

## Planck early results. V. The Low Frequency Instrument data processing<sup>★</sup>

A. Zacchei<sup>36</sup>, D. Maino<sup>22,38</sup>, C. Baccigalupi<sup>54</sup>, M. Bersanelli<sup>22,38</sup>, A. Bonaldi<sup>34</sup>, L. Bonavera<sup>54,5</sup>, C. Burigana<sup>37</sup>, R. C. Butler<sup>37</sup>, F. Cuttaia<sup>37</sup>, G. de Zotti<sup>34,54</sup>, J. Dick<sup>54</sup>, M. Frailis<sup>36</sup>, S. Galeotta<sup>36</sup>, J. González-Nuevo<sup>54</sup>, K. M. Górski<sup>48,58</sup>, A. Gregorio<sup>23</sup>, E. Keihänen<sup>15</sup>, R. Keskitalo<sup>48,15</sup>, J. Knoche<sup>52</sup>, H. Kurki-Suonio<sup>15,31</sup>, C. R. Lawrence<sup>48</sup>, S. Leach<sup>54</sup>, J. P. Leahy<sup>49</sup>, M. López-Cañiego<sup>47</sup>, N. Mandolesi<sup>37</sup>, M. Maris<sup>36</sup>, F. Matthai<sup>52</sup>, P. R. Meinhold<sup>18</sup>, A. Mennella<sup>22,36</sup>, G. Morgante<sup>37</sup>, N. Morisset<sup>40</sup>, P. Natoli<sup>24,2,37</sup>, F. Pasian<sup>36</sup>, F. Perrotta<sup>54</sup>, G. Polenta<sup>2,35</sup>, T. Poutanen<sup>31,15,1</sup>, M. Reinecke<sup>52</sup>, S. Ricciardi<sup>37</sup>, R. Rohlf<sup>40</sup>, M. Sandri<sup>37</sup>, A.-S. Suur-Uski<sup>15,31</sup>, J. A. Tauber<sup>29</sup>, D. Tavagnacco<sup>36</sup>, L. Terenzi<sup>37</sup>, M. Tomasi<sup>22,38</sup>, J. Valiviita<sup>45</sup>, F. Villa<sup>37</sup>, A. Zonca<sup>18</sup>, A. J. Banday<sup>57,7,52</sup>, R. B. Barreiro<sup>47</sup>, J. G. Bartlett<sup>4,48</sup>, N. Bartolo<sup>20</sup>, L. Bedini<sup>6</sup>, K. Bennett<sup>29</sup>, P. Binko<sup>40</sup>, J. Borrill<sup>51,55</sup>, F. R. Bouchet<sup>43</sup>, M. Bremer<sup>29</sup>, P. Cabella<sup>25</sup>, B. Cappellini<sup>38</sup>, X. Chen<sup>41</sup>, L. Colombo<sup>14,48</sup>, M. Cruz<sup>12</sup>, A. Curto<sup>47</sup>, L. Danese<sup>54</sup>, R. D. Davies<sup>49</sup>, R. J. Davis<sup>49</sup>, G. de Gasperis<sup>25</sup>, A. de Rosa<sup>37</sup>, G. de Troia<sup>25</sup>, C. Dickinson<sup>49</sup>, J. M. Diego<sup>47</sup>, S. Donzelli<sup>38,45</sup>, U. Dörl<sup>52</sup>, G. Efstathiou<sup>44</sup>, T. A. Enßlin<sup>52</sup>, H. K. Eriksen<sup>45</sup>, M. C. Falvela<sup>3</sup>, F. Finelli<sup>37</sup>, E. Franceschi<sup>37</sup>, T. C. Gaier<sup>48</sup>, F. Gasparo<sup>36</sup>, R. T. Génova-Santos<sup>46,27</sup>, G. Giardino<sup>29</sup>, F. Gómez<sup>46</sup>, A. Gruppuso<sup>37</sup>, F. K. Hansen<sup>45</sup>, R. Hell<sup>52</sup>, D. Herranz<sup>47</sup>, W. Hovest<sup>52</sup>, M. Huynh<sup>41</sup>, J. Jewell<sup>48</sup>, M. Juvela<sup>15</sup>, T. S. Kisner<sup>51</sup>, L. Knox<sup>17</sup>, A. Lähteenmäki<sup>1,31</sup>, J.-M. Lamarre<sup>50</sup>, R. Leonardi<sup>28,29,18</sup>, J. León-Tavares<sup>1</sup>, P. B. Lilje<sup>45,9</sup>, P. M. Lubin<sup>18</sup>, G. Maggio<sup>36</sup>, D. Marinucci<sup>26</sup>, E. Martínez-González<sup>47</sup>, M. Massardi<sup>34</sup>, S. Matarrese<sup>20</sup>, M. T. Meharga<sup>40</sup>, A. Melchiorri<sup>21</sup>, M. Migliaccio<sup>25</sup>, S. Mitra<sup>48</sup>, A. Moss<sup>13</sup>, H. U. Nørgaard-Nielsen<sup>10</sup>, L. Pagano<sup>48</sup>, R. Paladini<sup>56,8</sup>, D. Paoletti<sup>37</sup>, B. Partridge<sup>30</sup>, D. Pearson<sup>48</sup>, V. Pettorino<sup>54</sup>, D. Pietrobon<sup>48</sup>, G. Prézeau<sup>8,48</sup>, P. Procopio<sup>37</sup>, J.-L. Puget<sup>42</sup>, C. Quercellini<sup>25</sup>, J. P. Rachen<sup>52</sup>, R. Rebolo<sup>46,27</sup>, G. Robbers<sup>52</sup>, G. Rocha<sup>48,8</sup>, J. A. Rubiño-Martín<sup>46,27</sup>, E. Salerno<sup>6</sup>, M. Savelainen<sup>15,31</sup>, D. Scott<sup>13</sup>, M. D. Seiffert<sup>48,8</sup>, J. I. Silk<sup>19</sup>, G. F. Smoot<sup>16,51,4</sup>, J. Sternberg<sup>29</sup>, F. Stivoli<sup>39</sup>, R. Stompor<sup>4</sup>, G. Tofani<sup>32</sup>, L. Toffolatti<sup>11</sup>, J. Tuovinen<sup>53</sup>, M. Türlér<sup>40</sup>, G. Umam<sup>33</sup>, P. Vielva<sup>47</sup>, N. Vittorio<sup>25</sup>, C. Vuerli<sup>36</sup>, L. A. Wade<sup>48</sup>, R. Watson<sup>49</sup>, S. D. M. White<sup>52</sup>, and A. Wilkinson<sup>49</sup>

(Affiliations can be found after the references)

Received 9 January 2011 / Accepted 19 July 2011

### ABSTRACT

We describe the processing of data from the Low Frequency Instrument (LFI) used in production of the *Planck* Early Release Compact Source Catalogue (ERCSC). In particular, we discuss the steps involved in reducing the data from telemetry packets to cleaned, calibrated, time-ordered data (TOD) and frequency maps. Data are continuously calibrated using the modulation of the temperature of the cosmic microwave background radiation induced by the motion of the spacecraft. Noise properties are estimated from TOD from which the sky signal has been removed using a generalized least square map-making algorithm. Measured  $1/f$  noise knee-frequencies range from  $\sim 100$  mHz at 30 GHz to a few tens of mHz at 70 GHz. A destriping code (Madam) is employed to combine radiometric data and pointing information into sky maps, minimizing the variance of correlated noise. Noise covariance matrices required to compute statistical uncertainties on LFI and *Planck* products are also produced. Main beams are estimated down to the  $\approx -10$  dB level using Jupiter transits, which are also used for geometrical calibration of the focal plane.

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

### 1. Introduction

*Planck*<sup>1</sup> (Tauber et al. 2010; Planck Collaboration 2011a) is a third generation space mission to measure the anisotropy of the cosmic microwave background (CMB). It observes the sky in nine frequency bands covering 30–857 GHz with high sensitivity and angular resolution from 31' to 5'. The Low Frequency Instrument (LFI) (Mandolesi et al. 2010; Bersanelli et al. 2010; Mennella et al. 2011) covers the 30, 44, and 70 GHz bands with amplifiers cooled to 20 K. The High Frequency Instrument (HFI)

(Lamarre et al. 2010; Planck HFI Core Team 2011a) covers the 100, 143, 217, 353, 545, and 857 GHz bands with bolometers cooled to 0.1 K. Polarization is measured in all but the highest two bands (Leahy et al. 2010; Rosset et al. 2010). A combination of radiative cooling and three mechanical coolers produces the temperatures needed for the detectors and optics (Planck Collaboration 2011b). Two Data Processing Centres (DPCs), conceived as interacting and complementary since the earliest design of the *Planck* scientific ground segment (Pasian & Gispert 2000); check and calibrate the data and make maps of the sky, this paper and (Planck HFI Core Team 2011b). *Planck*'s sensitivity, angular resolution, and frequency coverage make it a powerful instrument for galactic and extragalactic astrophysics, as well as cosmology. Early astrophysics results are given in Planck Collaboration, 2011c–z.

