE}$ acoustic
peak at $\ell \approx 310$, while the two faint rings are a striking illustration of the detection of multiple acoustic peaks in the $TE$ power spectrum. The large-scale anticorrelation is suppressed due to the scale-dependent bias which results from the fact that peaks are defined by the second derivatives of the temperature field (e.g., Desjacques 2008).

We are now in a position to discuss the consistency of the Plenck results with the predictions of a $\Lambda$CDM cosmology. For simplicity, further analysis is focused on the angular profiles, and specifically the mean, $\bar{\mu}(\theta)$, estimated as the average of the angular profiles around all hot (cold) peaks above (below) a certain threshold $\nu$. This analysis is performed directly on the sphere to avoid any repixelization error. Note that the expected value of the mean temperature angular profile is proportional to $\int \ell \, d\ell \, C_T^{\ell \ell} J_0(\ell \theta)$, while the expected values of the $Q_\ell$ and $U_\ell$ mean angular profiles are approximately proportional to $\int \ell \, d\ell \, C_T^{\ell \ell} J_0(\ell \theta)$ and $\int \ell \, d\ell \, C_Q^{\ell \ell} J_0(\ell \theta)$, respectively. Since $T$ has even parity and $B$ has odd parity, the expectation value for $C_Q^{\ell \ell}$ is zero, and the $U_\ell$ mean angular profile is therefore expected to vanish.

A $\chi^2$ estimator is used to quantify the differences between the profiles obtained from the data and the expected values estimated with simulations:

$$\chi^2 = (\bar{\mu}(\theta) - \tilde{\mu}(\theta)) C^{-1} (\bar{\mu}(\theta) - \tilde{\mu}(\theta))^T,$$

with the covariance matrix defined as

$$C(i,j) = \frac{1}{N-1} \sum_{k=1}^{N} \left[ \mu_i(\theta_k) - \tilde{\mu}(\theta_k) \right] \left[ \mu_j(\theta_k) - \tilde{\mu}(\theta_k) \right],$$

where the sum is over the $N$ simulations used to estimate this matrix and $\tilde{\mu}(\theta)$ is the ensemble average. Note that although the profiles in Fig. 41 are derived from data at a resolution $N_{\text{side}} = 1024$, faster convergence of the $\chi^2$ statistic is achieved using maps at a lower resolution. We have verified that the results remain unchanged when adopting data with $N_{\text{side}} = 512$.

Figure 43 presents a comparison between the profiles obtained from the component-separated data and the mean value estimated from simulations processed through the SEVEM pipeline. Note that the error bars for the temperature profiles are asymmetric due to a bias in the selection of the peaks above a given threshold. Results for hot and cold spots are shown for two different thresholds, $\nu = 0$ and $\nu = 3$. There is generally excellent agreement between the different component-separation methods. A systematic deviation between the data and the simulations for the hot peaks in temperature ($\nu = 0$) is seen at a level greater than $1\sigma$. This discrepancy increases at higher thresholds, reaching values of about $2\sigma$ for the $\nu = 3$ case. Similar behaviour is seen for the cold spots. For the $Q_\ell$ angular profiles, the most striking differences appear around $\theta = 2^\circ$ in the $\nu = 3$ case for hot peaks, and around $\theta = 1.5^\circ$ for the cold peaks. For the $U_\ell$ angular profiles, where a null signal is expected (i.e., only noise is expected to be present), deviations at similar levels are seen.

Table 38 presents the corresponding $p$-values for this comparison. A theoretical $\chi^2$ distribution is used to determine the probability that a sky drawn from the $\Lambda$CDM cosmology has a value larger than that derived from the data. We have verified this approach by comparing the empirical $\chi^2$ distribution estimated from 100 simulations (in which the mean value and the covariance matrix are computed from a further 900 simulations) with the theoretical distribution with the corresponding degrees of freedom (see Fig. 44). The $\chi^2$ value of the data is then estimated using the mean value and the covariance matrix determined from simulations. Although some differences are found among the
component-separation methods, a general consistency between model and data is found. Although the \( \chi^2 \) test has the advantage of being sensitive to different types of deviations between model and data, does not assume prior knowledge about possible departures from the model, and can account for correlations between the various tests from which it is constructed, it can nevertheless be suboptimal under certain conditions. This appears to be the case when considering the systematic shift between data and simulations seen in the temperature profiles \( \mu_T \) — the \( \chi^2 \) statistic is not particularly sensitive to systematic deviations of constant sign. We therefore consider an alternative quantity, the integrated profile deviation, defined as

\[
\Delta \mu_T(W) = \int_0^R [\mu_T(\theta) - \bar{\mu}_T(\theta)] W(\theta) \, d\theta, \tag{72}
\]

Table 38. \( p \)-values of the \( T \), \( Q \), and \( U \) angular profiles computed from the stacking of hot and cold spots selected above the \( v = 0 \) and \( v = 3 \) thresholds.

<table>
<thead>
<tr>
<th>Probability [%]</th>
<th>Comm.</th>
<th>NILC</th>
<th>SEVEM</th>
<th>SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v = 0 ) (hot spots)</td>
<td>( T )</td>
<td>8</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>( Q )</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>( U )</td>
<td>93</td>
<td>28</td>
<td>75</td>
<td>44</td>
</tr>
<tr>
<td>( v = 3 ) (hot spots)</td>
<td>( T )</td>
<td>18</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>( Q )</td>
<td>34</td>
<td>23</td>
<td>31</td>
<td>19</td>
</tr>
<tr>
<td>( U )</td>
<td>28</td>
<td>61</td>
<td>21</td>
<td>50</td>
</tr>
<tr>
<td>( v = 0 ) (cold spots)</td>
<td>( T )</td>
<td>23</td>
<td>38</td>
<td>29</td>
</tr>
<tr>
<td>( Q )</td>
<td>86</td>
<td>85</td>
<td>63</td>
<td>78</td>
</tr>
<tr>
<td>( U )</td>
<td>14</td>
<td>11</td>
<td>39</td>
<td>34</td>
</tr>
<tr>
<td>( v = 3 ) (cold spots)</td>
<td>( T )</td>
<td>24</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td>( Q )</td>
<td>21</td>
<td>51</td>
<td>29</td>
<td>52</td>
</tr>
<tr>
<td>( U )</td>
<td>30</td>
<td>13</td>
<td>30</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 39. \( p \)-values of \( \Delta \mu_T \) computed from the stacking of hot and cold spots selected above the \( v = 0 \) and \( v = 3 \) thresholds from the Commander, NILC, SEVEM, and SMICA maps.

