
(Affiliations can be found after the references)

Received 26 March 2013 / Accepted 24 March 2014

ABSTRACT

We present the current estimate of instrumental and systematic effect uncertainties for the Planck-Low Frequency Instrument relevant to the first release of the Planck cosmological results. We give an overview of the main effects and of the tools and methods used to assess residuals in maps and power spectra. We also present an overall budget of known systematic effects, which are dominated by sidelobe straylight pick-up and imperfect calibration. However, even these two effects are at least two orders of magnitude weaker than the cosmic microwave background fluctuations as measured in terms of the angular temperature power spectrum. A residual signal above the noise level is present in the multipole range ℓ < 20, most notably at 30 GHz, and is probably caused by residual Galactic straylight contamination. Current analysis aims to further reduce the level of spurious signals in the data and to improve the systematic effects modelling, in particular with respect to straylight and calibration uncertainties.

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

1. Introduction

This paper, one of a set of papers documenting the 2013 release of data from the Planck1 mission (Planck Collaboration I 2014), describes the Planck-LFI instrument systematic effects and their related uncertainties in cosmic microwave background (CMB) temperature maps and power spectra. Systematic effects in Planck-HFI data are discussed in Planck Collaboration VI (2014) and Planck Collaboration X (2014).

The LFI implements a pseudo-correlation differential design similar to WMAP (Jarosik et al. 2003a,b) to suppress 1/f amplifier gain and noise fluctuations (Seiffert et al. 2002; Mennella et al. 2003; Bersanelli et al. 2010) as well as correlated effects from thermal and electrical variations affecting both the sky signal and reference loads. The reference signal is provided by stable 4.5 K blackbodies thermally and mechanically connected to

* Corresponding author: A. Mennella
  e-mail: anieilo.mennella@fisica.unimi.it

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

Article published by EDP Sciences

A3, page 1 of 23
the external structure of the High Frequency Instrument (HFI) 4 K box (Valenziano et al. 2009; Lamarre et al. 2010). The offset between the sky and reference signals, of the order of 1–2 K, is balanced in software during data processing on the ground (Mennella et al. 2003; Zacchei et al. 2011). The differential time streams are characterised by 1/f noise knee frequencies in the range 10–100 mHz (Mennella et al. 2010, 2011), leaving residual correlated low-frequency fluctuations in gain and signal that are removed during calibration and map-making.

The LFI is also an excellent polarimeter, with very low systematic effects. Depolarisation by the optics and by imperfections in the orthomode transducers, separating the orthogonal linear polarisations, has been accurately measured on the ground and is almost negligible (Leahy et al. 2010).

Asymmetrical bandpass response in the two radiometers is the main source of $I \rightarrow (Q, U)$ leakage in the foreground-dominated sky regions, especially at low frequencies. Although accurate knowledge of the bandpass response allows us, in principle, to correct for this effect during data analysis, the ground bandpass measurements were not accurate enough to maintain this residual below 1% (Zonca et al. 2009). For this reason the spurious polarisation from bandpass mismatch was estimated and removed using flight data, as described in Planck Collaboration II (2014).

Optical effects arise mainly from Galactic and CMB dipole pick-up caused by primary and secondary mirror spillovers (Tauber et al. 2010; Sandri et al. 2010). This is relevant especially for polarisation measurements at 30 GHz, where Galactic emissions are stronger.

In this paper we provide a preliminary overview of the instrument systematic effects and the uncertainties they cause on CMB temperature maps and power spectra (see Sect. 2). In Sect. 3 we outline and discuss the known instrumental effects, separating them into two broad categories: (i) effects that do not depend on the sky signal and impact the radiometric measurements as an additive spurious fluctuation or a gain variation; and (ii) effects that do depend on the sky signal, i.e., on its amplitude and/or on the scanned sky region. Some of these effects are removed in the data processing pipeline by means of algorithms described in Planck Collaboration II (2014). The assessment of the residual uncertainty, discussed in Sect. 4, was performed according to two different strategies. Null tests were the primary tool to check for systematic effect residuals exceeding the white noise level. We also assessed their impact on radiometric time streams, even if below the white noise limit, by exploiting in-flight housekeeping and scientific data.

Some of the effects discussed in this paper are also relevant for calibration, and are discussed in detail in Planck Collaboration V (2014). In this case we provide here only a brief discussion of the most relevant points and results, deferring to the dedicated paper any further details.

Throughout this paper we follow the naming convention described in Appendix A of Mennella et al. (2010) and also available on-line on the Explanatory Supplement (Planck Collaboration 2013).

2. Summary of uncertainties due to systematic effects

In this section we provide a top-level overview of the uncertainties due to systematic effects in the Planck-LFI CMB temperature maps and power spectra. Table 1 provides a list of these effects, with short descriptions of their cause, strategies for their removal and references to sections and/or papers where more information can be found. This section also provides a summary of the main results of our analysis, as detailed in Sect. 4 and corresponding subsections.

The impact of 1/f noise has been assessed using “half-ring” noise maps (see Sect. 4.1.2) normalised to the white noise estimate at each pixel obtained from the white noise covariance matrix, so that a perfectly white noise map would be Gaussian and isotropic with unit variance. Deviations from unity trace the contribution of residual 1/f noise in the final maps, which ranges from 0.06% at 70 GHz to 2% at 30 GHz, as detailed in Sect. 12.2 of Planck Collaboration II (2014).

Pixel uncertainties due to other systematic effects have been calculated on simulated maps degraded to $N_{\text{side}} = 128$ at 30 and 44 GHz and $N_{\text{side}} = 256$ at 70 GHz in order to approximate the optical beam size. In Table 2 we list the rms and the difference between the 99% and the 1% quantiles in the pixel value distributions. For simplicity we refer to this difference as the peak-to-peak (p-p) difference, although it neglects outliers but effectively approximates the peak-to-peak variation of the effect on the map.

Angular power spectra have been obtained from full resolution ($N_{\text{side}} = 1024$) systematic effect maps at each frequency using the HEALPix Anafast routine (Górski et al. 2005). We have then evaluated the propagation of the various effects in the final CMB map by assuming a simple internal linear combination component separation, as explained in Sect. 4.5. In Fig. 1 we show how the power spectra of the various effects compare with the Planck temperature spectrum, with the noise level coming from the half-ring difference maps (see Sect. 4.1.2) and with the residual map obtained from a difference map between survey 1 and survey 22 (see Sect. 4.1.3). The large plot in the top panel shows the power spectra obtained from frequency-independent maps resulting from the weighted-average of frequency maps using the weights specified in Sect. 4.5. Spectra in the three small plots in the lower panel, instead, show contributions of systematic effects from individual frequency maps.

Our analysis is based on a combined assessment of known and unknown systematic effects via simulations and null-maps. It is worth underlining that some effects might not be detected in difference maps, although none of these effects are likely to significantly affect the results of our analysis, as discussed in detail in Sect. 4.1.1. Our assessment shows that the global impact of systematic effect uncertainties is at least two orders of magnitude less than the CMB power spectrum, and demonstrates the robustness of Planck-LFI temperature anisotropy measurements. Comparison between the total simulated systematic effects and the residual signal obtained by differencing survey 1 and survey 2 maps highlights an excess signal in the multipole range $\ell \lesssim 20$ that is not completely accounted for in our simulations. This excess comes mainly from the 30 GHz channel and is likely to be caused by Galactic emissions picked up by beam sidelobes. Also the 44 GHz and 70 GHz channels show residuals at low multipoles, although smaller than at 30 GHz. Understanding this excess and further reducing the level of residual systematic uncertainties is the primary goal of our current analysis to obtain polarisation measurements with a level of purity comparable to what has been achieved for temperature anisotropies.

2 Time periods relative to individual surveys are defined in Table 11 of Planck Collaboration II (2014).
Table 1. List of known instrumental systematic effects in Planck-LFI.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Source</th>
<th>Control/removal</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>White noise correlation</td>
<td>Phase switch imbalance</td>
<td>Diode weighting</td>
<td>3.1.1</td>
</tr>
<tr>
<td>1/f noise</td>
<td>RF amplifiers</td>
<td>Pseudo-correlation and destriping</td>
<td>3.1.1</td>
</tr>
<tr>
<td>Bias fluctuations</td>
<td>RF amplifiers, back-end electronics</td>
<td>Pseudo-correlation and destriping</td>
<td>3.1.3</td>
</tr>
<tr>
<td>Thermal fluctuations</td>
<td>4 K, 20 K and 300 K thermal stages</td>
<td>Calibration, destriping</td>
<td>3.1.2, 4.2.1</td>
</tr>
<tr>
<td>1 Hz spikes</td>
<td>Back-end electronics</td>
<td>Template fitting and removal</td>
<td>3.1.4, 4.2.3</td>
</tr>
</tbody>
</table>

Effects independent of sky signal

- Main beam ellipticity: Accounted for in window function Planck Collaboration IV (2014)
- Intermediate sidelobes: Masking of Galaxy and point sources Not treated in this release
- Far sidelobes pick-up: Model sidelobes removed from timelines (not implemented in this release) 3.2.1, 4.3.1
- Bandpass asymmetries: Spurious polarisation removal Planck Collaboration II (2014)
- Analogue-to-digital converter non linearity: Template fitting and removal 3.2.2, 4.3.2
- Imperfect photometric calibration: Calibration using the 4 K reference Planck Collaboration II (2014)
- Pointing: Negligible impact on temperature anisotropy measurements 3.3, 4.4

Effects dependent on the sky signal

Table 2. Summary of systematic effects uncertainties on maps in $\mu K_{\text{CMB}}$.