The Low Frequency Instrument LFI on *Planck* comprises a set of 11 radiometer chain assemblies (RCAs), each composed of two independent, pseudo-correlation radiometers. There are two

<sup>★</sup> Corresponding author: A. Zacchei,  
e-mail: zacchei@oats.inaf.it

<sup>1</sup> *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.

RCAs at 30 GHz, three at 44 GHz, and six at 70 GHz. Each radiometer has two independent diodes for detection, integration, and conversion from radio frequency signal to DC voltage. The LFI is cryogenically cooled to 20 K to reduce noise, while the pseudo-correlation design with reference loads at  $\approx 4$  K ensures good suppression of  $1/f$  noise (Mennella et al. 2011).

LFI produces full-sky maps centered near 30, 44, and 70 GHz, with significant improvements with respect to current CMB data in the same frequency range. Careful data processing is required in order to realize the full potential of LFI and the ambitious science goals of *Planck*, which require that systematic effects be limited to a few  $\mu$ K per resolution element.

In this paper we describe the processing steps implemented to create LFI data products, with particular attention to the needs of the first set of astrophysics results.

The structure of the paper follows the flow of the data through the analysis pipeline. Section 2 describes the creation of time ordered information (TOI) from telemetry packets, time stamping, pointing reconstruction, and data flagging. Section 3 describes the main operations performed on the TOI, including removal of frequency spikes, creation of differenced data, determination of the gain modulation factor, and diode combination. Beam reconstruction is discussed in Sect. 4, calibration in Sect. 5, and noise in Sect. 6. Map-making, covariance matrices, and tests based on jackknife analysis and Monte Carlo simulations are described in Sect. 7. Section 8 reports on colour corrections. Section 9 describes how the CMB was removed from LFI and HFI maps. Finally, Sect. 10 gives an overview of the software infrastructure at the LFI DPC.

## 2. Creation of time ordered information

The task of the Level 1 DPC pipeline is to retrieve all necessary information from packets received each day from the Mission Operations Centre (MOC) and to transform the scientific TOI and housekeeping (H/K) data into a form that is manageable by the scientific pipeline.

During the  $\sim 3$  h daily telecommunication period (DTCP), the MOC receives telemetry from the previous day, archived in on-board mass memory, together with real-time telemetry. Additional auxiliary files, such as the attitude history file (AHF) of the satellite, are produced.

The MOC consolidates the data for each day, checking for gaps or corrupted telemetry packets, then provides the data, together with additional auxiliary data, to the DPCs through a client/server application called the data disposition system (DDS).

The data are received at the DPC as a stream of packets, which are handled automatically by four Level 1 pipelines: Data Receipt, Telemetry Handling, Auxiliary Data, and Command History.

The Data Receipt pipeline implements the client side of the interface with the DDS. It requests a subset of data provided through this interface. A finite-state machine model has been used in the design of this pipeline for better formalization of the actions required during interaction with the DDS server.

The Telemetry Handling pipeline is triggered when a new segment of telemetry data is received. The first task (Telemetry Unscrambler) discriminates between scientific and housekeeping telemetry packets. Scientific packets are grouped according to radiometer, detector source, and processing type, then uncompressed and decoded (see next paragraph). The on-board time of each sample is computed based on the packet on-board time and the detector sampling frequency. Housekeeping telemetry

packets are also grouped according to packet type, and each housekeeping parameter within the packet is extracted and saved into TOI. Subsequent tasks of the pipeline perform calibration of housekeeping and scientific TOIs together with additional quality checks (e.g., out of limits, time correlation). The last task, FITS2DMC, ingests the TOIs into the Data Management Component (DMC), making them available to the Level 2 and Level 3 pipelines.

The Auxiliary Data pipeline ingests the AHF provided by Flight Dynamics into the DMC. Finally the Command History pipeline requests and stores the list of telecommands sent to the satellite during the DTCP.

The four pipelines are implemented as perl scripts, scheduled every 5 min. Trigger files are created to activate the processing in the Auxiliary Data and Command History pipelines, and a pipeline monitoring facility displays information about the status of each pipeline. The entire Level 1 pipeline was heavily tested and validated before the start of *Planck* operations (see Frailis et al. 2009, for more details).

### 2.1. Scientific data processing

When creating TOI, the Level 1 pipeline must recover accurately the values of the original (averaged) sky and load samples acquired on-board. The instrument can acquire scientific data in several modes or “PTypes”; we describe here only the nominal one (PType 5) (see Zacchei et al. 2009). The key feature is that two independent differenced time streams are created from the sky and load signals with two different gain modulation factors (GMFs).

Data of PType 5 are first uncompressed. The lossless compression applied on-board is inverted, and the number of samples obtained is checked against auxiliary packet information. Decompressed data  $Q_{i=1,2}$  are then subject to a dequantization step to recover the original signals  $P_i$  according to

$$P_i = \frac{Q_i}{\text{SECOND\_QUANT}} - \text{OFFSET\_ADJUST}, \quad (1)$$

where SECOND\_QUANT and OFFSET\_ADJUST are parameters of the readout electronics box assembly (REBA), calibration of which is described by Maris et al. (2009).

After dequantization, data are demixed to obtain  $\bar{S}_{\text{sky}}$  and  $\bar{S}_{\text{load}}$  using as inputs the gain modulation factors  $R_1$  and  $R_2$  determined during REBA calibration (Maris et al. 2009):

$$\bar{S}_{\text{sky}} = \frac{R_2 P_1 - R_1 P_2}{R_2 - R_1}, \quad (2)$$

$$\bar{S}_{\text{load}} = \frac{P_1 - P_2}{R_2 - R_1}. \quad (3)$$

Conversion from ADU (analog-to-digital units) to volts is achieved by

$$\bar{V}_i = \frac{\bar{S}_i - Z_{\text{DAE}}}{G_{\text{DAE}}} - O_{\text{DAE}}, \quad (4)$$

where  $G_{\text{DAE}}$ ,  $O_{\text{DAE}}$ , and  $Z_{\text{DAE}}$  are data acquisition electronics (DAE) gain, offset, and small tunable offset, respectively, whose optimal values were determined during ground tests (Maris et al. 2009).

## 2.2. On-board time reconstruction

A time stamp is assigned to each data sample. If the phase switch (Mennella et al. 2011) is off (not switching), the packet contains consecutive values of either sky or load samples. Then

$$t_{i_{\text{smp}}}^{\text{obt}} = t_0^{\text{obt}} + i_{\text{smp}} \frac{N_{\text{aver}}}{f_{\text{samp}}}, \quad (5)$$

where  $i_{\text{smp}} \geq 0$  is the sample index within the packet and  $t_0^{\text{obt}}$  is the mean time stamp of the first averaged sample.  $N_{\text{aver}}$  is the number of fast samples averaged together to obtain a single detector sample, and  $f_{\text{samp}} \approx 4$  kHz is the detector sampling frequency.

If the phase switch is on (nominal case), consecutive pairs of either sky–load or load–sky samples are stored in the packet. Then consecutive pairs of samples have the same time stamp and

$$t_{i_{\text{smp}}}^{\text{obt}} = t_0^{\text{obt}} + 2 \text{trunc} \left[ \frac{i_{\text{smp}}}{2} \right] \frac{N_{\text{aver}}}{f_{\text{samp}}}. \quad (6)$$

On-board time information is stored in the form of TOI and directly linked to its scientific sample.

## 2.3. Data flagging

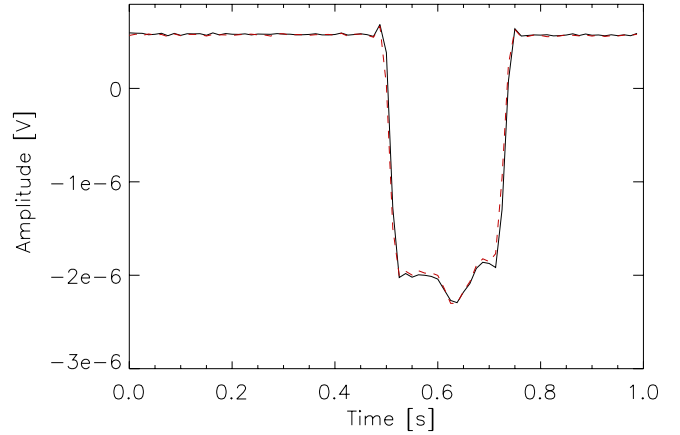
For each sample we define a 32-bit flag mask to identify potential inconsistencies in the data and to enable the pipeline to skip that sample or handle it differently during further processing. Currently flags that are checked include: those identifying the stable pointing period (determined from the AHF); science data that cannot be recovered (e.g., because of saturation); samples artificially created to fill data gaps; and samples affected by planet transits and moving objects within the Solar system.