<table>
<thead>
<tr>
<th>Probability [%]</th>
<th>Comm.</th>
<th>NILC</th>
<th>SEVEM</th>
<th>SMICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot spots</td>
<td>( T ) (( v = 0 ))</td>
<td>96.0</td>
<td>95.8</td>
<td>96.2</td>
</tr>
<tr>
<td>( Q )</td>
<td>98.6</td>
<td>98.2</td>
<td>98.3</td>
<td>98.7</td>
</tr>
<tr>
<td>( U )</td>
<td>97.1</td>
<td>96.9</td>
<td>98.1</td>
<td>97.9</td>
</tr>
<tr>
<td>Cold spots</td>
<td>( T ) (( v = 0 ))</td>
<td>92.0</td>
<td>90.6</td>
<td>90.6</td>
</tr>
</tbody>
</table>

In what follows, we use the \( N_{\text{side}} = 1024 \) component-separated maps at a resolution of 10' FWHM. The use of this higher resolution as compared to the \( N_{\text{side}} = 512 \) data used in Sect. 8.1 is motivated by the smaller-scale features that are expected to result from the oriented stacking.

We also introduce the concept of the noise-free ensemble average (NFEA), which is defined as the ensemble average of stacked CMB-only maps for a fiducial cosmology. Recall that the fiducial model for the simulated sky maps, the Planck 2013

\[ W(\theta) = \sum_{i=1}^{N_{\text{side}}} W_i(\theta) \]

where \( R \), the size of stacking patches, is taken to be 3.5 in this case. The weighting function is chosen to be proportional to the expected profile, but the results are robust for other choices, e.g., \( W = 1 \). The \( p \)-values obtained in this case are given in Table 39. These are consistent with what might be expected from visual inspection of the plots, i.e., the deviations are typically close to \( 2\sigma \). These deviations are likely to be connected to the deficit in the observed power spectrum at low multipoles, as may be seen in Fig. 42. Here, the purple dashed line indicates the reduction in \( \bar{\mu}_T \) if the theoretical \( C_{TT} \) values are reduced by 10% over the range \( 2 \leq \ell \leq 50 \).

\[ \chi^2 \]

8.2. Generalized stacking

In this section, a much wider class of stacking methods is introduced, with particular emphasis on oriented stacking, a novel approach that has not previously been explored in the literature. We regard the stacking as oriented if the orientation of the local coordinate frame, and in particular the \( \phi = 0 \) axis, is correlated with the map that is being stacked. Thus, the stacking methodology in Sect. 8.1 is considered unoriented, because the orientation is defined relative to the local meridian pointing towards the Galactic south, rather than any property of the data themselves. Alternative approaches to unoriented stacking can also be considered. In this subsection, the orientation of each patch is chosen randomly from a uniform distribution in \([0, 2\pi])\). The unoriented \( T \) and \( Q \) images can then be directly compared with previous sections.

For unoriented stacking, the ensemble average of stacked fields cannot result in any intrinsic \( \phi \)-dependence, as this would be averaged out by the uncorrelated orientation choices. The \( \phi \)-dependence due to a specific choice of representation can always be removed via a local rotation. For example, the ensemble averages of \( Q + iU \) around unoriented temperature peaks are proportional to \( e^{i \phi} \). A local rotation \((Q, U) \rightarrow (Q', U')\) (Kamionkowski et al. 1997) removes the \( e^{i \phi} \) factor and compresses the information into a single real map \( Q' \). For oriented stacking, the \( \phi \)-dependence can be a mixture of a few Fourier modes (\( e^{im\phi} \), for integer \( m \)). Each \( m \) mode corresponds to a radial (\( \sigma \)-dependent) function.

In what follows, we use the \( N_{\text{side}} = 1024 \) component-separated maps at a resolution of 10' FWHM. The use of this higher resolution as compared to the \( N_{\text{side}} = 512 \) data used in Sect. 8.1 is motivated by the smaller-scale features that are expected to result from the oriented stacking.

We also introduce the concept of the noise-free ensemble average (NFEA), which is defined as the ensemble average of stacked CMB-only maps for a fiducial cosmology. Recall that the fiducial model for the simulated sky maps, the Planck 2013

\[ \chi^2 \]
best-fit ΛCDM model (Planck Collaboration XVI 2014), differs from the updated Planck 2015 best-fit ΛCDM model (Planck Collaboration XIII 2016). In previous sections, this mismatch was partially accommodated by rescaling the CMB signal by a fixed scale factor. Here, we instead specifically adopt the 2015 best fit as a fiducial model for the data. When comparing the data to the simulations, we subtract the corresponding NFEA to minimize any bias resulting from cosmology dependence.

In the context of random Gaussian fields, the NFEA can be computed straightforwardly following Bardeen et al. (1986):

$$\langle M \rangle = \left( \mathcal{M} W^H (\omega \omega^T)^{-1} \omega \right),$$  \hspace{1cm} (73)

where $M$ is the map (around the central peak) to be stacked, and $\omega$ is the collection of Gaussian variables (on the central peak) that are related to peak selection and orientation determination. Equation (73) is only valid for Gaussian random variables. If the patch is rotated before stacking, the field value evaluated at a dynamic coordinate is, in general, not a random Gaussian variable. However, statistical isotropy guarantees that the rotation of patches is equivalent to an orientation constraint on the non-zero-spin field. For example, orienting each patch in the direction where $U = 0$ and $Q > 0$ is equivalent to the unoriented stacking case where only peaks satisfying the additional constraint $-\epsilon/2 < \arg(Q + iU) < \epsilon/2$ ($\epsilon \to 0^+$) are selected.

A further source of statistical bias can arise from noise mismatch between the simulations and the data. Since the effect of noise is to introduce random shifts in the peaks and hence suppress patterns in the stacked images, any noise mismatch can lead to pattern mismatch between the data and simulations. For the temperature data, the contribution due to noise mismatch is estimated to be at the sub-percent level, lower than the cosmic variance. For stacking on polarization peaks, the impact of the noise mismatch cannot be safely ignored. Thus, for quantitative comparisons in this paper, we only consider stacking on temperature peaks.

### 8.2.1. Oriented temperature stacking

The most straightforward way to orient a patch centred on a temperature peak is to align the horizontal axis with the major axis defined by a local quadratic expansion of the temperature field around the peak. The disadvantage of doing so is that the orientation is dominated by small-scale fluctuations that are noise-sensitive. A better choice is to make use of the inverse Laplacian $\nabla^2 T$ that filters out the small-scale power. The inverse Laplacian is defined as:

$$\nabla^2 T (\hat{\nu}) = -\sum_{l=1}^{\infty} \sum_{m=-l}^{l} \frac{a_{lm}^T}{l(l+1)} Y_{lm} (\hat{\nu}),$$  \hspace{1cm} (74)

where $a_{lm}^T$ are the harmonic coefficients of the masked temperature map. Spin-2 maps $Q_T, U_T$ are then defined by:

$$(Q_T \pm iU_T) (\hat{\nu}) = \sum_{l=2}^{\infty} \sum_{m=-l}^{l} a_{lm}^T [\pm 2] Y_{lm}(\hat{\nu}) .$$  \hspace{1cm} (75)

In the flat-sky limit, $Q_T \approx (\hat{\nu}^2 - \hat{\nu}_0^2)(\nabla^2 T)$ and $U_T \approx -2\hat{\nu}_0 \hat{\nu}^2 \nabla^2 T$. To align the $\nabla^2 T$ axes of the patches, we rotate each patch so that $U_T$ vanishes and $Q_T \geq 0$ for the central peak.