<table>
<thead>
<tr>
<th></th>
<th>30 GHz</th>
<th>44 GHz</th>
<th>70 GHz</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p-p rms</td>
<td>p-p rms</td>
<td>p-p rms</td>
</tr>
<tr>
<td>Bias fluctuations</td>
<td>0.08</td>
<td>0.01</td>
<td>0.10</td>
</tr>
<tr>
<td>Thermal fluctuations</td>
<td>0.61</td>
<td>0.11</td>
<td>0.40</td>
</tr>
<tr>
<td>1 Hz spikes</td>
<td>0.87</td>
<td>0.17</td>
<td>0.03</td>
</tr>
<tr>
<td>Sidelobes pick-up</td>
<td>18.95</td>
<td>4.53</td>
<td>1.92</td>
</tr>
<tr>
<td>ADC non-linearity</td>
<td>3.87</td>
<td>1.01</td>
<td>0.19</td>
</tr>
<tr>
<td>Calibration</td>
<td>4.33</td>
<td>1.16</td>
<td>4.74</td>
</tr>
<tr>
<td>Total($^a$)</td>
<td>21.02</td>
<td>4.83</td>
<td>5.61</td>
</tr>
<tr>
<td></td>
<td>1.13</td>
<td>7.87</td>
<td>2.00</td>
</tr>
</tbody>
</table>

Notes. ($^a$) Calculated on a pixel size approximately equal to the average beam FWHM. ($^b$) The total has been computed on maps resulting from the sum of individual systematic effect maps.

3. Overview of LFI systematic effects

Known systematic effects in the Planck-LFI data can be divided into two broad categories: effects independent of the sky signal, which can be considered as additive or multiplicative spurious contributions to the measured timelines, and effects which are dependent on the sky and that cannot be considered independently of the observational strategy.

These effects can generate correlations in the data, and should be removed from timelines before noise is assessed and maps are generated. For this release, based on temperature data only, we have removed from timelines three of these effects: diode-diode correlations (Sect. 3.1.1), ADC non linearity (Sects. 3.2.2 and 4.3.2) and 1 Hz frequency spikes (Sects. 3.1.4 and 4.2.3). The remaining effects have been treated as noise, and their effect assessed via the noise covariance matrices (Planck Collaboration II 2014) and half-ring difference maps (see Planck Collaboration II 2014, and Sect 4.1). The future release will include a deeper assessment and removal of instrumental effects to match the required accuracy for polarisation.

3.1. Effects independent of sky signal

3.1.1. Noise correlations and 1/f noise

Each Planck-LFI receiver is a pseudo-correlation system viewing a scalar feed directed through the telescope at the sky, together with a reference cold load thermally stable near 4 K. Non-white noise from the cold front-end amplifiers is reduced via the correlation, while fluctuations in the later stages of the receiver are minimised by modulating a phase switch in the correlation section at 8192 Hz. The LFI receiver design, construction, ground performance and initial flight performance have been extensively documented (Bersanelli et al. 2010; Mennella et al. 2010, 2011).

The noise properties of the receivers play an important role in downstream data analysis. In particular, we need good
Fig. 1. Angular power spectra of the various systematic effects compared to the Planck temperature anisotropy spectrum. The black dashed curve, representing the total contribution, has been derived from a map where all the systematic effects have been summed. Top panel: power spectra obtained from frequency independent maps resulting from the weighted-average of individual systematic effect frequency maps. Bottom panel: contributions of systematic effects from individual frequency maps. The CMB curve corresponds to the Planck best-fit model presented in Planck Collaboration XV (2014). In the bottom panels the CMB spectrum has been filtered by the beam window function for each frequency.
estimates of the white noise level, long term stability (1/$f$-type noise) and knowledge of any correlated noise components.

The receiver architecture is symmetric, with two complementary detector diodes as output for each receiver channel. As described in Seiffert et al. (2002) and Mennella et al. (2011) imperfect matching of components limits isolation between the complementary diodes of a receiver between $-10$ and $-15$ dB. This imperfect isolation leads to a small anti-correlated component in the white noise that is cancelled by a weighted average of the time ordered data from the two diodes of each receiver as the first step of analysis. This avoids the complication of tracking the anti-correlated white noise throughout the analysis.

We treat the combined diode data as the raw data, and calibration, noise estimation, map-making etc. are performed on these combined data. The weights were determined from some initial estimates of the calibrated noise for each detector, and are kept fixed for the entire mission.

Noise parameters were reported in Mennella et al. (2011). A longer data set, some thermal instabilities in the instrument (particularly during survey 3), and refinements of the data analysis (map making and noise covariance matrix) all require a more detailed look at the long term evolution of the noise characteristics of the receivers.

The noise power spectral density $P(f)$ of the receivers is generally well described by

$$P(f) = \sigma^2 \left[ 1 + \left( \frac{f}{f_k} \right)^\alpha \right], \tag{1}$$

where $\sigma$ characterises the white noise component, the knee frequency, $f_k$, denotes the frequency where white noise and 1/$f$ contribute equally in power to the total noise, and $\alpha$ characterises the slope of the power spectrum for frequencies $f < f_k$. In the following, low frequency power-law noise will referred to as 1/$f$ noise, regardless of its slope, $\alpha$.

We estimate the signal-subtracted noise power spectrum of each receiver on five-day time periods. Except for specific, mostly well understood events, shorter time scale noise estimation does not produce any evident trends. For nearly all the radiometers our noise model is a very good approximation of the power spectrum. We plot a representative comparison in Fig. 2. A few channels show features that have not been captured well by this simple model; the worst is displayed in Fig. 3.

Over the course of the nominal mission, the noise is well fit by the model, with the exception of the early parts of sky survey 3. During this time, thermal instabilities brought on by the switch-over from the nominal to the redundant sorption cooler cause poor fits and some changes in the parameters. In Figs. 4 through 6 we show the behaviour of the three noise parameters in Eq. (1) estimated on 5-day sections of data over the nominal time period. White noise and knee frequency are stable, while the slope starts increasing in absolute value after day 300, as a result of larger temperature fluctuations in the 20 K focal plane (FP). The jump in slope after day 500 is correlated with the sorption cooler switch-over (see Sect. 3.1.2 for further details).
Knee frequency [mHz]  
Days after launch  

Fig. 5. Fitted knee frequencies over the nominal survey for representative radiometers at 30, 44 and 70 GHz. Values are estimated on 5-day sections of data.

Fitted power-law slopes for low frequency noise. Here we note significant instability after day 300. This is due to substantially greater thermal instability of the 20 K stage just before and after switch over between the two sorption coolers, which occurred at day 460.

Fig. 6.

3.1.2. Thermal effects

The LFI is susceptible to temperature fluctuations in the 300 K back-end modules, in the 4 K reference loads, and in the 20 K FP. Figure 7 provides an overview of the main temperatures during the period between day 91 (the start of nominal operations) and 563 after launch.

The two topmost plots show the reference load temperatures at the level of the 70 GHz and 30–44 GHz channels, respectively. The temperature of the 70 GHz reference loads is actively controlled by a proportional-integral-derivative (PID) system and is very stable (δT_{rms} ∼ 0.13 mK, see the zoomed plot in the inset). Reference loads of the 30 and 44 GHz channels, instead, do not benefit from active thermal control. Their temperature is consequently more unstable and susceptible to major system-level events like, for example, the switch over to the redundant sorption cooler.

The third plot from the top of Fig. 7 shows the 20 K LFI FP temperature measured by a sensor placed on the feed horn flange of the LFI28 receiver. The temperature during the first sky survey was very stable, with a δT_{rms} ≤ 1 mK. Towards the end of the first year of operations the sorption cooler performance started to degrade and its stability was maintained with a series of controlled temperature changes. The switch over to the redundant cooler was performed on August 11, 2010, leaving a clear signature on all the main LFI temperatures. After this operation the level of temperature fluctuations in the FP increased unexpectedly, and this was later understood to be the effect of liquid hydrogen that was still present in the cold end of the nominal cooler, because the degraded compressor system was not able to absorb all the hydrogen that was present in the cooler line. Although this effect was later mitigated by a series of dedicated operations, most of the third sky survey suffered from a higher-than-nominal level of temperature variation.