## 3. TOI processing

### 3.1. Electronic spikes

The clock in the housekeeping electronics is inadequately shielded from the data lines, resulting in noise “spikes” in the frequency domain at multiples of 1 Hz (Meinhold et al. 2009). The spikes are synchronous with the on-board time, with no change in phase over the entire survey, allowing construction of a piecewise-continuous template by summing the data for a given detector onto a one second interval (Fig. 1) The amplitude and shape differ from detector to detector; differences between detectors of different frequency tend to be larger than between those of the same frequency. The amplitude also varies with time. This variation is estimated by constructing templates like Fig. 1 summed over the entire survey to obtain the shape of the signal, and then fitting the amplitude of a signal of this shape for each hour of data. This amplitude is smoothed with a 20-day boxcar window function to reduce the noise. Because of noise, this is likely to be an overestimate of the true variations.

To estimate the effect of spikes on the science data, we generate three simulated maps at each frequency. The first is a noise map, generated from the instrument white noise levels as measured in the data and the scan strategy of *Planck*, but no spikes or correlated noise. This is a best case scenario, with the lowest noise level possible in a real map. The second is a “spike” map, calculated assuming the square wave template for each detector modulated by a time-varying amplitude measured from the data, as described above. Because the variation of amplitudes is an



**Fig. 1.** A square wave template for both sky (black) and load (red/dashed line) for one of the 44 GHz detectors, computed by adding data between operational day (OD) 91 and 389 in phase over a 1-h interval. Individual templates are directly subtracted from the un-differenced data.

overestimate, as described above, this is a worst-case scenario of the effect of spikes. The third map is a “spike-subtracted” map, the same as the second, but with a constant spike template subtracted. This gives an estimate of the residual effect of electronic spikes that would be left in the maps if the spike template were subtracted. The 30 GHz maps are at HEALPix resolution  $N_{\text{side}} = 512$ ; the 44 GHz and 70 GHz maps were produced at  $N_{\text{side}} = 1024$ .

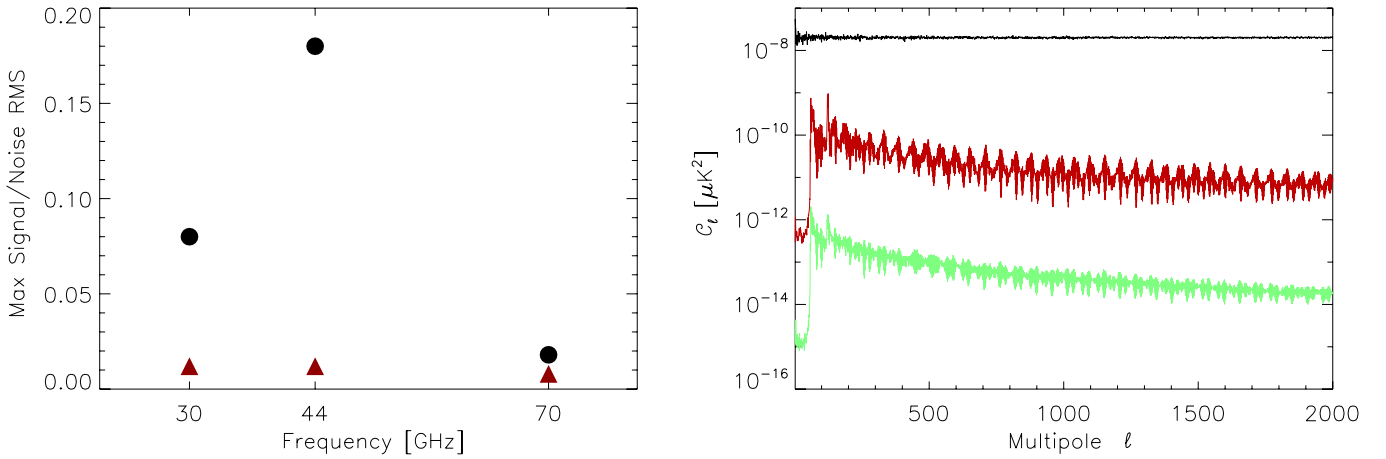
We scale the spike and spike subtracted maps to the noise, i.e.,  $\text{Map2}/\text{rms}(\text{Map3})$  and  $\text{Map2}/\text{rms}(\text{Map3})$ , where the rms is calculated from the global rms of the noise map scaled as appropriate for the relative number of observations (“hits”) in that pixel. Figure 2 (left) shows the maximum value of these ratios over the whole sky. At 44 GHz, the most affected frequency, the effect in the worst pixel is less than 20% of the noise. At 70 GHz the effect in the worst pixel is an insignificant 2% of the noise.

Figure 2 (right) gives angular power spectra of the three 44 GHz maps. The effect is everywhere well below the noise, and subtraction of a constant amplitude square-wave template reduces the effect by almost two orders of magnitude.

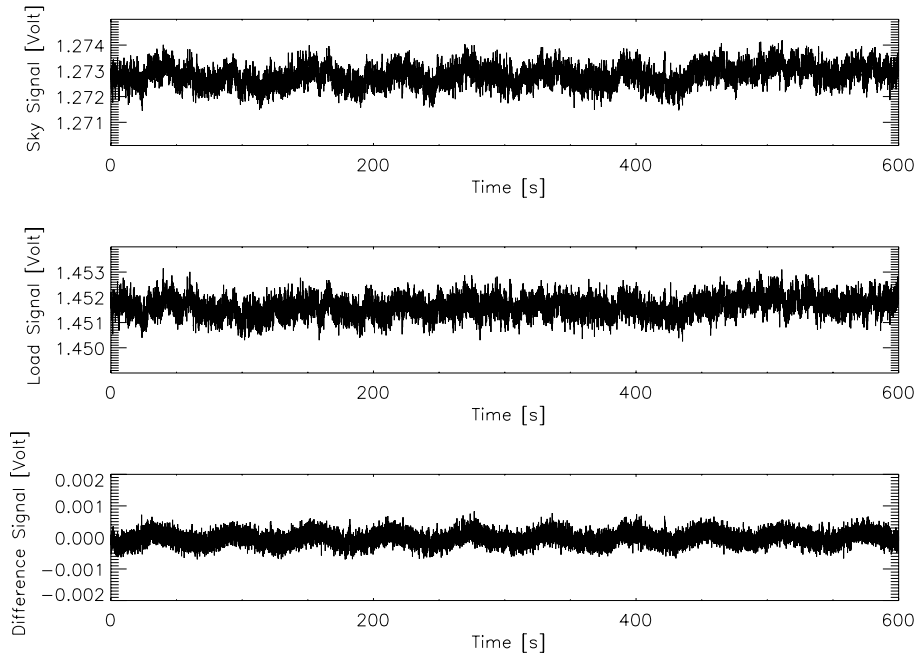
We decided to remove a square-wave template only at 44 GHz. This reduces the spike residual from 20% of the noise to 1% of the noise. At 70 GHz the effect of spikes is extremely small without correction, and at 30 GHz uncertainty in the template combined with the small size of the effect argued against removal.

### 3.2. Gain modulation factor and differenced data

The output of each detector (diode) switches at 4096 Hz (Mennella et al. 2010) between the sky,  $V_{\text{sky}}$ , and the 4 K reference load,  $V_{\text{load}}$ .  $V_{\text{sky}}$  and  $V_{\text{load}}$  are dominated by  $1/f$  noise, with knee frequencies of tens of hertz. This noise is highly correlated between the two streams, a result of the pseudo-correlation design (Bersanelli et al. 2010), and differencing the streams results in a dramatic reduction of the  $1/f$  noise. The two arms of the radiometer are slightly unbalanced, as one looks at the 2.7 K sky and the other looks at the  $\sim 4.5$  K reference load. To force the mean of the difference to zero, the load signal is multiplied by the GMF,  $R$ , which can be computed in several ways (Mennella et al. 2003). The simplest method, and the one implemented in



**Fig. 2.** The effect of electronic spikes on the data. *Left* – Maximum pixel value in the simulated “spike” (black dots) and “spike-subtracted” maps (red triangles; see text), scaled to the local pixel noise. In our data processing, the square wave signal is subtracted only at 44 GHz detectors. The black circles therefore represent the estimated highest spike signal level in the 30 GHz and 70 GHz maps, while the red triangle represents the estimated highest residual spike signal level in the 44 GHz map. *Right* – Angular power spectra of the 44 GHz simulations: the red (*middle*) line shows the power spectrum of the simulated spike map, and the green (*bottom*) line shows the power spectrum of the simulated spike-subtracted map. Subtraction reduces the power by a factor of about 100, from a small to an insignificant fraction of the white noise power, shown by the black (*top*) line.