Figure 45 presents the stacked images of SMICA temperature patches centred on temperature hot spots selected above the threshold $\nu = 0$, in both unoriented and oriented forms. These are seen to be in excellent agreement with their accompanying NFEAs, and, in the case of the unoriented stacks, with the results shown in Fig. 41, despite the different stacking methodologies adopted (and component separation method selected for visualization purposes).

The oriented $T$ image is notably different from the unoriented one. The alignment between the major axis (of $\nabla^2 T$) and the horizontal axis results in an ellipse elongated along the horizontal axis, rather than a central disc. Moreover, the quadratic-term contribution is suppressed along the horizontal axis where the temperature profile is smoother, and enhanced along the vertical axis where the temperature profile is sharper. As a consequence, the dark ring visible in the upper panel at $1^\circ$ splits into two cold blobs along the vertical axis.

To proceed with a quantitative analysis, we extract Fourier modes $T_m(\sigma)$ from the stacked map $T_{\text{stack}}(\sigma, \phi)$ as follows:

$$T_m(\sigma) = \frac{1}{(1 - \delta_{m0})^2} \int_0^{2\pi} T_{\text{stack}}(\sigma, \phi) \cos m\phi \, d\phi,$$  \hspace{1cm} (76)

where $\delta_{m0}$ is the Kronecker delta function. For odd $m$, the NFEA $\langle T_m \rangle$ vanishes due to statistical isotropy. For even $m$, a straightforward calculation shows that only $T_0(\sigma)$, which is equivalent to $\mu_T(\sigma)$, and $T_2(\sigma)$ have nonzero NFEAs.

As discussed previously in Sect. 8.1, there are some shortcomings of the standard $\chi^2$ procedure that is generally used to assess the consistency of the data with simulations. The problem is simplified by studying the statistics of an integrated profile deviation:

$$T_m(W) = \int_0^R [T_m(\sigma) - \langle T_m(\sigma) \rangle] W(\sigma) \, d\sigma,$$  \hspace{1cm} (77)

where $R$, the size of the stacking patches, is taken to be $2^\circ$ in our examples. The purpose of removing the NFEA, $\langle T_m(\sigma) \rangle$, which
Table 40. $T_m$, as defined in Eqs. (77) and (78), for different thresholds $\nu$.

<table>
<thead>
<tr>
<th>Method</th>
<th>Hot spots</th>
<th>Cold spots</th>
<th>Hot spots</th>
<th>Cold spots</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_0$</td>
<td></td>
<td>$T_2$</td>
<td></td>
</tr>
<tr>
<td>Threshold $\nu = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>-0.03(0.04+0.06+0.15)</td>
<td>-0.04(0.04+0.06+0.16)</td>
<td>0.07(0.04+0.01+0.05)</td>
<td>0.06(0.04+0.01+0.03)</td>
</tr>
<tr>
<td>NILC</td>
<td>-0.04(0.04+0.06+0.15)</td>
<td>-0.05(0.04+0.06+0.16)</td>
<td>0.06(0.03+0.01+0.02)</td>
<td>0.05(0.03+0.01+0.03)</td>
</tr>
<tr>
<td>SEVEM</td>
<td>-0.03(0.04+0.06+0.16)</td>
<td>-0.04(0.04+0.06+0.16)</td>
<td>0.06(0.04+0.01+0.03)</td>
<td>0.04(0.04+0.02+0.04)</td>
</tr>
<tr>
<td>SMICA</td>
<td>-0.03(-0.01+0.03+0.07)</td>
<td>-0.04(0.04+0.03+0.07)</td>
<td>0.06(-0.04+0.02-0.04)</td>
<td>0.04(0.04+0.02-0.04)</td>
</tr>
</tbody>
</table>

| Threshold $\nu = 1$ |          |            |          |            |
| Commander | -0.06(0.05+0.09+0.22) | -0.06(0.05+0.09+0.21) | 0.06(0.03+0.02+0.03) | 0.04(0.03+0.02+0.03) |
| NILC      | -0.06(0.05+0.09+0.22) | -0.07(0.05+0.09+0.21) | 0.06(0.02+0.02+0.03) | 0.04(0.02+0.02+0.03) |
| SEVEM     | -0.06(0.06+0.09+0.22) | -0.06(0.06+0.09+0.22) | 0.06(0.03+0.02+0.03) | 0.04(0.03+0.02+0.03) |
| SMICA     | -0.06(-0.01+0.03+0.06) | -0.06(-0.01+0.03+0.06) | 0.06(-0.02+0.02-0.04) | 0.04(-0.03+0.02-0.04) |

Notes. The expected values, together with the $1\sigma$ (68% CL) and $2\sigma$ (95% CL) ranges, determined from simulations are given in brackets.

Table 41. $T_0$, as defined in Eqs. (77) and (78), for different thresholds $\nu$ and hemispheres.

<table>
<thead>
<tr>
<th>Method</th>
<th>“North” $T_0$</th>
<th>Cold spots</th>
<th>“South” $T_0$</th>
<th>Cold spots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold $\nu = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commander</td>
<td>-0.02(0.03+0.07+0.16)</td>
<td>-0.03(0.02+0.07+0.16)</td>
<td>-0.05(0.03+0.07+0.18)</td>
<td>-0.06(0.03+0.07+0.18)</td>
</tr>
<tr>
<td>NILC</td>
<td>-0.02(0.02+0.07+0.16)</td>
<td>-0.03(0.02+0.07+0.17)</td>
<td>-0.05(0.02+0.07+0.18)</td>
<td>-0.06(0.02+0.07+0.18)</td>
</tr>
<tr>
<td>SEVEM</td>
<td>-0.02(0.03+0.07+0.17)</td>
<td>-0.03(0.03+0.07+0.18)</td>
<td>-0.05(0.03+0.07+0.18)</td>
<td>-0.06(0.03+0.07+0.18)</td>
</tr>
<tr>
<td>SMICA</td>
<td>-0.02(-0.01+0.04+0.09)</td>
<td>-0.03(-0.01+0.04+0.09)</td>
<td>-0.05(-0.01+0.04+0.08)</td>
<td>-0.06(-0.01+0.04+0.08)</td>
</tr>
</tbody>
</table>

| Threshold $\nu = 1$ |          |            |          |            |
| Commander | -0.04(0.03+0.09+0.23) | -0.05(0.02+0.09+0.23) | -0.08(0.04+0.09+0.25) | -0.08(0.04+0.09+0.24) |
| NILC      | -0.05(0.03+0.10+0.23) | -0.06(0.02+0.10+0.10) | -0.08(0.03+0.06+0.11) | -0.08(0.03+0.06+0.10) |
| SEVEM     | -0.04(0.04+0.10+0.23) | -0.05(0.03+0.10+0.10) | -0.08(0.04+0.09+0.25) | -0.08(0.04+0.09+0.24) |
| SMICA     | -0.04(-0.02+0.05+0.13) | -0.05(-0.02+0.05+0.13) | -0.08(-0.02+0.05+0.12) | -0.09(-0.02+0.05+0.12) |

Notes. The “north” hemisphere is centred on the Galactic coordinate $(l, b) = (212^\circ, -13^\circ)$ and the “south” hemisphere in the opposite direction. The expected values, together with the $1\sigma$ (68% CL) and $2\sigma$ (95% CL) ranges, determined from simulations are given in brackets.

differs for the data and the simulations, is to minimize the impact of the cosmology dependence. A natural choice for the filter is $\langle T_m(\sigma) \rangle$ itself with a proper normalization:

$$W(\sigma) = \frac{\langle T_m(\sigma) \rangle}{\int R \langle T_m(\sigma) \rangle^2 d\sigma}. \quad (78)$$

For the filter given by Eq. (78), the integrated profile deviation $T_m$ describes the relative deviation from the NFEA. If $\Lambda$CDM is the correct model, the deviation is due to cosmic variance and noise. The distribution of $T_m$ is obtained from simulations.