The last plot shows the temperature of the 300 K electronics box, measured by one of its temperature sensors. During the first sky survey the back-end temperature suffered from a daily fluctuation caused by the satellite transponder that was switched on daily during contact with the ground station. After day 258 the system was left continuously on, and the modulation disappeared. This operation caused an increase in the absolute temperature level. The second temperature change occurred in correspondence to the sorption cooler switch-over operation. The plot...
also shows a yearly temperature modulation due to the satellite rotation around the Sun and a temperature spike at day 191 after launch. This was caused by an operational anomaly that led the satellite to fail to re-point for an entire day with a corresponding temperature increase of the warm units.

More details about the thermal stability performance of Planck can be found in Planck Collaboration II (2011), while the susceptibility of the LFI to temperature variations is discussed in Terenzi et al. (2009b).

### 3.1.3. Bias fluctuations

The signal detected by the radiometers can vary because of fluctuations in the front-end and back-end amplifier bias voltages. In the LFI these fluctuations occurred according to two time scales:

- slow electric drifts, due to thermal changes in the power supply, in the RF amplifiers, and in the detector diodes;
- fast and sudden electric instabilities, arising in the warm electronics or from electromagnetic interference effects, and affecting both the cold amplifiers and the warm detector diodes.

The effect of slow drifts is suppressed by the pseudo-correlation architecture of the differential radiometers. Fast electric changes produce quasi-random fluctuations and abrupt steep drops or jumps in the signal. If jumps are caused by instabilities in the front-end bias voltage, then the effect involves the output voltage of both diodes in the radiometer. When the jumps occur in the back-end detector diodes (so-called “popcorn noise”), they impact only the output voltage of the corresponding diode and affect sky and reference load samples. In both cases the differentiated signal is largely immune from these effects.

### 3.1.4. 1 Hz spikes

This effect is caused by pick-up from the housekeeping electronics clock that occurs after the detector diodes and before the analogue-to-digital converter (ADC; Meinhold et al. 2009; Mennella et al. 2010, 2011). This spurious signal is detected in the radiometer time-domain outputs as a 1 s rectangular wave, with a rising edge near 0.5 s and a falling edge near 0.75 s in on-board time. In the frequency domain it appears at multiples of 1 Hz.

Frequency spikes are present at some level in the output from all detectors, but affect the 44 GHz data most strongly because of the low voltage output and high post-detection gain values in that channel. For this reason spikes are removed from the 44 GHz time-ordered data via template fitting, as described in Planck Collaboration II (2014).

### 3.2. Effects dependent on sky signal

#### 3.2.1. Sidelobe pick-up

Straylight contamination arises from the spurious signal pick-up from the telescope far sidelobes. Main sources of straylight contamination are the Galaxy, especially at 30 GHz, and the cosmological dipole, mainly detected in the directions of the main and sub-reflector spillover, as sketched in Fig. 8. In principle we should also include the stralight contribution from the orbital dipole, but its effect is a factor ten lower than the cosmic dipole, so that it can safely be neglected in this framework (but it has been considered in the calibration pipeline).

Intermediate sidelobes, i.e., the lobes in the pattern at angles less than 5° from the main beam, represent another source of systematic effects. The fraction of power intercepted by intermediate sidelobes ranges from 0.02% to 0.08% of the total beam power, which is about ten times less than the fraction in far sidelobes (ranging from 0.18% to 0.68%). Their effect is therefore correspondingly smaller, of ~1.5 μK on the maps. Moreover, because intermediate lobes involve sky regions very close to the main beam, their effect can be controlled by masking the Galaxy and point sources. In this paper we have therefore neglected the effect from intermediate sidelobes, which will be addressed in detail in a future paper dedicated to analysis of the full mission dataset.

Straylight affects the measured signal in two ways: (i) through direct contamination and coupling with the main beam sky signal; and (ii) in the photometric calibration of the radiometer detected signal. In this paper we concentrate on the direct detection, while the impact on calibration and the adopted mitigation strategies are described in Planck Collaboration V (2014).

Because of the beam orientation, the stralight fingerprint is different in odd surveys compared to even surveys. The Galaxy, for example, is detected by the sub-reflector spillover in the odd surveys and by the main-reflector spillover in the even surveys. Because the sub-reflector spillover points approximately in the main beam direction, the Galaxy stralight pattern is close to the Galactic plane. The main-reflector spillover, instead, points at about 85° from the main beam so that the Galaxy is re-imaged onto a ring (see figures in Sect. 4.3.1).

Further details about the Planck optical system are reported in Tauber et al. (2010), while the LFI and HFI beams and window functions are provided in Planck Collaboration IV (2014) and Planck Collaboration VII (2014), respectively.

#### 3.2.2. ADC non linearity

The ADC linearity requires that the voltage step sizes between successive binary outputs are constant over the entire input dynamic range. If these steps are not constant (see the sketch in Fig. 9) we have a non-linearity in the ADC response that leads to calibration errors. A brief description of the mathematical model of this effect is provided in Appendix A.

When there is a linear response, the voltage output of a coherent receiver scales linearly with the white noise. The typical
fingerprints of ADC non-linearity is a variation of the detector voltage output white noise not paired by a detectable variation in the voltage level. This effect was observed in the LFI radiometer data for the first time in flight, where drops of a few percent were observed in the voltage white noise but not in the output level over periods of a few weeks. Figure 10 shows this effect as a plot of relative white noise variation versus the detector output voltage for one of the most affected radiometer channels (the 44 GHz detector LFI25M-01).

The grey points represent an average over each pointing (about 40 min), while the solid line has been obtained by further binning the data in 200 bins over the plotted range in order to reduce the scatter and show more detail. The figure shows that the typical amplitude of the region where the non-linearity occurs is of the order of 1 mV, corresponding to about three bits in the ADC channels, because of their lower detector voltages.

The ADC non-linearity effect has been characterised from flight data and removed from the data streams according to the procedure described in Planck Collaboration II (2014). In Fig. 11 we show the same data as in Fig. 10 after the correction has been applied. The figure clearly shows that the anomalous white noise dips disappear after correction.

In general we cannot exclude other causes of this anomalous scaling of voltage with noise. The ADC linearity tests performed before launch were not sensitive enough to highlight this effect, and we could not perform post-launch tests on similar devices. On the other hand, the effect occurs repeatedly at specific values of the input ADC voltage and the ADC non-linearity model applied to correct the data proved effective. These facts give us confidence that this hypothesis is sound.

### 3.2.3. Imperfect photometric calibration
An important set of systematic effects are those related to the photometric calibration of the radiometers. Such effects are discussed in Planck Collaboration V (2014); here we will only provide the most important information to put the results of that paper within the context of this work.

There are three different kinds of systematic effects that can affect the calibration.

1. **Incorrect assumptions regarding the calibration signal.** In the case of LFI, the signal used for the calibration is the dipolar field caused by the motion of the solar system with respect to the CMB rest frame and by the motion of the spacecraft around the Sun. We model the former using the values quoted by Hinshaw et al. (2009) and the latter using the spacecraft’s attitude information. Any error in the numbers would lead directly to an error in the calibration of Planck-LFI data.

2. **Incorrect treatment of the calibration signal.** To actually use any previous knowledge of the CMB dipole, we need to convolve the signal with the beam response of the LFI radiometers. Any error in this step would produce a systematic effect in the map, not only because of the wrong shape expected for the calibration signal, but also because of the removal of the (wrong) dipole from the calibrated maps done by the Planck-LFI pipeline (Planck Collaboration II 2014). Possible types of errors include: wrong convolution of the expected dipole with the radiometer beams, incorrect masking of the Galaxy when fitting the observed signal with the dipole, etc.

3. **Incorrect reconstruction of gain fluctuations.** Some of the algorithms we used in calibrating LFI data for this release use the radiometer equation and the recorded variations of the radiometers total-power output to track gain changes. In principle, any deviation in the behaviour of the radiometer from the implemented model can induce systematic effects in the gain curves.

The calibration strategy and uncertainties for HFI are discussed in Planck Collaboration VIII (2014).

### 3.3. Pointing effects
Pointing uncertainties are translated into uncertainties in pixel temperature measurements. If pointing uncertainties are not constant in time, then the statistics of the sky anisotropy measurements are not preserved, with a consequent impact on power spectrum and cosmological parameters. For Planck-LFI, pointing uncertainties arise from two main effects:

1. **Satellite pointing determination.** The Planck Attitude Control Movement System guarantees a pointing accuracy of about 2″ (Planck Collaboration I 2014; Planck Science Office 2010), which is well within scientific requirements.
However, small non-idealities in the system and errors in the attitude reconstruction (caused, for example, by thermoelastic effects) can affect the data.

2. **Uncertainties in the FP geometry reconstruction.** The measurement of the Planck-LFI FP geometry is based on the determination of the beam pointing with respect to the nominal line of sight exploiting Jupiter observations. The peak of each beam has been determined by fitting data with a bi-variate Gaussian function which may not be representative of the real beam centre.