**Fig. 3.** Effect of the gain modulation factor (GMF) on sky and load signals for flight data. *The upper and middle panels* show 10 min of sky and load signals of the LFI27S-11 detector: they are highly correlated with clear signatures of low-frequency noise. After application of the GMF in taking the difference (Eq. (7)), such fluctuations are dramatically reduced, revealing the presence of a sky signal dominated by the CMB dipole (*lower panel*). Note the change in the  $y$ -axis scale.

the processing pipeline, is to take the ratio of DC levels from sky and load outputs obtained by averaging the two time streams, i.e.,  $R = \langle V_{\text{sky}} \rangle / \langle V_{\text{load}} \rangle$ . Then

$$\Delta V(t) = V_{\text{sky}}(t) - \frac{\langle V_{\text{sky}} \rangle}{\langle V_{\text{load}} \rangle} V_{\text{load}}(t). \quad (7)$$

We compute  $R$  from unflagged data for each pointing period identified from the AHF information.

To verify the accuracy of this approach, we started with a time stream of real differenced data, then generated two time streams of undifferenced data using a constant (typical) value of  $R$ . We then ran these two time streams through the pipeline, and compared the results with the original time stream.

Deviations between the pipeline values of  $R$  and the constant input value used to generate the undifferenced data were at the 0.002% level.

The  $R$  factor has been stable over the mission so far, with overall variations of 0.03–0.04%. To keep the pipeline simple, we apply a single value of  $R$  to each pointing period.

Figure 3 shows the effect of applying Eq. (7) with the  $R$  factor to flight data. The correlated  $1/f$  noise in sky and load streams (evident in the two upper plots of the figure) is reduced dramatically. The residual  $1/f$  noise has a knee frequencies of 25 mHz, and little effect on maps of the sky, as described in Sect. 7.

### 3.3. The diode combination

Having two diodes for each radiometer enables observation of both sky and load with a combined duty cycle of almost 100%. In combining the outputs, however, we must take into account the effects of imperfect isolation and differences in noise between the two diodes.

Isolation between diodes was measured for each radiometer in ground tests and verified in flight using the CMB dipole, planets, and Galactic plane crossings. Typical values range from  $-13$  to  $-20$  dB. This is within specifications, and does not compromise LFI sensitivity. It does, however, produce a small anti-correlation of the white noise of the two diodes of a given radiometer. When data from the two diodes are averaged, the white noise of the resulting TOI is lower than would be the case if they were statistically independent. A complete mathematical description of this behaviour is given in [Mennella et al. \(2011\)](#). This causes no difficulty in subsequent calibration and further processing; however, the effect must be taken into account in inferring the noise properties of individual detector chains from the combined outputs.

To take account of differences in noise in combining the diode outputs, we assign relative weights to the uncalibrated diode time-streams based on their calibrated noise. Specifically, we make a first order calibration of the timelines,  $G_0$  and  $G_1$ , subtract a signal estimate, and calculate the calibrated white noise levels,  $\sigma_0$  and  $\sigma_1$ , for the two diodes. The weights for the two diodes ( $i = 0$  or  $1$ ) are

$$W_i = \frac{\sigma_i^2}{G_{01} \sigma_0^2 + \sigma_1^2} \quad (8)$$

where the weighted calibration constant  $G_{01}$  is given by

$$G_{01} = \frac{1}{\sigma_0^2 + \sigma_1^2} [G_0 \sigma_1^2 + G_1 \sigma_0^2], \quad (9)$$

and is the same for each diode.

The weights are fixed to a single value per radiometer for the entire dataset. Since all calibrations, noise estimation, and other tests are done on these combined data streams, small errors in the weights cause inconsequential losses in sensitivity, and no systematic errors.

### 3.4. Detector pointing

Detector pointing is a fundamental ingredient in data processing that requires knowledge of the spacecraft attitude and the location of the horns in the focal plane. The AHF gives the orientation of the spacecraft spin axis in quaternions sampled at 8 Hz, as well as beginning and ending times for a single pointing period. It specifies with appropriate flags the periods of spin-axis maneuvers during which star tracker positions are unreliable. Horn locations within the focal plane are determined from both ground measurements and planet crossings.

The orientation of the spacecraft spin axis at the time of each data sample is determined by linear spherical interpolation of the 8 Hz quaternions. Individual detector pointings are determined by simple rotations from the spin-axis reference frame to the telescope optical axis, then to the relevant horn position, with an additional rotation to account for the orientation of the horn in the focal plane.

In some cases small extrapolations of the quaternions are necessary at the end of a pointing period. Simulations verify that these introduce no significant degradation of the pointing accuracy.

## 4. Main beams and the geometrical calibration of the focal plane

Knowledge of the beams is of paramount importance in CMB experiments. Errors and uncertainties, and the details of complex non-Gaussian shapes, directly affect cosmological parameters.

We determine the main beam parameters and the position of each horn in the focal plane from planet observations. Jupiter gives the best results, but other planets and bright celestial sources have been used as well. Inputs to the calculations include TOI from each radiometer throughout the planet crossing, the AHF for the same period, and the time-dependent position of the planet as seen by *Planck*, provided by the JPL Horizons system, which accounts for both spacecraft and planet motion.

### 4.1. Algorithm and testing

We create a 2D map of the footprint of the focal plane on the sky by selecting data within  $10^\circ$  of the telescope line of sight. This comprises the whole extension of the LFI focal plane. To minimize the effects of  $1/f$  noise on weak sources, we use TOI from which offsets per ring derived by the Madam destriper (Sect. 7) have already been removed.

We fit a bivariate Gaussian beam model to these data ([Burigana et al. 2001](#)):

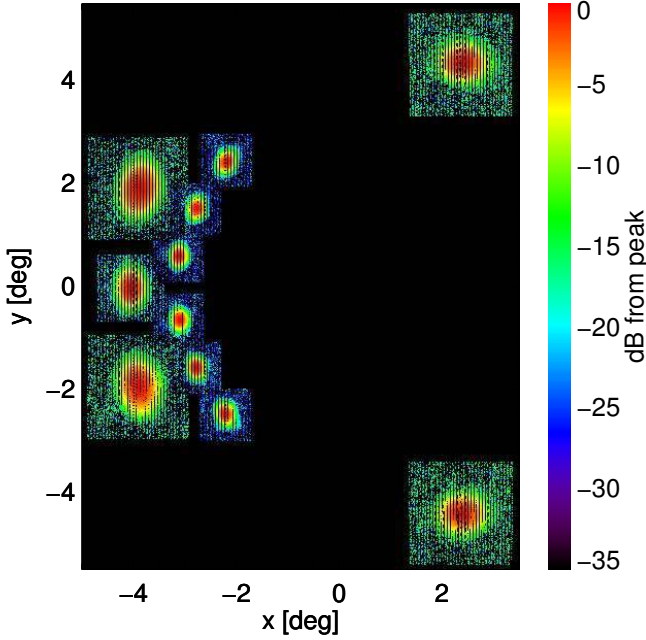
$$B(x_i, y_i) = \frac{A}{d^2} \exp \left\{ -\frac{1}{2} \left[ \frac{(\Delta x_i \cos \alpha + \Delta y_i \sin \alpha)^2}{\sigma_x^2} + \frac{(-\Delta x_i \sin \alpha + \Delta y_i \cos \alpha)^2}{\sigma_y^2} \right] \right\}. \quad (10)$$

Here  $A$  is an overall amplitude.  $x_i$  and  $y_i$  are Cartesian coordinates, with  $x_c$ ,  $y_c$  the position of the centre of the beam, and  $\Delta x_i \equiv x_i - x_c$  and  $\Delta y_i \equiv y_i - y_c$ .  $\sigma_x$  and  $\sigma_y$  are the beamwidth parameters of the elliptical approximation of the beam shape, and the angle  $\alpha$  is the reconstructed orientation of the beam in the focal plane and  $d$  is the actual distance (in astronomical units) of the planet.

We tested our technique with simulations using the measured beam patterns together with a detailed model of the *Planck* telescope. The simulations included the nominal main and far beam patterns, the effects of smearing caused by the motion of the satellite, and pointing uncertainties. Using these simulations of Jupiter crossings (including instrumental noise and complete sky signal), we are able to reconstruct the main beam shape down to  $-20$  dB and to recover the main beam properties at the 1% level or better for all LFI beams. Table 1 reports results for the main beam properties for a sample of the LFI beams. These figures are representative of our expected accuracy for in-flight beam and focal plane reconstruction.

Figure 4 shows the footprint of the LFI focal plane obtained during the first season of Jupiter observations, from 24 October to 1 November 2009. Figure 5 shows beam images for LFI28M, LFI25M and LFI21M from those observations. As expected, all beams are asymmetric but with no significant departures from an elliptical shape visible down to the  $\sim -10$  dB level. For lower levels, aberration starts to distort the beam response, creating non-elliptical shapes.

We also constructed a planet mask, including Jupiter, Mars, and Saturn, the most luminous planets at LFI frequencies. The planet mask is radiometer-dependent, since each horn observes a planet at different times. The planet masking algorithm assigns an appropriate flag to data that lie within an ellipse, centred at the position of the planet and with an orientation that matches



**Fig. 4.** LFI focal plane as determined from the first season of Jupiter observations, 24 October to 1 November 2009. Contour levels are in dB from the peak. All beams are well approximated by an elliptical Gaussian down to the  $-10$  dB level.