Table 40 presents a comparison of the $T_m$ values derived from the Planck data and the FFP8 simulations. No inconsistencies in excess of the 3$\sigma$ level have been found, although tensions around 2$\sigma$ are seen.

The $m = 0$ projection kernel $J_0[(l + 1/2)\sigma]$ peaks at low $\ell$. Thus $T_m$ is cosmic-variance sensitive and the apparent discrepancy in it could be related to a low-$\ell$ power deficit. An example is shown in Fig. 42 for illustration. To test the robustness of this result, we have tried three additional filters: a top-hat filter $W = 1$, a linear filter $W = \sigma$, and a Gaussian filter $W = \exp(-\sigma^2/\sigma_g^2)$ with $\sigma_g = 1^\circ$. In all cases, the power deficit remains at about the 2$\sigma$ level.

Although the $T_0$ deficit is not significant enough to falsify the $\Lambda$CDM model, further investigation of its properties may still be interesting and help us understand the other anomalies discussed in this paper. We consider two possibilities. Firstly the amplitude of the $T_0$ deficit is of order $5$–$10\%$, which coincides with the level of hemispherical power asymmetry discussed in Sect. 6.1. To test whether the $T_0$ deficit is localized on one hemisphere, we define the “north” direction to be aligned with the power asymmetry direction at $(l, b) = (212^\circ, -13^\circ)$ (Akrami et al. 2014) and compute $T_0$ on the northern and southern hemispheres separately. The results are presented in Table 41. Although the $T_0$ deficit is more significant for the southern hemisphere, it remains consistent with the $\Lambda$CDM prediction. Secondly, it is of interest to determine whether the $T_0$ deficit is related to the Cold Spot discussed in Sect. 5.7. We therefore mask out the Cold Spot using a disc of radius $6^\circ$ and repeat the calculation. The impact of this region on the $T_0$ deficit is insignificant.

A16, page 52 of 62
Tensions at the $2\sigma$ level are also seen for $T_2$. However, due to the additional $\ell^2$ factor in the projection kernel, the oriented ($m = 2$) component $T_2$ is more sensitive to high-$\ell$ power where the cosmic variance is small, and an understanding of the noise properties of the data is more important. The former implies that the related uncertainty in $T_2$ is, in general, smaller than that in $T_0$. However, a mismatched cosmology, perhaps arising from a different primordial power amplitude $A_s$, can then lead to significant tension between the data and the simulations. Indeed, we find that without application of our cosmology calibration (i.e., the subtraction of the NFEA in Eq. (77)) the $T_2$-tension between the data and simulations increases by about $0.5\sigma$, whereas the variation of the $T_0$-tension is $\lesssim 0.2\sigma$. The high-$\ell$ sensitivity of $T_2$ also requires the use of an accurate noise model, and it is possible that the $1-2\sigma$ tension in $T_2$ may be alleviated once improved noise simulations are available.

8.2.2. Oriented polarization stacking

The stacked $Q$ and $U$ images can be decomposed into Fourier modes, $Q+iU = \sum_{m=-\infty}^{\infty} P_m(\sigma)e^{im\phi}$. For unoriented $Q+iU$ stacking on temperature peaks, only $P_2(\sigma)$ has a non-zero NFEA, and it can be linked to the conventional $Q$ stacking via $P_2 = -Q$. Figure 46 shows that the stacked $Q$ image is in excellent agreement with its NFEA and the corresponding stacked image (fourth panel) in Fig. 41, despite the different stacking methodologies adopted (and component-separation method selected for visualization purposes). The length and orientation of the headless vectors represent the polarization amplitude, $P_{\text{stack}} \equiv \sqrt{Q_{\text{stack}}^2 + U_{\text{stack}}^2}$, and direction.

We next consider oriented stacking of the polarization maps, again using $Q_T$, $U_T$ to define the orientation of the patches. The stacked polarization images around temperature peaks have $m = 0, 2, 4$ Fourier components. We can also choose to stack the polarization maps on $P_T$ peaks, where $P_T = \sqrt{Q_T^2 + U_T^2}$. This picks up $m = 0, 4$ Fourier modes with no circularly symmetric ($Q_T, m = 2$) mode. In Fig. 47 we compare the $(Q, U)$ images stacked centred either on $T$ peaks (top panel) or on $P_T$ peaks (bottom panel) with their corresponding NFEAs, and find excellent agreement.

For a quantitative comparison, we only consider stacking on temperature peaks and define the polarization integrated profile deviation

$$P_m(W) = \int_0^R (P_m(\sigma) - \langle P_m(\sigma) \rangle) W(\sigma) \, d\sigma,$$

where by default the filter is

$$W(\sigma) = \frac{\langle P_m(\sigma) \rangle}{\int_0^R (P_m(\sigma))^2 \, d\sigma}.$$

The comparison of $P_m$ ($m = 0, 2, 4$) between the data and the simulations is shown in Table 42, where the results are seen to be in excellent agreement.

Finally, we note that the peak selection does not have to be made from the temperature map. In Fig. 48 we show a few examples of stacking on polarization peaks using the $N_{\text{side}} = 512$ maps. The higher-resolution polarization data are too noisy for peak selection. In the upper panels, we compare stacked images of the $E$-mode map centred around $E$-mode peaks with the corresponding NFEA. We find that the noise impact is relatively minor for 20' FWHM maps and the plots are in qualitatively good agreement. Another possibility, shown in the lower panels, is to stack polarization maps centred on peaks determined from the corresponding polarization amplitude map. In this case the peaks are strongly biased by the quadratic noise contribution and quite visible deviation from the NFEA is observed in the stacked image.

9. Conclusions

In this paper, we have presented a study of the statistical isotropy and Gaussianity of the CMB using the Planck 2015 data, including the full mission for temperature. We do not claim that our results support or refute any particular physical model. Rather,
Table 42. $P_m$, as defined in Eqs. (79) and (80), for different thresholds $\nu$.