### 4. Assessing residual systematic effect uncertainties in maps and power spectra

In this section we discuss our assessment of the impact of residual systematic effects on maps and power spectra. This assessment has been performed according to two strategies: by "null maps" obtained by differencing maps with the same sky signal, in order to highlight residuals, and by simulating known systematic effects in the timelines, thereby exploiting a combination of flight data and measured instrumental properties.

#### 4.1. Null tests

We define a "null test" as any difference between two independent data sets which are anticipated to give nearly the same signal, on the assumption of perfect calibration, pointing reconstruction, and systematic effects removal. Null tests are a powerful means to assess the validity and self-consistency of Planck data on various time scales and across different dimensions (detector, frequency, time), and to highlight systematic effects above the white noise level.

The design of Planck and its observing strategy provides a wide range of opportunities for null tests, with sensitivity to different systematic effects and implications for the scientific outputs. Although we refer to these tests as null tests, the results are generally not featureless. Some of these features are caused by beam orientation and ellipticity that in turn cause spurious effects in odd minus even survey difference maps in correspondence to point sources and in the galactic plane. The analysis of the effect of beam ellipticity on the CMB power spectrum is provided in Planck Collaboration IV (2014) and will not be repeated here. A part of these residuals is caused by signal pick-up from beam sidelobes. Our beam model captures, at least partly, these fingerprints. Future work will be aimed at exploiting this model to remove the sidelobe signal from the data. A fraction of this large-scale residual has not been captured yet by our instrument model and will require further investigation to be understood and properly removed from the data.

In the Planck-LFI collaboration each internal data release is accompanied by a comprehensive set of null tests as a check of our processes and ongoing improvement in terms of systematic errors. In this section we report the results from the main tests supporting the systematic effect analysis for the first Planck public data release. Unless otherwise noted, the maps presented in this section are masked to remove point sources and to include only pixels measured in both maps. Difference maps are divided by 2 to be statistically consistent with average maps, and are smoothed to 2° FWHM to enhance large scale features.

#### 4.1.1. Systematic effects that are insensitive to null tests

Null maps are powerful means to understand residual systematic effects in the data, both of known and unknown origin, but they do not capture all possible effects. For example, fluctuations occurring on 20 min time scale would be undetected in half-ring difference maps, and fast fluctuations (like 1 Hz spikes and short time scale temperature variations) and effects arising from near sidelobes would not be revealed by survey difference maps. None of these effects, however, are likely to affect the results of our analysis significantly:

- Twenty-minute spurious fluctuations, if present, can be detected in power spectra calculated from time-ordered data (which are routinely calculated and assessed to derive noise properties). Their effect is strongly reduced by short (one second) baseline destriping map-making. In-flight LFI noise properties have been presented and discussed in Mennella et al. (2010).
- Effects at 1 Hz have been assessed from in-flight timelines by stacking data from all the mission (for each detector) in one second time windows. This allowed us to produce time-domain templates of the 1 Hz spurious signal that have been removed from the data at 44 GHz, which is the channel most affected.
- Short time scale temperature fluctuations at the level of the radiometers and of the 4 K reference loads are not expected since the LFI and HFI FPs act as lowpass thermal filters. Measurements from the LFI and HFI temperature sensors confirm that only the slow temperature fluctuations propagate from the cooler cold ends to the FP detectors and reference loads. Thermal transfer functions have been derived both before launch (Terenzi et al. 2009a) and in-flight (Gregorio et al. 2013). These have been used to produce the thermal systematic effect maps discussed in this paper.
- Near sidelobes can also produce spurious effects that are undetected in survey difference maps. The fraction of power intercepted by intermediate sidelobes ranges from 0.02% to 0.08% of the total beam power, about 10 times smaller than the fraction contained in far sidelobes (ranging from 0.18% to 0.68%). Their effect is therefore correspondingly smaller, of ~1.5 μK in the maps. In this paper we therefore neglect the effect from intermediate sidelobes, but these will be addressed in detail in a future full-mission analysis paper.

#### 4.1.2. Half-ring difference null tests

Half-ring difference null tests, constructed by taking a weighted difference between the first and second halves of each pointing period, are useful to assess the data noise properties and systematic effects on time scales smaller than about 20 min. Weights are calculated as explained in Sect. 9.2 of Planck Collaboration II (2014).

In Fig. 12 we show the half-ring difference maps for the three LFI frequencies. A simple quantitative test was performed by dividing them pixel-by-pixel by the square root of the white noise covariance maps (Planck Collaboration 2013) and checking the standard deviation of the resulting maps. We found this rms value to be very close to unity: 1.0211, 1.0089, and 1.0007 for 30, 44, and 70 GHz, respectively. The deviation from unity is consistent with the different level of 1/f noise in the three frequency channels (see Tables 1 and 10 of Planck Collaboration II 2014). A more complete quantitative analysis of these maps including cross-spectra analysis is reported in the “Data validation” section of Planck Collaboration II (2014).
4.2. Assessment of timeline-additive systematic effects

4.2.1. Thermal effects

Method. Thermal systematic effects maps have been generated using a simulation strategy that combines in-flight temperature sensor measurements (Mennella et al. 2010), thermal modelling of the propagation of temperature fluctuations (Tomasi et al. 2010) and radiometric transfer functions measured during ground tests (Terenzi et al. 2009b). Here we sketch the procedure used to combine these data into systematic effect maps.

For each temperature effect and for each receiver detector diode we choose the most representative sensor, generally the closest to the receiver. Housekeeping data are low-pass filtered to reduce noise contribution. The half-ring difference map noise is slightly lower than the survey difference map noise, simply because half-ring difference maps cover a longer time period than survey difference maps. For this reason we rescale the half-ring difference spectra by \( t_{HR}/t_{SD} \), where \( t_{HR} \) and \( t_{SD} \) are the average integration times in half-ring and survey difference maps, respectively.

For multipoles \( \ell \geq 30 \), the survey difference power spectra closely match the instrumental noise. For \( \ell < 30 \) there are additional residuals, especially at 30 GHz, which are partially captured by our simulations. These residuals, however, are at least two orders of magnitudes below the CMB power spectrum.

As a demonstration of the accuracy of our sidelobe model, we show in Fig. 14 the 30 GHz survey 1 minus survey 2 difference map after removing the residual sidelobe signal as predicted by our model (see Fig. 24). The blue ring structure disappears, confirming both the nature of this spurious feature in the map and the ability of our sidelobe model to capture a significant part of the large-scale residuals. Further investigation is needed to fully understand and remove the remaining level of spurious large-scale structures in our maps.

In Fig. 15 we show survey 1 minus survey 3 null test maps. These two surveys cover the sky in nearly identical orientations, and would be consistent with noise if calibration and other systematics were perfectly controlled. However, as seen in this figure, there are large-scale features also in these difference maps, and these are still under investigation. Because the first Planck cosmological release is based only on data from the first two surveys, a detailed study of effects present in data beyond survey 2 is outside the scope of this paper, and will be discussed in the second data release.

Next, we quantify the impact of residuals seen in our difference maps through angular power spectrum analyses. In Figs. 16 through 18 we compare the pseudo-spectra of odd-even survey difference maps (left column of Fig. 13) with the CMB spectrum filtered by the beam window function of each LFI channel; to spectra from simulated survey difference maps (right column of Fig. 13); and to spectra from half-ring difference maps, which estimate the noise contribution. The half-ring difference map noise is slightly lower than the survey difference map noise, simply because half-ring difference maps cover a longer time period than survey difference maps. For this reason we rescale the half-ring difference spectra by \( t_{HR}/t_{SD} \), where \( t_{HR} \) and \( t_{SD} \) are the average integration times in half-ring and survey difference maps, respectively.

For multipoles \( \ell \geq 30 \), the survey difference power spectra closely match the instrumental noise. For \( \ell < 30 \) there are additional residuals, especially at 30 GHz, which are partially captured by our simulations. These residuals, however, are at least two orders of magnitudes below the CMB power spectrum.

As a demonstration of the accuracy of our sidelobe model, we show in Fig. 14 the 30 GHz survey 1 minus survey 2 difference map after removing the residual sidelobe signal as predicted by our model (see Fig. 24). The blue ring structure disappears, confirming both the nature of this spurious feature in the map and the ability of our sidelobe model to capture a significant part of the large-scale residuals. Further investigation is needed to fully understand and remove the remaining level of spurious large-scale structures in our maps.

In Fig. 15 we show survey 1 minus survey 3 null test maps. These two surveys cover the sky in nearly identical orientations, and would be consistent with noise if calibration and other systematics were perfectly controlled. However, as seen in this figure, there are large-scale features also in these difference maps, and these are still under investigation. Because the first Planck cosmological release is based only on data from the first two surveys, a detailed study of effects present in data beyond survey 2 is outside the scope of this paper, and will be discussed in the second data release.