**Table 1.** Simulation of the reconstruction of the beams and focal plane geometry from observations of Jupiter, including realistic models of the beams, instrument noise, beam smearing, and star tracker uncertainties.

Parameter	Input	Reconstruction	$\delta$ [%]
<b>LFI19S – 70 GHz</b>			
FWHM	12:83	12:97	1.12
ellipticity	1.280	1.276	0.98
$x_0$	$-2^{\circ}8715$	$-2^{\circ}8704$	0.36
$y_0$	$-1^{\circ}5678$	$-1^{\circ}5829$	0.96
<b>LFI25S – 44 GHz</b>			
FWHM	29:33	30:23	3.07
ellipticity	1.170	1.230	5.12
$x_0$	$-2^{\circ}8227$	$-2^{\circ}8293$	0.23
$y_0$	$-5^{\circ}1369$	$-5^{\circ}0844$	1.02
<b>LFI27M – 30 GHz</b>			
FWHM	32:42	32:89	1.45
ellipticity	1.380	1.384	0.32
$x_0$	$-4^{\circ}7788$	$-4^{\circ}7798$	0.02
$y_0$	$2^{\circ}4903$	$2^{\circ}3958$	3.79

the beam orientation, with axes  $\sim 3$  times larger than the beam widths derived from beam fitting. These flags are used in the map-making and ensuing data analysis to discard samples affected by planet transits

## 5. Photometric calibration

### 5.1. First steps

The ideal source for photometric calibration, i.e., conversion of the data from volts to kelvin, should be constant, perfectly known, present during all observations, and have the same

frequency spectrum as the CMB. In the frequency range of the LFI, the CMB dipole, caused by the motion of the Solar system with respect to the CMB reference frame, satisfies nearly all of these requirements, lacking only in that it is well, but not perfectly, known. The modulation induced on the CMB dipole by the orbital motion of *Planck* around the Sun satisfies even this last requirement, and will be the ultimate calibration source for the LFI; however, it cannot be used effectively until data for a full orbit of the Sun are available. For this paper, therefore, we must use the CMB dipole. We follow essentially the calibration procedure used for the WMAP first year data (Hinshaw et al. 2003). For the  $k$ th pointing period, the signal from each detector can be written as

$$\Delta V_k = g_k(\Delta T_{\text{sky}} + n) + b_k, \quad (11)$$

where  $\Delta T_{\text{sky}}$  is the sky signal,  $n$  is the noise, and  $g_k$  and  $b_k$  are the gain and baseline solution. The dominant sky signal on short time scales is the CMB dipole (Galactic plane crossings produce a localized spike that is easy to exclude). This is modeled as

$$\Delta V_m(g_k, b_k) = g_k(\Delta T_d + \Delta T_v) + b_k, \quad (12)$$

where we have considered both the cosmological dipole  $\Delta T_d$  and the modulation from the spacecraft motion  $\Delta T_v$ . We fit for  $g_k$  and  $b_k$  for each pointing period  $k$  by minimising

$$\chi^2 = \sum_{i \in k} \frac{[\Delta V(t_i) - \Delta V_m(t_i|g_k, b_k)]^2}{\text{rms}_i^2}. \quad (13)$$

The sum includes unflagged samples within a given pointing period  $k$  that lie outside a Galactic mask.

The mask is created from simulations of microwave emission provided by the Planck Sky Model (PSM)<sup>2</sup>. Of the LFI frequencies, 30 GHz has the strongest diffuse foreground emission. The mask excludes all pixels that in the 30 GHz PSM are more than  $5 \times 10^{-4}$  times the expected rms of the CMB. It also excludes point sources brighter than 1 Jy found in a compilation of all radio catalogues available at high frequencies (the *Planck* Input Catalogue, see Massardi 2006). The Galactic and point source masks preserve  $\sim 82\%$  of the sky.

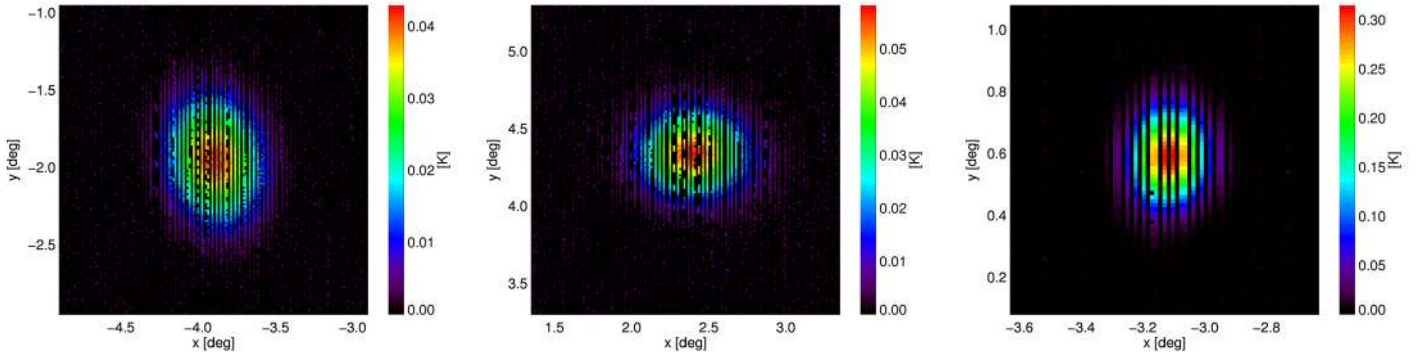
The *Planck* scan strategy is such that the instrument field of view describes nearly great circles on the sky. The signal mean is therefore almost zero and nearly constant from one circle to the next. This reduces the correlation between the gain and baseline solutions, a feature also taken advantage of by WMAP (Hinshaw et al. 2003).

As pointed out by Hinshaw et al. (2003) and Cappellini et al. (2003), the largest source of error in Eq. (13) arises from unmodelled sky signal  $\Delta T_a$  from CMB anisotropy and emission from the Galaxy. To correct this, we solve iteratively for both  $g'_k$  and  $\Delta T'_a$ . If  $g'_k$  is the solution at a certain iteration, the next solution is derived using Eq. (13) with

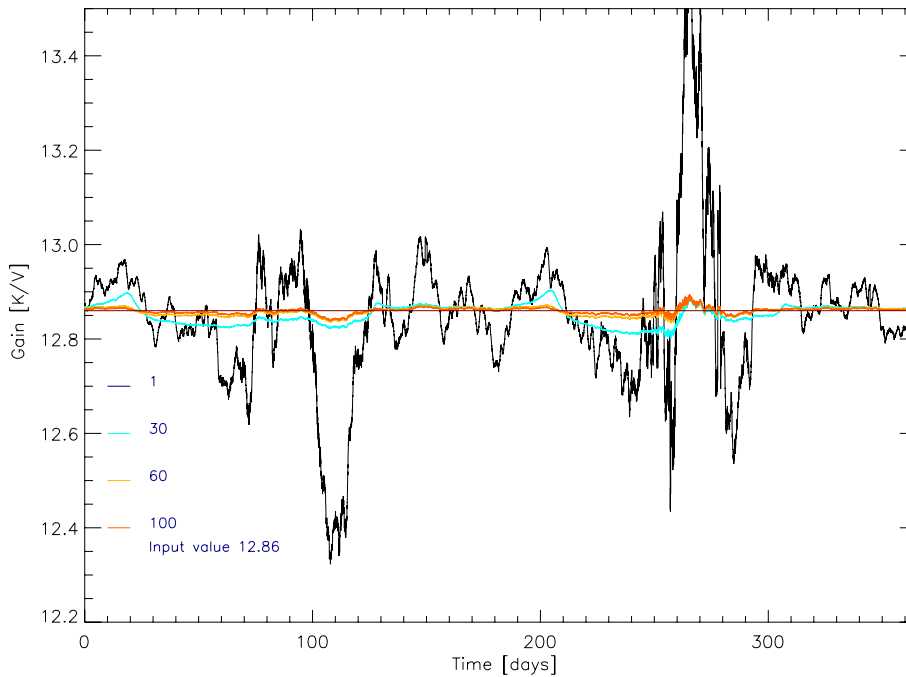
$$\Delta V' = \Delta V - g'_k \Delta T'_a, \quad (14)$$

where  $\Delta T'_a$  is the sky signal (minus dipole components) estimated from a sky map built from the previous iteration step. This is repeated to convergence, typically after a few tens of iterations. Figure 6 shows the gain error induced by unmodelled sky signal in a one-year simulation of one 30 GHz detector. The simulation includes CMB anisotropies, the CMB dipole,

<sup>2</sup> The *Planck* Sky Model is available at: [http://www.apc.univ-paris7.fr/APC\\_CS/Recherche/Adamis/PSM/psky-en.html](http://www.apc.univ-paris7.fr/APC_CS/Recherche/Adamis/PSM/psky-en.html)



**Fig. 5.** Beam map of LFI data around the Jupiter observations (24 October–1 November 2009) for LFI horns LFI28M (*left*), LFI25M (*middle*) and LFI21M (*right*).