<table>
<thead>
<tr>
<th>Method</th>
<th>$P_0$</th>
<th>$P_2$</th>
<th>$P_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commander</td>
<td>0.06(0.02$^{+0.03+0.05}_{-0.03-0.06}$)</td>
<td>-0.01(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
<td>0.04(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
</tr>
<tr>
<td>NILC</td>
<td>0.05(0.02$^{+0.03+0.05}_{-0.03-0.06}$)</td>
<td>-0.02(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
<td>0.03(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
</tr>
<tr>
<td>SEVEM</td>
<td>0.05(0.02$^{+0.03+0.05}_{-0.03-0.06}$)</td>
<td>0.01(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
<td>0.04(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
</tr>
<tr>
<td>SMICA</td>
<td>0.05(0.03$^{+0.03+0.05}_{-0.03-0.06}$)</td>
<td>-0.02(0.00$^{+0.01+0.02}_{-0.01-0.02}$)</td>
<td>0.03(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
</tr>
</tbody>
</table>

Cold spots, threshold $\nu = 0$

<table>
<thead>
<tr>
<th>Method</th>
<th>$P_0$</th>
<th>$P_2$</th>
<th>$P_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commander</td>
<td>0.06(0.02$^{+0.03+0.05}_{-0.03-0.06}$)</td>
<td>-0.01(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
<td>0.03(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
</tr>
<tr>
<td>NILC</td>
<td>0.06(0.02$^{+0.03+0.05}_{-0.03-0.06}$)</td>
<td>-0.01(0.00$^{+0.01+0.02}_{-0.01-0.02}$)</td>
<td>0.04(0.01$^{+0.02+0.04}_{-0.02-0.04}$)</td>
</tr>
<tr>
<td>SEVEM</td>
<td>0.06(0.03$^{+0.03+0.05}_{-0.03-0.06}$)</td>
<td>0.01(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
<td>0.03(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
</tr>
<tr>
<td>SMICA</td>
<td>0.05(0.03$^{+0.03+0.05}_{-0.03-0.06}$)</td>
<td>-0.02(0.00$^{+0.01+0.02}_{-0.01-0.02}$)</td>
<td>0.03(0.02$^{+0.02+0.04}_{-0.02-0.04}$)</td>
</tr>
</tbody>
</table>

Hot spots, threshold $\nu = 1$

<table>
<thead>
<tr>
<th>Method</th>
<th>$P_0$</th>
<th>$P_2$</th>
<th>$P_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commander</td>
<td>0.04(0.02$^{+0.03+0.06}_{-0.04-0.07}$)</td>
<td>-0.02(0.00$^{+0.02+0.05}_{-0.02-0.05}$)</td>
<td>0.05(0.01$^{+0.02+0.06}_{-0.02-0.06}$)</td>
</tr>
<tr>
<td>NILC</td>
<td>0.06(0.02$^{+0.03+0.05}_{-0.04-0.07}$)</td>
<td>-0.02(0.01$^{+0.02+0.05}_{-0.02-0.05}$)</td>
<td>0.05(0.01$^{+0.02+0.06}_{-0.02-0.06}$)</td>
</tr>
<tr>
<td>SEVEM</td>
<td>0.05(0.02$^{+0.04+0.07}_{-0.04-0.07}$)</td>
<td>0.01(0.00$^{+0.01+0.03}_{-0.01-0.03}$)</td>
<td>0.05(0.01$^{+0.02+0.06}_{-0.02-0.06}$)</td>
</tr>
<tr>
<td>SMICA</td>
<td>0.04(0.03$^{+0.03+0.07}_{-0.03-0.07}$)</td>
<td>-0.02(0.00$^{+0.01+0.02}_{-0.01-0.02}$)</td>
<td>0.06(0.04$^{+0.04+0.06}_{-0.04-0.06}$)</td>
</tr>
</tbody>
</table>

Cold spots, threshold $\nu = 1$

<table>
<thead>
<tr>
<th>Method</th>
<th>$P_0$</th>
<th>$P_2$</th>
<th>$P_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commander</td>
<td>0.07(0.02$^{+0.03+0.06}_{-0.04-0.07}$)</td>
<td>0.00(0.01$^{+0.02+0.03}_{-0.02-0.03}$)</td>
<td>0.01(0.01$^{+0.02+0.03}_{-0.02-0.03}$)</td>
</tr>
<tr>
<td>NILC</td>
<td>0.08(0.02$^{+0.03+0.06}_{-0.04-0.07}$)</td>
<td>-0.01(0.01$^{+0.02+0.03}_{-0.02-0.03}$)</td>
<td>0.01(0.01$^{+0.02+0.03}_{-0.02-0.03}$)</td>
</tr>
<tr>
<td>SEVEM</td>
<td>0.09(0.02$^{+0.03+0.06}_{-0.03-0.07}$)</td>
<td>0.00(0.00$^{+0.01+0.02}_{-0.01-0.02}$)</td>
<td>0.02(0.01$^{+0.02+0.03}_{-0.02-0.03}$)</td>
</tr>
<tr>
<td>SMICA</td>
<td>0.06(0.03$^{+0.03+0.06}_{-0.03-0.07}$)</td>
<td>-0.01(0.01$^{+0.02+0.03}_{-0.01-0.02}$)</td>
<td>0.02(0.01$^{+0.02+0.03}_{-0.02-0.03}$)</td>
</tr>
</tbody>
</table>

Notes. The expected values, together with 1$\sigma$ (68% CL) and 2$\sigma$ (95% CL) ranges, determined from simulations are given in brackets.

Fig. 48. Top: $E$-mode maps stacked on the unoriented $E$-mode maxima computed above the null threshold $\nu = 0$. Bottom: $Q$ stacked around oriented polarization amplitude ($P$) maxima. In this case, no threshold is used and the orientation is chosen such that $U = 0$ and $Q \geq 0$ on the central peak. The left panels are the stacked SMICA maps, and the right panels their corresponding NFEAs. See Fig. 46 for the meaning of the headless vectors (black dashed lines). The image units are $\mu K$.

In Sect. 4, we examined aspects of the Gaussianity of the CMB fluctuations. Tests of skewness, kurtosis, multi-normality, $N$-point functions, and Minkowski functionals yielded no indications of significant departures from Gaussianity, while the very nature of such a model-independent approach to leave the detailed interpretation to the reader. However, we do address the important subject of a posteriori correction where possible.

The statistical tests are performed on maps of the CMB anisotropy that result from the application of the four component-separation methods described in Planck Collaboration IX (2016). All of the results presented here are robust with respect to the choice of component-separated CMB map. This is important since it demonstrates the high quality and equivalence of the Planck component-separated data products rendered by different methodologies under varying assumptions.