Next, we quantify the impact of residuals seen in our difference maps through angular power spectrum analyses. In Figs. 16 through 18 we compare the pseudo-spectra of odd-even survey difference maps (left column of Fig. 13) with the CMB spectrum filtered by the beam window function of each LFI channel; to spectra from simulated survey difference maps (right column of Fig. 13); and to spectra from half-ring difference maps, which estimate the noise contribution. The half-ring difference map noise is slightly lower than the survey difference map noise, simply because half-ring difference maps cover a longer time period than survey difference maps. For this reason we rescale the half-ring difference spectra by \( t_{HR}/t_{SD} \), where \( t_{HR} \) and \( t_{SD} \) are the average integration times in half-ring and survey difference maps, respectively.

For multipoles \( \ell \geq 30 \), the survey difference power spectra closely match the instrumental noise. For \( \ell < 30 \) there are additional residuals, especially at 30 GHz, which are partially captured by our simulations. These residuals, however, are at least two orders of magnitudes below the CMB power spectrum.

As a demonstration of the accuracy of our sidelobe model, we show in Fig. 14 the 30 GHz survey 1 minus survey 2 difference map after removing the residual sidelobe signal as predicted by our model (see Fig. 24). The blue ring structure disappears, confirming both the nature of this spurious feature in the map and the ability of our sidelobe model to capture a significant part of the large-scale residuals. Further investigation is needed to fully understand and remove the remaining level of spurious large-scale structures in our maps.

In Fig. 15 we show survey 1 minus survey 3 null test maps. These two surveys cover the sky in nearly identical orientations, and would be consistent with noise if calibration and other systematics were perfectly controlled. However, as seen in this figure, there are large-scale features also in these difference maps, and these are still under investigation. Because the first Planck cosmological release is based only on data from the first two surveys, a detailed study of effects present in data beyond survey 2 is outside the scope of this paper, and will be discussed in the second data release.

Next, we quantify the impact of residuals seen in our difference maps through angular power spectrum analyses. In Figs. 16 through 18 we compare the pseudo-spectra of odd-even survey difference maps (left column of Fig. 13) with the CMB spectrum filtered by the beam window function of each LFI channel; to spectra from simulated survey difference maps (right column of Fig. 13); and to spectra from half-ring difference maps, which estimate the noise contribution. The half-ring difference map noise is slightly lower than the survey difference map noise, simply because half-ring difference maps cover a longer time period than survey difference maps. For this reason we rescale the half-ring difference spectra by \( t_{HR}/t_{SD} \), where \( t_{HR} \) and \( t_{SD} \) are the average integration times in half-ring and survey difference maps, respectively.

For multipoles \( \ell \geq 30 \), the survey difference power spectra closely match the instrumental noise. For \( \ell < 30 \) there are additional residuals, especially at 30 GHz, which are partially captured by our simulations. These residuals, however, are at least two orders of magnitudes below the CMB power spectrum.

As a demonstration of the accuracy of our sidelobe model, we show in Fig. 14 the 30 GHz survey 1 minus survey 2 difference map after removing the residual sidelobe signal as predicted by our model (see Fig. 24). The blue ring structure disappears, confirming both the nature of this spurious feature in the map and the ability of our sidelobe model to capture a significant part of the large-scale residuals. Further investigation is needed to fully understand and remove the remaining level of spurious large-scale structures in our maps.

In Fig. 15 we show survey 1 minus survey 3 null test maps. These two surveys cover the sky in nearly identical orientations, and would be consistent with noise if calibration and other systematics were perfectly controlled. However, as seen in this figure, there are large-scale features also in these difference maps, and these are still under investigation. Because the first Planck cosmological release is based only on data from the first two surveys, a detailed study of effects present in data beyond survey 2 is outside the scope of this paper, and will be discussed in the second data release.

Next, we quantify the impact of residuals seen in our difference maps through angular power spectrum analyses. In Figs. 16 through 18 we compare the pseudo-spectra of odd-even survey difference maps (left column of Fig. 13) with the CMB spectrum filtered by the beam window function of each LFI channel; to spectra from simulated survey difference maps (right column of Fig. 13); and to spectra from half-ring difference maps, which estimate the noise contribution. The half-ring difference map noise is slightly lower than the survey difference map noise, simply because half-ring difference maps cover a longer time period than survey difference maps. For this reason we rescale the half-ring difference spectra by \( t_{HR}/t_{SD} \), where \( t_{HR} \) and \( t_{SD} \) are the average integration times in half-ring and survey difference maps, respectively.

For multipoles \( \ell \geq 30 \), the survey difference power spectra closely match the instrumental noise. For \( \ell < 30 \) there are additional residuals, especially at 30 GHz, which are partially captured by our simulations. These residuals, however, are at least two orders of magnitudes below the CMB power spectrum.

As a demonstration of the accuracy of our sidelobe model, we show in Fig. 14 the 30 GHz survey 1 minus survey 2 difference map after removing the residual sidelobe signal as predicted by our model (see Fig. 24). The blue ring structure disappears, confirming both the nature of this spurious feature in the map and the ability of our sidelobe model to capture a significant part of the large-scale residuals. Further investigation is needed to fully understand and remove the remaining level of spurious large-scale structures in our maps.

In Fig. 15 we show survey 1 minus survey 3 null test maps. These two surveys cover the sky in nearly identical orientations, and would be consistent with noise if calibration and other systematics were perfectly controlled. However, as seen in this figure, there are large-scale features also in these difference maps, and these are still under investigation. Because the first Planck cosmological release is based only on data from the first two surveys, a detailed study of effects present in data beyond survey 2 is outside the scope of this paper, and will be discussed in the second data release.

Next, we quantify the impact of residuals seen in our difference maps through angular power spectrum analyses. In Figs. 16 through 18 we compare the pseudo-spectra of odd-even survey difference maps (left column of Fig. 13) with the CMB spectrum filtered by the beam window function of each LFI channel; to spectra from simulated survey difference maps (right column of Fig. 13); and to spectra from half-ring difference maps, which estimate the noise contribution. The half-ring difference map noise is slightly lower than the survey difference map noise, simply because half-ring difference maps cover a longer time period than survey difference maps. For this reason we rescale the half-ring difference spectra by \( t_{HR}/t_{SD} \), where \( t_{HR} \) and \( t_{SD} \) are the average integration times in half-ring and survey difference maps, respectively.

For multipoles \( \ell \geq 30 \), the survey difference power spectra closely match the instrumental noise. For \( \ell < 30 \) there are additional residuals, especially at 30 GHz, which are partially captured by our simulations. These residuals, however, are at least two orders of magnitudes below the CMB power spectrum.
Fig. 13. Survey 1 minus survey 2 difference maps calculated from actual measurements (left column) and from simulations (right column), for 30, 44 and 70 GHz (top to bottom).

to remove high frequency sensor noise, Fourier-filtered to obtain the estimated temperature fluctuation at the receiver location, and then multiplied by the radiometric transfer function to obtain the simulated antenna temperature fluctuation on the undifferenced sky and reference load channels.

For each pointing period the average measured sky and reference load voltages are added to the two antenna temperature fluctuation data streams. After a weighted average of the two detector data values of each radiometer, we take the sky-load difference using the gain modulation factor, $r$, and multiply the resulting stream by the photometric constant, $G$. The weights, $r$, and $G$ are the same as used in the nominal pipeline to produce sky maps for that radiometer.

After oversampling to the receiver sampling frequency using linear interpolation, we use these data to build maps that account for the same in-flight pointings and map-making procedure as used to produce the final scientific products.
Results. In Fig. 19 we show the peak-to-peak amplitude of the various effects on final maps. Back-end temperature fluctuations have a sub-μK effect on maps. This low level can be understood if we consider that these fluctuations impact sky and reference load signals symmetrically and are effectively suppressed in the differential measurement. Furthermore the residual present in the data is a purely multiplicative effect, so it is essentially calibrated out through our gain model (Planck Collaboration V 2014).

Temperature variations in the 4 K reference loads couple with the radiometric output as an asymmetric additive spurious signal. In this case the relative calibration model provides no benefit, leaving a residual of about 1 μK peak-to-peak at 30 and 44 GHz. At 70 GHz, this effect is largely suppressed by the active thermal control system present on the HFI FP, close to the reference loads of this frequency channel.

Front-end 20 K temperature variations couple with the radiometric measurements through both gain and noise temperature fluctuations. For this reason the effect can only be partially calibrated out. Moreover, the asymmetry of the receiver chain before the orthomode transducer is such that the suppression provided by the sky-load differing is not optimal. The residual effect

Fig. 15. Survey 1 minus survey 3 difference maps: 30 GHz (top), 44 GHz (middle), and 70 GHz (bottom).