**Fig. 6.** Simulation showing convergence of the gain solution for one year of observations of one 30 GHz detector. The simulations include CMB dipole(s), CMB anisotropies, and Galactic emission. The input gain was 12.86 K/V and the sky included all diffuse components as well as nominal instrument noise. The first iteration shows large errors caused by Galactic and CMB anisotropy emissions; however, after one hundred iterations, convergence is achieved with an overall deviation from the input value of less than 0.01%. The various curves show the solution after 1, 30, 60 and 100 iterations.

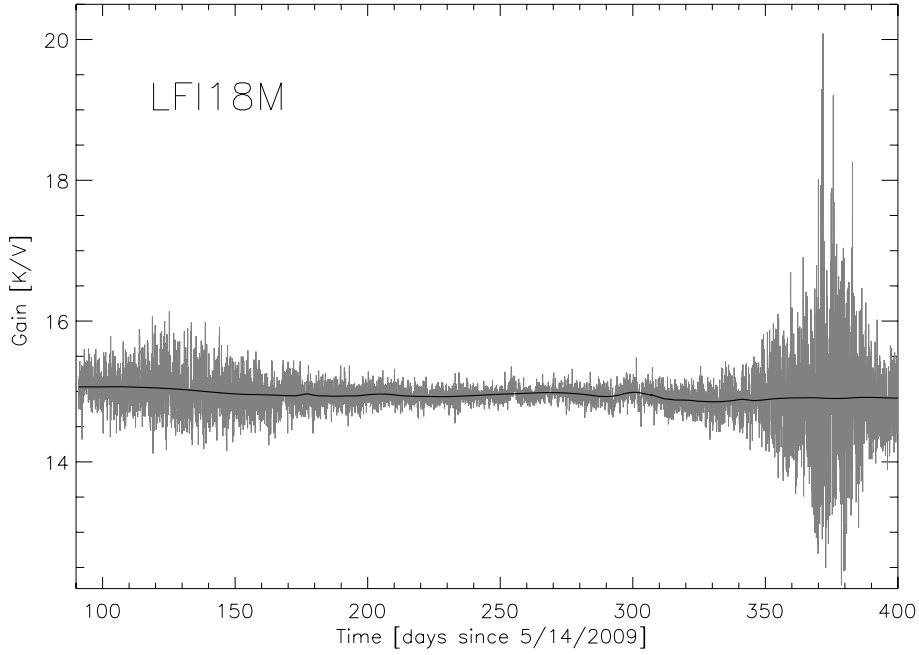
and Galactic emission. Gain errors in this example are  $\sim 5\%$  after one iteration. After a few tens of iterations, the residual errors are  $< 0.01\%$  over the entire year.

The algorithm alternates between dipole fitting and map-making. Maps are made with (Madam Sect. 7) ignoring polarisation, with no noise prior and baseline length equal to the pointing period length. To improve calibration and reduce noise, calibration is performed simultaneously for both radiometers of each single horn. In the presence of real noise, the actual scatter from one gain solution to the other is quite large. Figure 7 shows an example of the hourly gain solution (grey line) derived from the iterative scheme described above for LFI18M, one of the 70 GHz radiometers. Apart from the scatter induced by instrument noise, the gain solution is quite stable throughout the observation period. Around the dipole maxima, typical noise-induced variations are  $\sim 0.8\%$  (rms). Nonetheless, the stability of the gain solution is poor compared to the stability of the instrument itself, as indicated by the stability of the uncalibrated white noise level of both differenced and undifferenced data, or the stability of the

total power from both sky and load signals. This is particularly evident during the minima of the dipole signal (see Mennella et al. 2011, for further details).

There are also specific things that affect the gain solution. To the extent that they can be measured and understood, their effects on the gain can be corrected directly. For example, a non-linearity in the analogue-to-digital converters (ADCs), discovered during data analysis, produces a multiplicative effect on the data that is recovered (erroneously) by the calibration pipeline as a gain variation. We have developed two independent, complementary methods to correct for this. In the first, we calibrate the data using the gain solution that follows the induced ADC gain variation. In the second, we model the nonlinearity and remove the effect at the raw TOI level.

Alternatively, temperature variations of the amplifiers can induce real gain variations on short time scales. For example, during the first 259 days after launch the downlink transponder was powered up only for downlinks. This induced rather sharp daily variations in the temperature and gain of the amplifiers in the



**Fig. 7.** Hourly gain solution (gray line) from flight data for LFI18M, as derived from our iterative calibration algorithm. The gain is quite stable over the observing period, although there is a lot of scatter due to noise, especially during dipole minima. The thick black line is the refined gain solution (see text) applied to create calibrated TOI and sky maps.

back-end unit (BEM). Starting on day 259, the transponder has been powered up continuously, eliminating this source of gain variations.

In the next section we discuss additional steps taken in the calibration procedure to deal with the effects of noise and gain changes induced by events such as the transponder cycle change.

### 5.2. Improving calibration accuracy

As shown in Fig. 7, the hourly gain solutions are strongly affected by noise. To reduce the effects of noise and recover more accurately the true and quite stable gains of the instrument, we process the hourly gain solution as follows:

- calculate running averages of length 5 and 30 days. The 5-day averages are still noisy during dipole minima, while the 30-day averages do not follow real but rapid gain changes accurately.
- further smooth the 5- and 30-day curves with wavelets;
- use the 30-day wavelet-smoothed curve during dipole minima;
- use the 5-day unsmoothed curve around day 259 (the down-link transponder change) to trace real gain variations;
- use the 5-day wavelet smoothed curve elsewhere.

A typical gain solution is plotted in Fig. 7 as the solid black line. From the 5- and 30-day gain curves we infer information on the actual gain stability of the instrument as the mission progresses, and also on the overall uncertainty in the gain reconstruction. Specifically, the rms of the  $g_k$  over a period of  $N$  pointings is

$$\delta g = \sqrt{\frac{\sum_{k=1}^N (g_k - \langle g \rangle)^2}{N - 1}}, \quad (15)$$

where  $\langle g \rangle$  is the average of the  $N$  gains. The effect of the wavelet smoothing filter is to average over a number of consecutive pointings. Ignoring the different weights in the average, the

**Table 2.** Summary of dipole-based gain statistics.

DETECTOR	MAIN ARM		SIDE ARM	
	$\langle g_k \rangle$ [K/V]	$\sigma_{g_k}$ [%]	$\langle g_k \rangle$ [K/V]	$\sigma_{g_k}$ [%]
<b>70 GHz</b>				
LFI 18	14.935	0.279	22.932	0.243
LFI 19	27.434	0.141	41.843	0.228
LFI 20	25.572	0.253	29.581	0.261
LFI 21	41.629	0.367	41.999	1.038
LFI 22	64.275	0.367	62.504	0.185
LFI 23	36.492	0.290	54.121	0.382
<b>44 GHz</b>				
LFI 24	282.295	0.349	175.728	0.306
LFI 25	123.141	0.358	123.958	0.279
LFI 26	167.364	0.398	142.061	0.411
<b>30 GHz</b>				
LFI 27	12.875	0.314	15.320	0.349
LFI 28	15.802	0.225	19.225	0.379

overall uncertainty can be approximated as

$$\delta g|_{\text{stat}} \simeq \frac{\delta g}{\sqrt{M}} = \frac{1}{\sqrt{M}} \sqrt{\frac{\sum_{k=1}^N (g_k - \langle g \rangle)^2}{N - 1}}. \quad (16)$$

Table 2 lists the largest statistical uncertainties and their associated mean gains out of four time windows (days 100–140, 280–320, 205–245, 349–389, the first two corresponding to minimum and the second two to maximum dipole response), for the main and side arms of the LFI radiometers. In order to provide conservative estimates, we have always chosen a value for  $M$  corresponding to the number of pointings in 5 days, even in cases where a 30-day smoothing window was used. Equations (15) and (16) and Table 2 are the same as Eqs. (12) and (13) and Table 9 of Mennella et al. (2011). Peak-to-peak variations in the daily gains



reach 10% (with mean 7%); however, the rms of the smoothed gain solution is generally in the  $\sim 0.3\text{--}0.4\%$  range. This can be taken as the current level of LFI calibration accuracy.

Although the current pipeline provides results approaching those expected from the stability of the instrument, we are working to improve it as much as possible. In particular, we would like to trace gain variations on time scales shorter than the pointing period. To achieve this, we are developing a detailed gain model (currently under test) based on calibration constants estimated from the pipeline and instrument parameters (temperature sensors, total power data), see [Mennella et al. \(2011\)](#) for further information.