We find that the CMB is largely consistent with statistical isotropy, although there are a few indications of anomalies with respect to the expectations of $\Lambda$CDM. Some of the tests we have performed are the same as those in PCIS13, in which case the results are consistent. Since many of these anomalies were also observed in the WMAP temperature data, we re-emphasize explicitly the statement we made in 2013 – that the agreement between the two independent experiments effectively rules out the possibility that the origin of these features can be found in residual systematic artefacts present in either data set. We have also performed a number of new tests, in order to try to narrow down the nature of the apparent violations of statistical isotropy. In addition, although the component-separated polarization maps contained in the Planck 2015 release are high-pass filtered, we have performed a stacking analysis that tests some aspects of the polarized sky while mitigating the impacts of noise and systematic effects.
variance of the CMB map was found to be low, in agreement with previous studies (PCIS13). First-order moments of filtered maps also exhibit the low-variance anomaly, as well as a kurtosis excess on certain scales associated with the Cold Spot. A new study of peak statistics finds results consistent with the expectations for a Gaussian random field, although the Cold Spot is again detected.

Section 5 provides an updated study of several previously familiar peculiarities. We study in detail the low variance anomaly, which appears to be associated with the known low-ℓ deficit in the angular power spectrum. We confirm the lack of large-scale angular correlations, relatively featureless northern ecliptic hemisphere 3- and 4-point functions, and indications of violations of point- and mirror-parity symmetry, although we make little or no attempt to correct these for a posteriori effects.

We place tight constraints on a quadrupolar power modulation. The Cold Spot is examined further, and, while we find variance, skewness, and kurtosis angular profiles consistent with the expectations of statistically isotropic simulations, both for oriented and unoriented stacking, with the observed small changes being consistent with the anomalies described in this paper will take us beyond the standard model of cosmology.

In Sect. 6 we perform a series of tests probing the well-known large-scale dipolar power asymmetry. We detect the asymmetry via pixel-to-pixel variance, as well as by measuring power explicitly or indirectly via ℓ to ℓ ± 1 mode coupling. The latter approach lends itself to a posteriori correction, which reduces the significance of the asymmetry substantially when no model for the anomaly is assumed. In addition, we perform two independent but related tests of directionality. One finds suggestions of anomalous clustering of directions out to relatively small scales while the other does not, evidently due to being optimized for slightly different forms of directionality.

Section 7 demonstrates that the significances of several large-angular-scale anomalies are robust to the use of larger sky coverage, with the observed small changes being consistent with expectations from random Gaussian statistics.

Finally, Sect. 8 presents the results of the stacking of temperature and polarization peaks. We find results that largely agree with statistically isotropic simulations, both for oriented and unoriented stacking. The exception is a low unoriented temperature profile, which seems to be yet another reflection of the large-scale power deficit.

With the Planck 2015 release, we are probably near the limit of our ability to probe the CMB anomalies with temperature fluctuations alone. The use of large-angular-scale polarization, expected for the final Planck release, should enable independent tests of these peculiar features. Importantly, this will reduce or eliminate the subjectivity and ambiguity in interpreting their statistical significance. It is a tantalizing possibility that some of the anomalies described in this paper will take us beyond the standard model of cosmology.

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); CSC, MINECO, JA, and RES (Spain); Tekes, AoF, and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); ERC and PRACE (EU). A description of the Planck Collaboration and a list of its members, indicating which technical or scientific activities they have been involved in, can be found at http://www.cosmos.esa.int/web/planck/planck-collaboration.

Some of the results in this paper have been derived using the HEALPix package.

References

Bennett, C. L., Hill, R. S., Hinshaw, g. et al., 2011, ApJS, 192, 17

citation
Appendix A: Generalized Savitzky-Golay polynomials

In the construction of optimal linear filters, one needs to combine information about the (statistically isotropic) CMB signal, anisotropic instrumental noise, masking to be applied for the elimination of foreground contributions, and a model for any non-Gaussian signal for matched filtering. These can be combined in a general framework of normalized convolutions (Knutsson & Westin 1993), where the filtered field is defined as

\[ U = \frac{aB \ast w T}{aB \ast B^T w} \]  

(A.1)

where \( B \) is the (multiscale) filtering beam function, \( T \) is the temperature, \( a \) and \( w \) their respective weights, and \( \ast \) denotes the usual convolution operation

\[ (aB \ast w T)(\xi) = \sum_x a(x) B(x) \cdot w(\xi - x) T(\xi - x). \]  

(A.2)

In the absence of a specific model for the non-Gaussian signal, the beam functions can be taken to be orthogonal polynomials on a disc, weighted by some smoothing function, while the weights applied to the temperature maps are determined by the CMB and noise covariance.

In a simple approach, the information about the CMB signal can be utilized by pre-whitening the map by convolving it with an isotropic beam function \( w_f = C_T^{-1/2} \) derived from the isotropic best-fit CMB power spectrum combined with a diagonal approximation to the instrumental noise covariance. After the component-separated CMB maps are pre-whitened, and the corresponding mask is applied to the resulting map, the multi-scale filtering kernel \( b_f \) is applied at various scales.

In this paper, the maps are pre-whitened with the 2013 best-fit cosmological parameter CMB spectrum (Planck Collaboration XV 2014), co-added to an isotropic noise power spectrum derived from the half-mission, half-diagonal noise maps appropriate for each component-separation method. No adjustment is made either for the recalibration of the 2015 data relative to the nominal results that the cosmological spectrum is derived from, or for the mismatch in noise level between the half-mission, half-diagonal and full-mission maps. This implies that the filtering is sub-optimal, but the data and simulations are treated consistently so there should be no significant impact on the results. The resulting pre-whitening beam function \( w_f \) for the SMICA temperature map is shown in Fig. A.1.

The peak detector wavelets are taken to be Savitzky-Golay polynomials (Savitzky & Golay 1964), generalized to be defined on a disc with a polynomial smoothing weight function applied, as shown in Fig. A.1. A generalized spherical Savitzky-Golay kernel of order \( n \) and smoothing weight \( k \) (referred to as SSGnk in the text) is defined by a polynomial function of a radial variable \( x = \sin(\theta/2)/\sin(\theta_{\text{max}}/2) \)

\[ F_{n,k}(x) = \left( \sum_{i=0}^{n/2} a_i x^i \right) \left( 1 - x^2 \right)^{\frac{k}{2}}, \]  

(A.3)

which is normalized to have unit mean on a disc and is orthogonal to all non-constant polynomials up to order \( n \),

\[ \int_0^1 x F_{n,k}(x) \, dx = 1, \quad \int_0^1 x^{n+1} F_{n,k}(x) \, dx = 0. \]  

(A.4)

These are essentially high-order low-pass filters in harmonic space, but have compact support on the sphere. A few representative Savitzky-Golay polynomials are compared to a Gaussian kernel in Fig. A.1. Combined with pre-whitening, the total effect of the filters applied is described by the composite beam functions shown in Fig. A.1.

One should note a slight \( \ell \)-space bandwidth mismatch between differently shaped kernels with the same FWHM value in real space, which is clear from the lower left panel of Fig. A.1. While not a problem in general, some care should be exercised when directly comparing results for different shape kernels. In particular, the \( \ell \) value at which the filter kernel coefficient reaches \( b_{\ell} = b_{\text{max}}/2 \) differs by a factor of 1.58 between the GAUSS and SSG84 kernels of the same FWHM.