Fig. 16. Angular power spectra for 30 GHz null tests. Pseudo-spectra are calculated on 80% of the sky, with the Galaxy and point sources masked. All spectra are corrected for sky fraction but not for beam smearing effects. For comparison, we also show the simulated odd-even survey difference spectrum and the best-fit cosmological model spectrum filtered by the beam window function. For \( \ell \geq 30 \) the spectrum of the survey 1 minus survey 2 map difference fully coincides with the half-ring difference spectrum calculated for the same time period, while for the 30 GHz channel there is a small sidelobe contribution at the 10 μK² level on larger scales.

Fig. 17. Angular power spectra for 44 GHz null tests. In this case the low-\( \ell \) spectrum of the survey 1 minus survey 2 map difference is closer to that of the half-ring difference spectrum than at 30 GHz. Letters in the plot follow the same convention of the legend in Fig. 16.
Fig. 18. Angular power spectra for 70 GHz null tests. This channel has the smallest large-scale residuals among the three LFI channels. Letters in the plot follow the same convention of the legend in Fig. 16.

Fig. 19. Peak-to-peak thermal effects in maps. Notice the logarithmic scale on the ordinate axis.

is similar for the three frequency channels, and of the order of 1 μK peak-to-peak.

Maps of the combined thermal effects at the three LFI frequency channels are shown in Fig. 20.

4.2.2. Bias fluctuations

Method. The effect of bias fluctuations in the front-end amplifiers has been computed for maps and power spectra using the measured drain currents and a linear transfer function that links the drain currents of the two amplifiers to the radiometric output in antenna temperature. Since we are interested in assessing purely electrical instabilities, we correct the drain current housekeeping data for variations induced by temperature changes in the 20 K and 300 K temperature stages, i.e.,

\[ I_{\text{corr, drain}}(t) = I_{\text{drain}}(t) - \alpha_{20\,\text{K}} \delta T_{20\,\text{K}}(t) - \alpha_{300\,\text{K}} \delta T_{300\,\text{K}}(t), \]  

(2)

where \( \delta T_{20\,\text{K}}(t) \) and \( \delta T_{300\,\text{K}}(t) \) are temperature variations on the 20 K and 300 K temperature units, respectively, and \( \alpha_{20\,\text{K}} \) and \( \alpha_{300\,\text{K}} \) are the corresponding drain current thermal susceptibility coefficients, calculated using an iterative linear fitting process. First we calculate the coefficients of the susceptibility to back-end temperature fluctuations, exploiting the temperature change induced by the change in the transponder state, which occurred at day 258 (see Fig. 7), and then we determine the coefficients of the susceptibility to front-end temperature fluctuations using data from a temperature susceptibility test run at the end of the in-flight calibration phase. The iterative process is closed by recalculating back-end thermal coefficients after correcting drain currents for front-end temperature fluctuations.

Following thermal correction we correlate drain current changes with antenna temperature variations in sky and reference load samples. We recall here that the LFI receiver architecture implies that the signal characteristics at each detector depend on both radiometer front-end amplifiers (Bersanelli et al. 2010). Thus, for each detector diode we first calculate the weight, \( w \), providing the maximum correlation between the output voltage and the linear combination of the drain currents of the two radiometer amplifiers, then a linear fit between this combination and the voltage output provided the required transfer function. Mathematically the relationship between the
corrected drain current fluctuations, \( \delta I_{\text{corr}, \text{drain}} \), and the voltage output variations, \( \delta V_{\text{sky(ref)}} \), reads
\[
\delta V_{\text{sky(ref)}}(t) = \alpha_{\text{sky(ref)}} \left[ w \delta I_{\text{corr}, 1}(t) + (k - w) \delta I_{\text{corr}, 2}(t) \right],
\] (3)
where \( k \) is a constant, \( w \) is the weight and \( \alpha_{\text{sky(ref)}} \) is the slope of the linear fit between the weighted combination of the two drain currents and the sky (reference load) voltage outputs. The time-ordered data obtained by Eq. (3) are finally projected onto the sky using flight pointings.

Results. Maps of the residual effect arising from bias fluctuations are shown in Fig. 21. This effect is smaller than 1 \( \mu \text{K} \) peak-to-peak at all frequencies and presents little structure apart from a stripe in the 70 GHz map, caused by a jump in the bias voltage occurring at day 258, following the change in transponder state. This jump affected in particular the 70 GHz radiometers, leaving a small signature in the maps.

Fig. 21. Maps of the systematic effect from drain current fluctuations at 30 GHz (top), 44 GHz (middle) and 70 GHz (bottom).

Fig. 22. Maps of 1 Hz spikes at 30 GHz (top), 44 GHz (middle) and 70 GHz (bottom). The map at 44 GHz represents the residual after the spike signal has been removed from the time-ordered data, while maps at 30 and 70 GHz represent the spike signal with no removal applied.

4.2.3. 1 Hz spikes

Method. Time ordered data containing the spike signal are generated using templates obtained from flight radiometric data. The details of this method are described in Sect. 7.1 of Mennella et al. (2011), and will not be repeated here.

Results. In Fig. 22 we show maps of the spike systematic effect at the three LFI frequencies. Because spikes are removed from the 44 GHz channel, but not from 30 and 70 GHz, the corresponding maps represent the residual effect after removal at 44 GHz, and the spike effect with no removal applied at 30 and 70 GHz. In all the three channels the rms effect is at the sub-\( \mu \text{K} \) level.
4.3. Assessment of effects dependent on the sky

4.3.1. Far sidelobes

Method. The external straylight contamination is evaluated with simulations in which the sky model includes the diffuse Galactic emission and the dipole, the two most important sources of external straylight contamination. At 30 GHz, the straylight assessment includes the beam frequency dependence and the receiver in-band response (see Zonca et al. 2009; Planck Collaboration IX 2014) by dividing the bandpass response into discrete frequency intervals. For each frequency interval a weight factor is calculated as the integral of the bandpass response over the interval itself. In Fig. 23 we show, as an example, the bandpass response of the LFI27 receiver (main and side arms) and the seven frequency intervals considered in the simulations. The weights correspond to the integral of the bandpass response curve over the frequency interval. In parallel, far sidelobes are computed using the GRASP MrGTD\footnote{Multi-reflector geometrical theory of diffraction.} software (www.ticra.com) at the frequencies indicated on the top of each slice reported in Fig. 23 (27, 28, 29, 30, 31, 32, and 33 GHz). The optical model used in GRASS simulations is reported in Planck Collaboration IV (2014). For each frequency interval an observation of the sky model is simulated for all 30 GHz detectors using the beam sidelobes and the real sky pointings, neglecting beam smearing effects and weighting the data stream with the above mentioned weight factors. Finally we run the Madam map-making code to generate maps from simulated data streams.

Results. In Figs. 24 through 26 we show the simulated sidelobe fingerprint on the sky after the destriping process, for the odd (left side) and even (right side) surveys, respectively. These figures show that the straylight from the cosmological dipole is similar in the two surveys, while the Galaxy straylight is, as expected, larger in the second. The ring-shaped fingerprint in the second survey is also observed at the expected level in the real data by taking the difference of even minus odd survey maps (see Figs. 13 and 14 in Sect. 4.1), thus confirming the accuracy of our simulations.

These results show that the most sensitive channel to straylight is 30 GHz, followed in order by 70 GHz and 44 GHz. This is consistent with the telescope optical performance at the various frequencies. The primary mirror is strongly under-illuminated by the 44 GHz horns, resulting in a low straylight sensitivity at the expense of a larger main beam, especially for the LFI25 and LFI26 horns. The 30 and 70 GHz horns are characterised by similar illumination properties, so that their straylight susceptibilities are comparable, with a slightly better performance of the 70 GHz horns with respect to the 30 GHz ones. If we also take the higher sensitivity of the 30 GHz channel to the Galactic signal into account, it is apparent that this channel is, overall, the most susceptible to straylight contamination.
4.3.2. ADC non-linearity

**Method.** The levels of the residuals due to ADC correction process are estimated by applying the correction algorithm to simulated data containing a known ADC effect and making difference maps with those produced from data with no ADC effect. The starting point of this analysis is time-ordered data for all individual detectors based on (ring-based) sky and reference load simulations. The noise component is simulated using average 1/f noise parameters over the nominal mission for each detector. Galactic and CMB signals are based on the observed maps, converted into time-ordered data using real pointings and successively uncalibrated using the inverse gain table. The same is done with a map of the WMAP dipole, while the orbital dipole is calculated from the pointing information and JPL ephemeris for the satellite velocity. Finally spline fits to the observed sky and reference voltage levels per pointing period are used together with estimated receiver temperatures, $T_{\text{CMB}} = 2.725 \text{ K}$ and $T_{\text{ref}} = 4.5 \text{ K}$, as a model for the gain evolution.