## 6. Noise estimation

Once data are calibrated, we evaluate the noise properties of each radiometer. We select data in chunks of 5 days each and then compute noise properties. This is done using the roma Iterative Generalized Least Square (IGLS) map-making algorithm ([Natoli et al. 2001](#); [de Gasperis et al. 2005](#)) which includes a noise estimation tool based on the iterative approach described in [Prunet et al. \(2001\)](#). IGLS map-making is time and resource intensive and cannot be run over the whole data set within the current DPC system. However since the TOD length considered here is only 5 ODs, it is possible to use the roma implementation of this algorithm which has a noise estimator built-in. The method implemented here is summarized as follows. Model the calibrated TOD as

$$\Delta T = \mathbf{P}\mathbf{m} + \mathbf{n}, \quad (17)$$

where  $\mathbf{n}$  is the noise vector, and  $\mathbf{P}$  is a projection matrix that relates a map pixel  $\mathbf{m}$  to a TOD measurement  $\Delta T$ . We obtain a zeroth order estimate of the signal through a rebinned map and then iterate noise and signal estimation:

$$\hat{\mathbf{n}}_i = \Delta T - \mathbf{P}\hat{\mathbf{m}}_i, \quad (18)$$

$$\hat{\mathbf{m}}_{i+1} = (\mathbf{P}^T \hat{\mathbf{N}}_i^{-1} \mathbf{P})^{-1} \mathbf{P}^T \hat{\mathbf{N}}_i^{-1} \Delta T, \quad (19)$$

where  $\hat{\mathbf{N}}_i$  is the noise covariance matrix in time domain estimated at iteration  $i$ . We have verified that convergence is reached in a few, usually three, iterations.

We calculate the Fourier transform of the noise time stream (with an FFT algorithm) and fit the resulting spectrum for the three parameters, the white noise level, the knee-frequency, and the slope of the  $1/f$  noise part:

$$P(f) = \sigma_{\text{WN}}^2 \left[ 1 + \left( \frac{f}{f_k} \right)^\beta \right]. \quad (20)$$

The white noise level is taken as the average of the last few percent of frequency bins. A linear fit to the log-log spectrum low frequency tail gives the slope of the  $1/f$  noise. The knee frequency,  $f_k$ , is the frequency at which these two straight lines intersect. We tested the accuracy with simulations that included sky signal and instrumental noise with known properties. The noise properties were recovered with typical deviations from input values of  $\sim 10\%$  for knee-frequency and slope, and less for white noise level. Examples of noise spectra and corresponding fits are shown in Fig. 8.

## 6.1. Noise constrained realizations and gap filling

The FFT-based noise power spectrum estimation method requires continuity of the noise time stream. As discussed in Sect. 3, we identify bad data (e.g., unstable spacecraft pointing, data saturation effects) and gaps in the data with appropriate flags. We fill in the flagged data with a Gaussian noise realization constrained by data outside the gap ([Hoffman & Ribak 1991](#)). Although in principle this method requires a pure noise time stream outside the gap, we have verified that given the low signal-to-noise ratio in the LFI TOD the procedure is not affected by the signal present in the time streams. We fill the gap with Gaussian noise whose properties match those of the noise power spectrum computed over the day immediately before the one with flagged data. An example is shown in Fig. 9.

## 7. The map-making pipeline

### 7.1. Frequency maps

The map-making pipeline produces sky maps of temperature and polarisation for each frequency channel. It takes as input the calibrated timelines and pointing information in the form of three angles ( $\theta, \phi, \psi$ ) describing the orientation of the feed horns for each data sample. An essential part of the map-making process is the reduction of correlated  $1/f$  noise, a large part which can be removed by exploiting redundancies in the scanning strategy. While the underlying sky signal remains the same, the observed signal varies due to noise. Statistical analysis of the signal variations allows one to distinguish between true sky signals and noise.

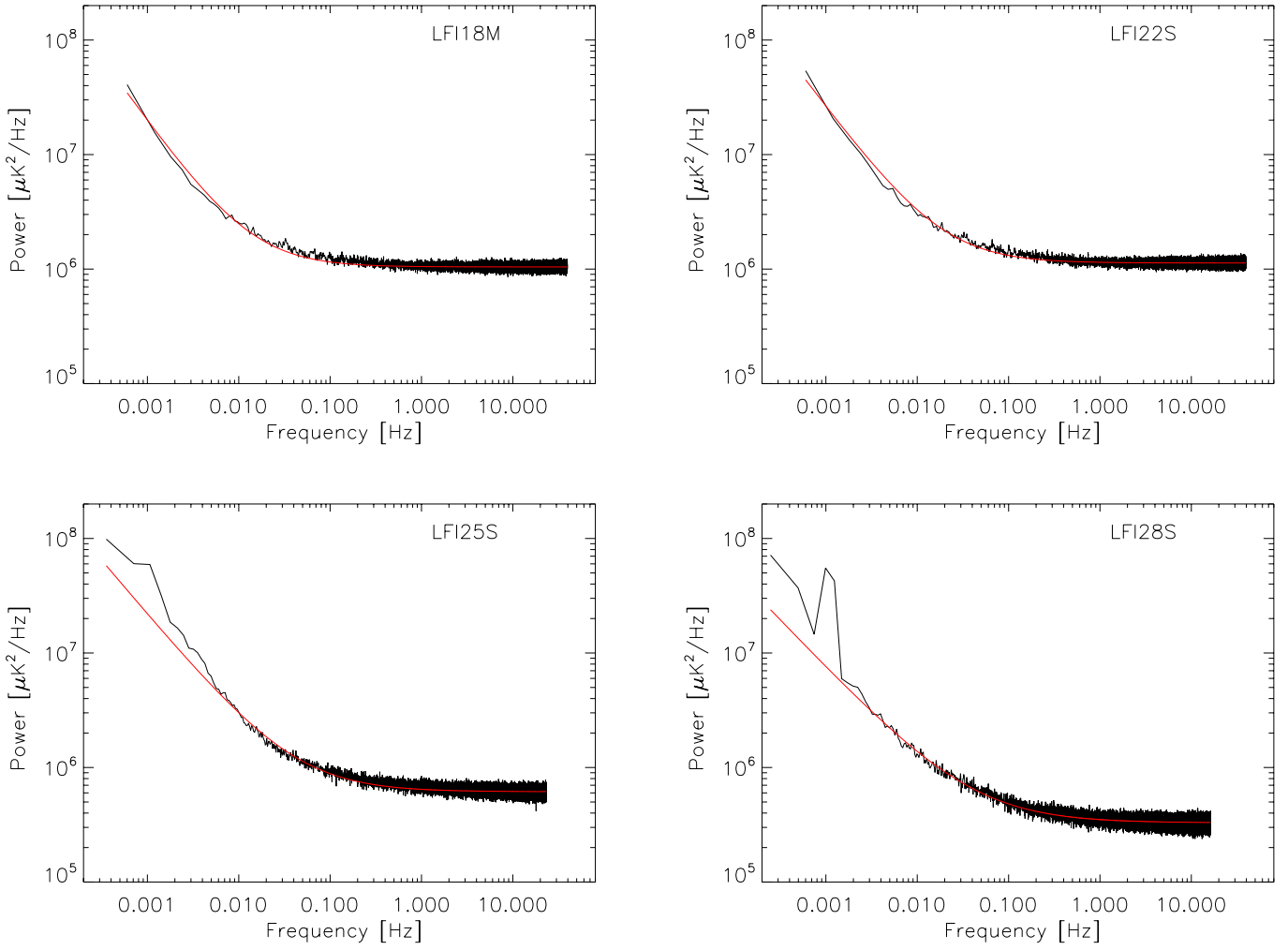
Among several map-making codes tested with simulated *Planck* data (see [Ashdown et al. 2007a,b, 2009](#)) the LFI baseline ([Mandolesi et al. 2010](#)) is to use the Madam destriping code ([Maino et al. 2002](#)). The algorithm and the underlying theory are described in detail in [Keihänen et al. \(2010\)](#); [Kurki-Suonio et al. \(2009\)](#); [Keihänen et al. \(2005\)](#). The basic idea is to model the correlated noise component by a sequence of constant offsets, called baselines. A key parameter in the code is the length of the baseline to be fitted to the data. Madam allows the use of an optional noise prior, if the noise spectrum can be reliably estimated, which further improves the accuracy of the output map. Without the noise prior, the optimal baseline length is of the order of the satellite spin period ( $\approx 1$  min). With an accurate noise prior, a much shorter baseline can be used. The shorter the baseline, the closer the Madam solution will be to the optimal Generalized Least Square solution (see Fig. 16 of [Ashdown et al. 2009](#)).

We are continually improving our knowledge of the instrument and its noise characteristics, and this information will eventually be used in the Madam algorithm. However, at this stage in the processing we decided to make two simplifications when running our map-making pipeline: no noise prior was used, and all radiometers were weighted equally. These choices lead to a simpler and faster map-making algorithm, which is sufficiently accurate for the *Planck* early results and avoids using detailed parameters describing the instrument which are under continual revision.