Appendix B: Doppler boosting

The main effect of our relative motion with respect to the CMB rest frame is a dominant contribution to the CMB dipole (\( C_l^\ell \)); this is boosting of the monopole and has been detected previously (Kogut et al. 2003; Fixsen et al. 1996; Hinshaw et al. 2009). A subtler consequence of our motion is the boosting of all other multipoles. In fact, there are really two effects at work. The first is a modulation effect which increases power by approximately 0.25% in the direction of our motion and decreases it by the same amount in the opposite direction. This can equivalently be thought of as coupling between the multipoles \( \ell \) and \( \ell \pm 1 \). The second is an aberration effect which shifts the apparent direction in which CMB photons arrive at our detectors toward the velocity direction.

Planck Collaboration XXVII (2014) reported a detection of this Doppler boosting, and an associated measurement of its velocity signature of \( 384 \pm 78 \) (statistical) \( \pm 115 \) (systematic) km s\(^{-1}\) in the known dipole direction, \((l, b) = (264^\circ, 48^\circ)\). Here, we demonstrate that the Planck 2015 data release remains in agreement with this result, by considering the angular scales 500 \( \leq \ell \leq 2000 \). However, since the simulations employed in the analysis partially contain the effects of Doppler boosting (as noted in Sect. 3 the aberration contribution was erroneously omitted), we report a consistency check rather than a detection.

It is useful to perform a harmonic transform on the peculiar velocity vector,

\[ \beta_{\ell M} = \int d\hat{n} Y_{\ell M}^* (\hat{n}) \beta \cdot \hat{n}, \]  

(B.1)

where only the \( L = 1 \) modes are non-zero. Following the convention in Planck Collaboration XXVII (2014), we rotate to an orthonormal basis, labelled \( \beta_n \) (along the expected velocity direction), \( \beta_n^L \) (parallel to the Galactic plane), and \( \beta_n^L \) (the remaining vector).

The peculiar velocity is detected using estimators that pick out the off-diagonal components of the CMB covariance matrix

\[ \langle T_{\ell m_1} T_{\ell m_2} \rangle_{\text{CMB}} = \sum_{LM} (-1)^M \left( \frac{\ell + 1}{m_1 m_2 L} \right) W_{\ell m_1 m_2} \beta_{\ell M}^L. \]  

(B.2)

The weight function \( W_{\ell m} \) is a sum of the modulation (\( b_{\ell} W^\ell \)) and aberration (\( W^\phi \)) effects. We quote results based on orthogonalized weight matrices,

\[ W^\phi = W^\ell - W^\ell R^\ell / R^\ell \]  

(B.3)

\[ W^\tau = W^\ell - W^\ell R^\ell / R^\phi \]  

(B.4)

A16, page 59 of 62
Table B.1. Significance measures for the $\beta$ estimates for the 143 × 217 data set.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>$\chi^2$</th>
<th>PTE [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>3.28</td>
<td>7.01</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.21</td>
<td>64.39</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.08</td>
<td>77.53</td>
</tr>
<tr>
<td>$\beta$</td>
<td>3.38</td>
<td>33.70</td>
</tr>
</tbody>
</table>

Notes. $\chi^2$ is formed from the three modes of $\beta$ using the covariance matrix measured from Doppler boosted simulations.

Due to the clear connection between the velocity estimators and those used for the lensing analysis, we adopt the same data (143 GHz and 217 GHz sky maps, with dust foregrounds removed using the 857 GHz data as a template) and mask as used in Planck Collaboration XV (2016). The results summarized in Table B.1 show a slight excess of signal in the dipole direction of the data compared to simulations. This is due to the simulations used containing the modulation, but not aberration, part of the Doppler boost signal.

Appendix C: Generalized modulation estimator

Consider a parameter $X$ that the (primary) CMB power spectrum is dependent on. Let $X$ have a dipolar dependence of the form $X(\hat{n}) = X_0 + \Delta X \hat{n} \cdot \hat{n}$ (this could correspond to a gradient in $X$ across our observable volume), where $X_0$ is the average value, $\hat{n}$ is the direction to the last scattering surface, and $\hat{m}$ is the gradient direction. To linear order in $\Delta X/X$, the measured spherical harmonics coefficients are given by

$$a_{\ell m} = a_{\ell m}^{iso} + \sum_M \Delta X_M \sum'_{\ell' m'} \frac{d_{\ell' m'}^{iso}}{d\ell m} \epsilon_{\ell'm'}^M,$$  \hspace{1cm} (C.1)

where the $a_{\ell m}^{iso}$ are the unmodulated statistically isotropic modes. The $\epsilon_{\ell'm'}^M$ are coupling coefficients given by

$$\epsilon_{\ell'm'}^{0} = \delta_{m'm} (\delta_{\ell'-1}A_{\ell-1,m} + \delta_{\ell'+1}A_{\ell,m}),$$  \hspace{1cm} (C.2)

$$\epsilon_{\ell'm'}^{1} = \delta_{m'm-1} (\delta_{\ell'-2}B_{\ell-2,m-1} - \delta_{\ell'+1}B_{\ell,m}),$$  \hspace{1cm} (C.3)

where

$$A_{\ell m} = \sqrt{\frac{(\ell + 1)^2 - m^2}{(2\ell+1)(2\ell+3)}},$$  \hspace{1cm} (C.4)

$$B_{\ell m} = \sqrt{\frac{(\ell + m + 1)(\ell + m + 2)}{2(2\ell+1)(2\ell+3)}}.$$  \hspace{1cm} (C.5)

From Eq. (C.1) we can find the covariance matrix to first order in the components $\Delta X_M$:

$$C_{\ell'm'} = C_{\ell'\ell} \delta_{m'm} + \frac{\delta C_{\ell'}{\ell}}{2} \sum_M \Delta X_M \epsilon_{\ell'm'}^M.$$  \hspace{1cm} (C.6)
where $\delta C_{\ell+1} = dC_{\ell}/dX + dC_{\ell+1}/dX$. To determine the best-fit parameters, we proceed by maximizing the CMB likelihood function
\begin{equation}
\mathcal{L} = \frac{1}{\sqrt{2\pi|C|}} \exp(-d^2C^{-1}d/2),
\end{equation}
where $d$ is the CMB temperature data. Equation (C.7) is maximized for the $\Delta X_M$ that satisfy
\begin{equation}
d^t C^{-1} \frac{dC}{d\Delta X_M} C^{-1} d = \text{Tr} \left( C^{-1} \frac{dC}{d\Delta X_M} C^{-1} \right).
\end{equation}

From Eq. (C.6) it is clear that the CMB covariance can be decomposed into an isotropic part ($C_\ell$) and a small anisotropic part proportional to $\Delta X_M$. By inverting Eq. (C.6) and using the orthogonality of the $\xi_{0 \ell}$, we can determine the best-fit parameters
\begin{equation}
\Delta X_0 = \sum_{\ell m} \frac{\delta c_{\ell \ell+1} \xi_{0 \ell} \xi_{0 \ell+1}^* m}{\sqrt{C_\ell C_{\ell+1}}},
\end{equation}
\begin{equation}
\Delta X_{\ell m} = \sum_{\ell' m'} \frac{\delta c_{\ell \ell+1} \xi_{0 \ell} \xi_{0 \ell+1}^* m}{\sqrt{C_{\ell \ell'} C_{\ell+1 \ell'+1}}},
\end{equation}
and $\Delta X_{-\ell m} = -\Delta X_{\ell m}$, to first order in the anisotropy. These estimators are the full-sky, no-noise versions of Eqs. (44) and (45).