The simulated ADC effect is induced by applying the inverse of the spline correction used in the real data. The same algorithm as used with the real data is then applied to the simulated data iteratively five times to ensure convergence. Intensity maps are constructed by simple binning into an $N_{\text{side}} = 1024$ map at each iteration, both for the simulation with and without the ADC effect. Some of these maps show a residual dipole caused by small changes in the overall slope of the temperature-voltage response curve due to the ADC correction. Since the calibration pipeline determines this response and does not give rise to a residual dipole, we correspondingly remove it here via a correlation fit with the input dipole map. The ADC effect maps are finally taken as the difference between the fifth iteration map and the no-ADC map. Maps for each frequency band are produced by averaging all maps for that frequency, taking the detector weighting into account.

Since some 70 GHz channels cannot be corrected due to “popcorn” noise, a separate method was used to estimate the likely level of ADC error for these channels, using the white noise level on the difference data. This is immune to the “popcorn” noise, but cannot be used to correct the ADC effect, since it is not known whether the effect is due to the sky or to reference voltages. In these cases we only estimate the ADC effect and do not apply any correction.

**Results.** Maps of the ADC effect at the three LFI frequency channels are shown in Fig. 28. The main effect of the ADC residuals is a small (<0.1% of the dipole signal) ring-based gain error which appears in the maps as stripes in the scan direction. The residuals are generally larger where the sky signal is stronger, i.e., following the CMB dipole and Galactic plane. The contribution from the Galactic plane becomes weaker at higher frequencies as expected. Broad stripes in the 30 GHz map are due to residual deviations from linearity on voltages ranges larger than the ADC peaks. These also occur at the other frequencies, but as the number of channels increases this effect averages out, leading to a well-determined correction placing it between 30 and 70 GHz in terms of the amplitude of residuals.

4.3.3. Imperfect photometric calibration

**Method.** We have developed an analytical model of the impact of the uncertainty in the dipole calibration algorithm due to the radiometer white noise and the loss of integration time due to Galactic masking. We ran this model to estimate how this effect propagates through the calibration and mapmaking pipeline. Such simulations scan a sky map (the *input map*) of pure astrophysical signal (without dipole) to produce a time-ordered data stream, which is then uncalibrated using gains inferred from the total-power output of the radiometers. These time-ordered data are then used as input in a simplified version of the LFI pipeline to produce a new calibrated map (the *output map*). The difference between the input and output maps should be mainly due
to dipole leakage, since the gains used in the decalibration phase differ from those calculated by the pipeline. Refer to Sect. 5 in (Planck Collaboration V 2014) for more information.

**Results.** Figure 29 shows the difference between the input and output maps. The shape of the features in these maps closely follows the scanning circles drawn by the pointing direction of the telescope towards the sky. (This is expected, since the calibration is performed on the time-ordered data.) The estimated impact of such systematic effects on the Planck-LFI maps is of a few $\mu K$ per pixel.

**4.4. Pointing uncertainties**

Method. To estimate the uncertainty introduced by the Gaussian fit in main beam measurements we perform tests using the radio-frequency model of the flight telescope (Planck Collaboration IV 2014) and compare the centre calculated by the fit with the beam maximum, which is uniquely determined in optical simulations. Typical differences between the centres are $1''$ for all the 70 GHz beams, $4''$ for LFI24, $18''$ for LFI25 and LFI26 (44 GHz horns), and $6''$ for the 30 GHz beams. These estimates are all smaller than the statistical uncertainty in the determination of the beam centre, which ranges from $4''$ at 70 GHz to $10''$ at 30 GHz.

The FP geometry was reconstructed using four Jupiter transits labelled as J1, J2, J3 and J4 (Planck Collaboration II 2014; Planck Collaboration IV 2014). When we compare the FP geometry obtained from the combination of J1 and J2 with the one obtained from the combination of J3 and J4 we find a difference of about $15''$ in pointing, mainly along the in-scan direction. On the other hand, the comparison of the focal plane geometries determined from single Jupiter transits (J1 against J2 and J3 against J4) shows differences within the expected uncertainty.
other for the period between day 540 and day 563. Details of the period ranging from day 91 to day 539 after launch, and the different instrument databases in the data analysis pipeline, one for instrument digital electronics, is compensated for using two different thermal control set-point of the data processing unit in the instrument.

A3, page 18 of 23

The 15° discrepancy, likely to be correlated with changes in the thermal control set-point of the data processing unit in the instrument digital electronics, is compensated for using two different instrument databases in the data analysis pipeline, one for the period ranging from day 91 to day 539 after launch, and the other for the period between day 540 and day 563. Details of the FP database solution and related uncertainties can be found in Planck Collaboration II (2014).

We assess the impact of this effect using dedicated simulations constructed according to the following procedure:

1. Generate time-ordered data by observing a CMB-only sky with flight detector pointing derived by applying the two-FP database solution.
2. Reconstruct the CMB map from the time-ordered data generated in step 1, applying each of the two-FP database solution in map reconstruction.
3. Repeat step 2 using the single FP database solution.
4. Compute the difference of the power spectra obtained from the two generated maps.

Results. Figure 30 shows that the relative difference of power spectra is of the order of $10^{-4}$, which is negligible.

4.5. Propagation of systematic uncertainties through component separation

A further step in our assessment has been to evaluate the impact of the various systematic effects on the CMB map independently from the frequency. To do this we computed a weighted sum of the three maps for each effect using weights obtained derived with a pixel-based ILC (internal linear combination) component separation method (Leach et al. 2008; Planck Collaboration XII 2014). The ILC method implements direct variance minimisation exploiting the fact that the CMB component (in thermodynamic temperature units) is constant across frequencies, while foregrounds are characterised by non-thermal spectra. The CMB temperature can then be estimated at each pixel, $p$, in terms of a simple weighted sum of the frequency maps,

$$ T_{\text{CMB}}(p) = \sum_{i=1}^{N} w_i T_i(p) \quad \text{where} \quad \sum_{i=1}^{N} w_i = 1. $$

(4)

The ILC coefficients are estimated including Planck frequencies between 30 and 353 GHz. However, only the three LFI channels are included in the total systematic error map, because we are only interested in residual LFI systematic effects in the CMB products. To propagate systematic effects through component separation, we therefore replace the frequency maps in Eq. (4) with the corresponding systematic effect maps,

$$ T_{\text{syst}}(p) = \sum_{i=1}^{3} w_i T_{\text{syst},i}(p). $$

(5)

Note that for simplicity, the ILC weights are uniformly distributed in pixel and harmonic domains, whereas in the Planck component separation pipeline, the variance minimisation is conducted in the needlet space, i.e., on sub-sets of the harmonic and pixel domains where foregrounds are relevant at various levels, resulting in a set of coefficients for each needlet domain (Planck Collaboration XII 2014).

4.6. Gaussianity statistical tests

Finally we assess non-Gaussianity induced by known systematic effects in the LFI maps. We present results derived with different non-Gaussianity tests carried out at each frequency using the map obtained by summing the various systematic effects considered in this paper.

For detailed information on the non-directional or targeted non-Gaussianity tests with the Planck data, we refer the interested reader to Planck Collaboration XXIII (2014); Planck Collaboration XXIV (2014); Planck Collaboration XIX (2014). Here we consider the subset consisting of Minkowski functionals (Schmalzing & Gorski 1998), statistical quantities derived from the one-point PDF (variance, skewness, kurtosis and the Kolmogorov-Smirnov or KS distance) and the skewness and kurtosis of the spherical Mexican hat wavelet (SMHW, Martínez-González et al. 2002). The properties of these estimators are described in Planck Collaboration XXIII (2014) and references therein.

We compare the values of our estimators derived from a set of ideal Gaussian CMB and noise realisations with those obtained from the same CMB and noise simulations to which the systematic effect maps are added. The CMB and noise maps were simulated following the Planck-LFI data processing pipeline (Planck Collaboration II 2014). Where the three estimators did not provide significant deviations between the maps with and without the systematic effects, we have carried out an additional test by rescaling the systematic effect maps with a constant factor in order to provide an estimate of the amplitude required to detect significant deviations with respect to the CMB signal (i.e., larger than 3σ or 99% confidence level).

Figure 31 shows the three Minkowski functionals for the three LFI frequency bands. In each panel we compare the ±1σ (68% confidence band centred on the mean corresponding to the Gaussian CMB plus noise simulations, with the same simulations with systematic effects added for HEALPix resolutions $N_{\text{side}} = 512, 256, \text{and } 128$. Our analysis shows that the estimators based on the Minkowski functionals are not affected by the presence of systematic effects in the maps.

Fig. 30. Relative difference between the spectra of the maps simulated with one and two instrument databases. The three curves represent power spectra relative to three different feed horn pairs in the 70 GHz frequency channel. A running average smoothing kernel has been applied to reduce the scatter and enhance any larger-scale trends. The relative uncertainties are of the order of $\Delta C_i / C_i \lesssim 10^{-4}$.
Fig. 31. The three Minkowski functionals computed for Gaussian CMB and noise simulations (black symbols) compared with the Minkowski functionals computed for the same simulations with systematic effects added (solid magenta line). From left to right: the area, contour length or perimeter, and the genus. From top to bottom: the three LFI frequencies, 30, 44, and 70 GHz. These are explicitly for: \( N_{\text{side}} = 512 \).

Table 3 contains the difference, \( \Delta \), of the mean of the two distributions (maps with and without systematic effects), normalised by its dispersion and multiplied by 100, corresponding to the standard deviation, skewness and kurtosis for the three LFI frequency bands, as detailed by Eq. (6):

\[
\Delta \equiv \frac{100}{\sigma(X_{\text{sys}})} \left( \langle X_{\text{sys}} \rangle - \langle X_{\text{clean}} \rangle \right)
\]

(6)

Here \( X_{\text{sys}} \) represents each of the considered statistics corresponding to the maps with systematic effects, pixelised at a HEALPix resolution of \( N_{\text{side}} = 1024 \), and \( X_{\text{clean}} \) represents each of the considered statistics corresponding to the maps without systematic effects. There are no significant deviations, as the distributions corresponding to the two types of map are virtually superimposed.

Table 4 shows the difference of the mean of the two distributions (maps with and without systematic effects), normalised by its dispersion and multiplied by 100 (see Eq. (6)), and corresponding to the skewness and the kurtosis of the SMHW for the LFI frequency bands. The list of angular scales selected for this analysis is the one used in Planck Collaboration XXIV (2014) for the \( f_{\text{gal}} \) estimation, and comprises 16 angular scales between 1.3' and 956.3' with logarithmic spacing. Again, no significant deviations are seen. As an additional check, the sum of the systematic effects at 70 GHz has been directly translated into a \( f_{\text{gal}} \) estimate for the local shape, resulting in a value of \( \Delta f_{\text{gal}} = -0.06 \) and a relative deviation of \( \Delta f_{\text{gal}} / \sigma(f_{\text{gal}}) = -0.003 \); the impact of known systematics effects at 70 GHz on primordial non-Gaussianity is negligible.

To conclude, we characterise the levels of detectability of the non-Gaussian contamination of these systematic effect maps. Adopting the one-point PDF and Minkowski functionals statistics, we employ simulations with different levels of systematic effects,

\[
\Delta T(n) = \Delta T_{\text{CMB}}(n) + \Delta T_{\text{noise}}(n) + f \Delta T_{\text{syst}}(n),
\]

(7)
to estimate the factor $f$ at which level the systematic effect is detectable. This level is taken to be the value of $f$ for which any of the estimators is outside the $3\sigma$ confidence level of the values corresponding to maps without systematic effects.

The results indicate that the minimum values of $f$ are $f \sim 8$ at 30 GHz, $f \sim 12$ at 44 GHz, and $f \sim 7$ at 70 GHz. To conclude, systematic effects do not generate significant levels of non-Gaussianity for the temperature maps at the three LFI frequencies.

5. Conclusions

In this paper we analysed and quantified the uncertainties on Planck-LFI CMB temperature anisotropy measurements arising from systematic effects along two complementary approaches. On the one hand, we adopted a top-down approach, in which surplus excess signals are highlighted by a series of dedicated null-tests in which maps containing the same sky signal are differentiated to obtain maps containing noise and systematic effects. On the other hand, we followed a bottom-up approach in which each known effect is simulated in terms of timelines and maps.

Our analysis shows that systematic effect uncertainties are at least two orders of magnitude below the CMB temperature anisotropy power spectrum. The two dominant effects are stray-light pick-up from far sidelobes and imperfect photometric calibration. In this current data release the sidelobe signal is not removed from the data, although the CMB dipole pick-up by far sidelobes is accounted for during the calibration process using a monochromatic model.

Statistical analyses performed on maps containing the sum of all the simulated systematic effects added to a simulated CMB map showed no detectable non-Gaussianity levels unless their level was artificially increased by a factor ranging from 7 to 12. This confirms that instrumental effects do not significantly impact Gaussianity studies.

Survey difference maps show a signal excess in the multipole range $\ell < 20$ that is only partially accounted for in the simulated maps. This excess could be caused by yet un-modelled stray-light pick-up affecting the measurements both directly and in the photometric calibration process.

Currently, analysis focusses on understanding, and further reducing, the level of systematic uncertainties in view of the 2014 data release, which will include polarisation data and results. Areas of activity include a more thorough in-band modelling of the sidelobe response at all frequencies, aimed at removing the spurious signal from timelines, and better correction in the calibration step.

Acknowledgements. The development of Planck has been supported by: ESA; CNES and CNRS/INSU-IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MICINN, JA and RES (Spain); Teles, AofC and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); and PRACE (EU). We acknowledge the computer resources and technical assistance provided by the Spanish Supercomputing Network nodes at Universidad de Cantabria and Universidad Politécnica de Madrid as well as the support provided by the Advanced Computing and e-Science team at IFCA. A description of the Planck Collaboration and a list of its members, including the technical or scientific activities in which they have been involved, can be found at http://www.sciops.esa.int/index.php?project=planck&page=Planck_Collaboration.

Appendix A: Theory of the ADC non-linearity effect

ADC non-linearity arises when the measured detector voltage differs from the true voltage in some repeatable manner, depending on the exact values of the voltage thresholds of the chip. By mapping the apparent voltage, $V'$, to the true voltage, $V$, the ADC effect can be corrected and this mapping is precisely the ADC response curve, $R(V')$, as measured through the LFI acquisition system. In a perfect radiometer this voltage is the product of the system temperature, $T_{\text{sys}}$, and radiometer gain, $G(t)$,

$$V = V'R(V') = G(t)T_{\text{sys}}.$$  \hfill (A.1)

Probing the response function requires tracking small known input voltage variations, $\Delta V$, in terms of a measured $\Delta V'$ at various working voltages, $V'$. This can be illustrated by differentiating the above equation with respect to $V'$,

$$\Delta V = \left( V' \frac{dR(V')}{dV'} + R(V') \right) \Delta V' = G(t) \Delta T.$$  \hfill (A.2)

Equation (A.2) shows the relation between the differential input and output signals, and illustrates how a localised gradient change can dominate via the $dR/dV'$ term. It also shows that small intrinsic thermal noise fluctuations, $\Delta T$, can be used as a test input temperature signal, assuming it is due to bandwidth limited noise power, $\Delta T = T_{\text{sys}}/\sqrt{\nu\tau}$, where $\nu$ and $\tau$ are channel bandwidth and sample integration time, respectively. By combining the two previous equations, the differential response can be expressed as

$$\frac{dR(V')}{dV'} = \left( \frac{1}{\sqrt{\nu\tau}} - \frac{1}{V} \right) R(V').$$  \hfill (A.3)

In the case of no ADC effects and voltage variations induced purely through gain fluctuations, we have $V' = \sqrt{\nu\tau} \Delta V'$, and
the differential response $dR(V')/dV'$ remains zero for all $V'$, as expected. Non-linearities are signaled where the thermal white noise does not follow detector voltages, revealing variations in the response curve. Since the radiometer gains drift very slowly, many estimates of white detector noise by Fourier analysis from the one minute scan rings are available, and by binning and averaging signal-to-noise ratios of ÷100 are achievable. The above equation can be integrated numerically, making use of these binned values, as $R(V') \approx 1$ is a good approximation. A discrete set of corrected voltages $V'_i$ for each binned measured voltages $V_i$ can be found via a trapezoidal summation,

$$V'_i = V'_0 + \frac{\delta V'}{2} \sum_{j=1}^{i} a \left( \frac{1}{\Delta V'_{i-1}} + \frac{1}{\Delta V'_{j}} \right) - \left( \frac{1}{V'_{i-1}} + \frac{1}{V'_{j}} \right).$$  \tag{A.4}

Here $V'_0$ is lowest voltage bin, $\delta V'$ is the voltage bin width, and $a = 1/\sqrt{\Delta V'}$ is fitted such that the top voltage bin, $V'_{\text{max}}$, is equal to $V'_0$ to maintain the same overall linear response. The tables of corrected voltages to measured voltages for each detector are stored in the DPC database as the ADC correction, and are implemented as spline fits when correcting the time-ordered data.

### Appendix B: ADC error before and after correction

An estimator of the magnitude of the ADC effect is the relative variation in the white noise ratio of “sky” samples to the “reference load” samples. In fact, this removes the effect of the noise variations by comparing the ADC linearity between two well separated voltage levels. These estimates are given in Table B.1 for all the LFI detectors, before and after the correction. In boldface we show the channels that were actually corrected for this effect.

Some channels that were not corrected for the ADC effect because the total power data were affected by so-called “popcorn” noise, i.e. random jumps in the total power voltage that were irrelevant for the map-making but not for the ADC removal algorithm that is based on the total power data.

### References


5 These jumps were present both in the sky and reference load data so that they were effectively removed by differencing.