Errors can easily be found by expanding the log-likelihood about the best-fit parameters. The Fisher matrix is defined as
\begin{equation}
F_{MM} \equiv \frac{1}{2} \text{Tr} \left( \frac{dC}{d\Delta X_M} C^{-1} \frac{dC}{d\Delta X_M} C^{-1} \right).
\end{equation}

Upon switching bases, we find
\begin{equation}
F_{0,0} = \frac{1}{2} \sum_{\ell m} \frac{\delta c_{\ell \ell+1}^2 A_{0 \ell} A_{0 \ell+1}^*}{C_\ell C_{\ell+1}},
\end{equation}
\begin{equation}
F_{R(\Delta X_{\ell m}),R(\Delta X_{\ell m'})} = \frac{1}{2} \sum_{\ell m} \frac{\delta c_{\ell \ell+1}^2 B_{\ell m} B_{\ell m'}^*}{C_{\ell \ell'} C_{\ell+1 \ell'+1}}.
\end{equation}

We can then assign the standard errors, $\sigma = \sqrt{F}^{-1}$.

**Appendix D: Weighted-variance modified shape function estimator**

The BipoSH representation characterizes the off-diagonal elements in the covariance matrix and is a generalization of the angular power spectrum, $C_\ell$.
\begin{equation}
A_{\ell m}^{LM} = \sum_{m''} (a_{\ell m''} c_{m''})^{-1} C_{\ell m''},
\end{equation}

In general, it is not possible to analyse the full sky even for component-separated maps, due to the presence of residual contributions from diffuse Galactic emission and point sources. However, the application of a mask leads to coupling between the spherical harmonic modes. Hence, the correlation function is no longer described only by $C(\theta)$ or the power spectrum $C_\ell$, and other quantities are required to completely quantify the statistical field.

We obtain an analytic expression for the observed BipoSH coefficients after the application of a mask in terms of the corresponding coefficients of the unmasked sky, and of the mask itself,
\begin{equation}
\tilde{A}_{\ell m}^{LM} = \sum_{\ell' m'} \frac{\Pi_{\ell' m'}}{\sqrt{\ell' m'}} \sum_{\ell'' m''} \frac{\Pi_{\ell'' m''}}{\sqrt{\ell'' m''}} C_{\ell'' m''}^{L0} C_{\ell'' m''}^{L0} \times \sum_{\ell' m' j k} \left\{ \begin{array}{cc}
L_{\ell' m'} & \ell' m' \\
j & k
\end{array} \right\} \Pi_{j k} W_{\ell' m'}^{LM} C_{\ell m j k}^{LM},
\end{equation}
where $\Pi_{\ell'} = \sqrt{2\ell' + 1}$, $A_{\ell m}^{LM}$ are the BipoSH coefficients of the masked sky map, $A_{\ell m}^{LM}$ correspond to the BipoSH coefficients of the unmasked sky, $W_{\ell m}^{LM}$ are the BipoSH coefficients of the mask itself, $C_{\ell m j k}^{LM}$ are the Clebsch-Gordon coefficients, and the term $[\{}$ in Eq. (D.2) is the $j$–symbol. This quantifies the coupling between the BipoSH coefficients of the CMB sky map and those of the mask itself.

The underlying CMB sky may have deviations from statistical isotropy, as discussed in Sect. 6.4, due either to a dipole modulation ($L = 1$) of unknown origin, or to Doppler boosting ($L = 1$) of the temperature field. The BipoSH coefficients of such statistical isotropy-violating fields can be given by
\begin{equation}
A_{\ell m}^{LM} = \tilde{A}_{\ell m}^{LM} + \phi_{LM} G_{\ell m}^{LM},
\end{equation}

Here $\tilde{A}_{\ell m}^{LM}$ corresponds to the BipoSH coefficients of the unknown but statistically isotropic CMB field. This couples with BipoSH coefficients of the mask to introduce a mean field linear bias $\langle A_{\ell m}^{LM} \rangle_{\text{mask}}$, which is estimated from simulations and subtracted from the BipoSH coefficients obtained from the masked sky. The $\phi_{LM}$ are the spherical harmonic coefficients of the field that breaks statistical isotropy, and $G_{\ell m}^{LM}$ is the shape function. Shape functions for dipole modulation and Doppler boosting are given in Eqs. (54) and (56), respectively.

Due to symmetries of the mask, which is largely defined by foreground residuals towards the Galactic plane, the dominant BipoSH modes of the mask correspond to $J = [0, 2], K = 0$. Hence, for all practical purposes, signal is retained in the $L = 1$ mode itself, although masking modifies the shape function, now defined as the modified shape function in the rest of the text. A weighted variance modified shape function is defined as
\begin{equation}
\tilde{\phi}_{LM} = \sum_{\ell' m'} \frac{\tilde{A}_{\ell' m'}^{LM} K_{\ell' m'}^{LM}}{K_{\ell m}^{LM}},
\end{equation}

where $\tilde{A}_{\ell m}^{LM}$ is $A_{\ell m}^{LM} - \langle A_{\ell m}^{LM} \rangle_{\text{mask}}$ and the weights are chosen such that $\sum_{\ell m} w_{\ell m}^{LM} = 1$.

Here $K_{\ell m}^{LM}$ is the MSF, which can be evaluated as
\begin{equation}
K_{\ell m}^{LM} = \sum_{\ell' m'} \frac{\Pi_{\ell' m'} G_{\ell' m'}^{LM}}{\sqrt{\ell' m'}} \sum_{\ell'' m''} \frac{\Pi_{\ell'' m''}}{\sqrt{\ell'' m''}} C_{\ell'' m''}^{L0} C_{\ell'' m''}^{L0} \times \sum_{\ell' m' j k} \left\{ \begin{array}{cc}
L_{\ell' m'} & \ell' m' \\
j & k
\end{array} \right\} \Pi_{j k} W_{\ell' m'}^{LM} C_{\ell m j k}^{LM}.
\end{equation}
The weights are then given by

\[ w^{LM}_{\ell_1\ell_2} = \frac{1}{\sum_M \left( \frac{\sigma_{\lambda_2 M \ell_2}}{K^{LM}_{\ell_1\ell_2}} \right)^2} \times \left[ \sum_{\ell_1'\ell_2'} \frac{1}{\sum_M \left( \frac{\sigma_{\lambda_1 M \ell_1'}}{K^{LM}_{\ell_1\ell_2}} \right)^2} \right]^{-1} \]  

(D.6)