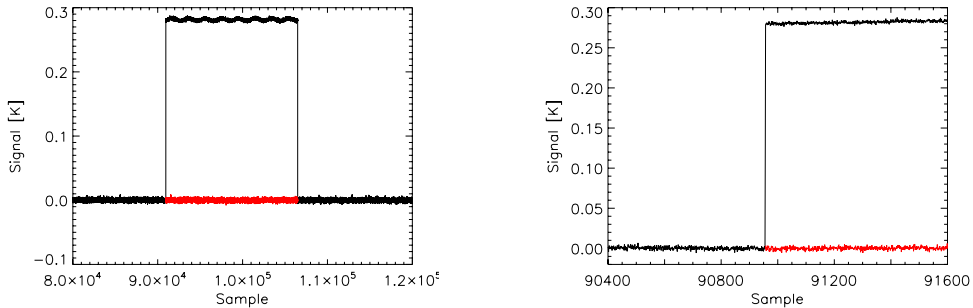
With these simplifications, the map-making equations can be written in a concise form. Technically, we are neglecting the baseline covariance,  $\mathbf{C}_a$ , and setting the white noise variance  $\mathbf{C}_n$  to unity. The basic model behind the algorithm is

$$\Delta T = \mathbf{P}\mathbf{m} + \mathbf{n}', \quad (21)$$

where  $\Delta T$  is the calibrated TOD,  $\mathbf{P}$  is the pointing matrix,  $\mathbf{m}$  is the pixelized sky map, and  $\mathbf{n}'$  is the instrumental noise. This last



**Fig. 8.** Noise spectra of radiometers LFI18M, LFI22S, LFI25S, and LFI28S) estimated by the noise pipeline (black lines). All spectra are well-fit by Eq. (20) with a single knee frequency and slope (red lines). An excess near 1 mHz is visible in LFI25S and LFI28S. This is approximately the bed-switching frequency of the sorption cooler, and the different slopes in LFI28S and LFI25S on the low-frequency side of the spectrum are possibly indications of thermal effects on the radiometer output.



**Fig. 9.** Gap filling procedure applied to LFI28M for day 239. *The upper panel* shows the original TOI (black) where a step is caused by a DAE gain change that produces saturated data. The (red) lines show the constrained noise realization used to replace those data. *The lower panel* shows a zoom around the position of the step to highlight the consistency of the gap filling data with the unflagged part of the TOI.

term can be written as

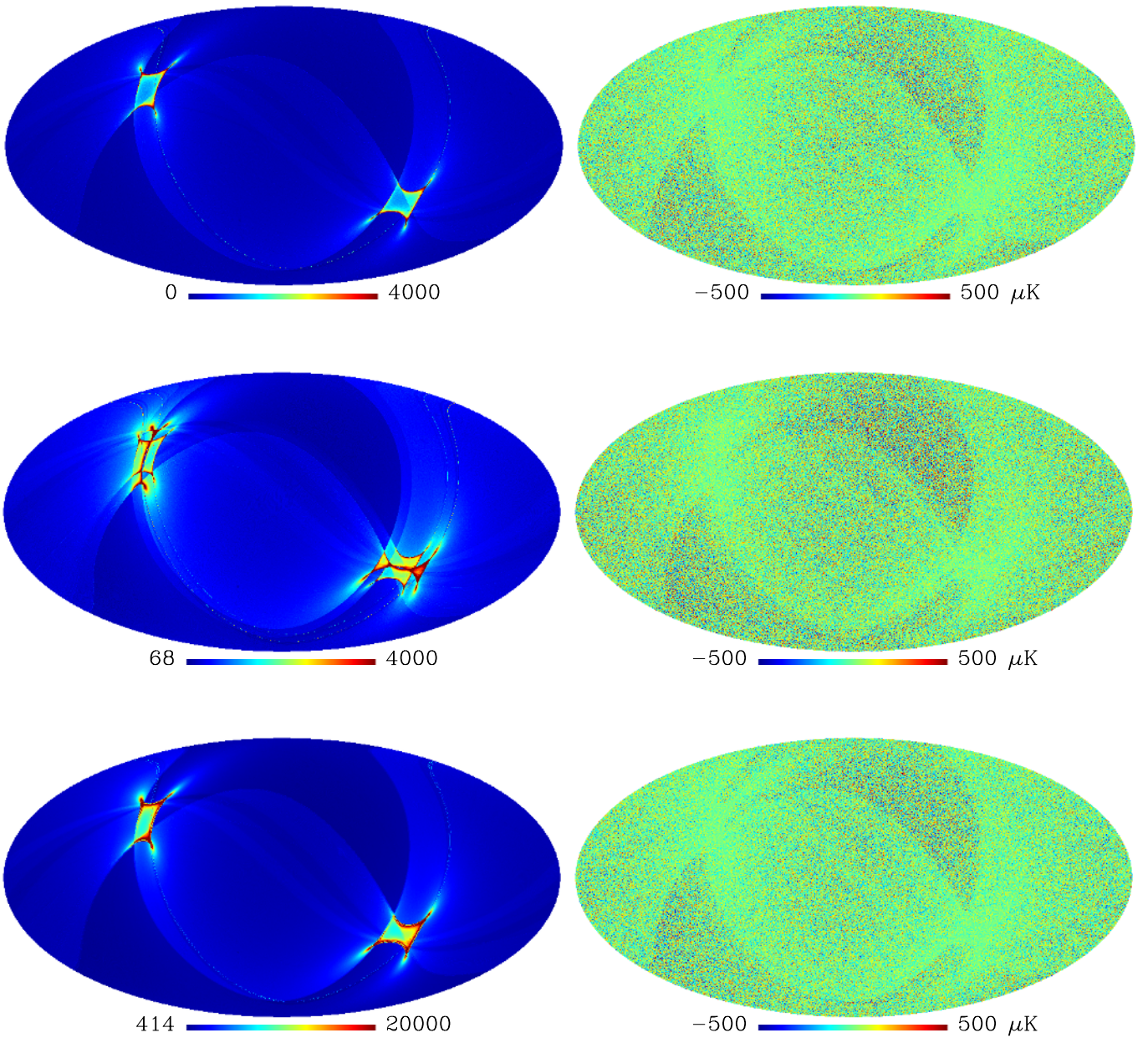
$$\mathbf{n}' = \mathbf{F}\mathbf{b} + \mathbf{n}, \quad (22)$$

where  $\mathbf{b}$  is the vector of unknown base function amplitudes and the matrix  $\mathbf{F}$  projects these amplitudes into the TOD. Since Madam uses uniform baselines, the matrix  $\mathbf{F}$  consists of ones and

zeros, indicating which TOD sample belongs to which baseline. Finally  $\mathbf{n}$  is a pure white noise stream assumed to be statistically independent of the baselines.

The maximum likelihood solution is obtained by minimizing

$$\chi^2 = (\Delta T - \mathbf{F}\mathbf{b} - \mathbf{P}\mathbf{m})^T (\Delta T - \mathbf{F}\mathbf{b} - \mathbf{P}\mathbf{m}), \quad (23)$$



**Fig. 10.** Hit count (*left*) and noise maps (*right*) at 30 GHz (*top*), 44 GHz (*middle*), and 70 GHz (*bottom*). The complex distribution around the ecliptic poles in the 44 GHz hit map is caused by the location of 44 GHz horns on the focal plane. The noise maps are derived from half-ring jackknife tests described in the text.

with respect to the quantities  $\mathbf{b}$  and  $\mathbf{m}$ . The baseline amplitudes  $\mathbf{b}$  are determined by solving

$$(\mathbf{F}^T \mathbf{Z} \mathbf{F}) \mathbf{b} = \mathbf{F}^T \mathbf{Z} \mathbf{y}, \quad (24)$$

where

$$\mathbf{Z} \equiv \mathbf{I} - \mathbf{P}(\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T. \quad (25)$$

Madam uses an iterative conjugate-gradient method to solve Eq. (24). An estimate for the map is finally obtained as

$$\mathbf{m} = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T (\Delta T - \mathbf{F} \mathbf{b}). \quad (26)$$

The map  $\mathbf{m}$  has as many elements as pixels in the sky. Each element is a Stokes parameter triplet  $(I, Q, U)$  for a pixel  $p$ . The

matrix  $\mathbf{P}^T \mathbf{P}$  is a  $3 \times 3$  block diagonal matrix that operates on map space. There is a block for each pixel  $p$ . A block can only be inverted if the pixel  $p$  is sampled with a sufficient number of different polarisation directions to allow determination of the three Stokes parameters for that pixel. This is gauged by the condition number of the block. For the present analysis, if the inverse condition number  $rcond$  (ratio of the smallest to largest eigenvalue) was less than 0.01, the pixel  $p$  was excluded from the  $(I, Q, U)$  map.

The  $\mathbf{P}^T \mathbf{P}$  blocks must be inverted when Eqs. (24) and (25) are solved for the baselines. These inversions are computed by eigenvalue decomposition. Eigenvalues whose magnitudes are less than  $10^{-6}$  times the largest eigenvalue are discarded; only the remaining part is inverted.

For the present analysis, we need only the  $I$ -component maps at the three LFI nominal frequencies, combining observations of all radiometers at a given frequency. Figure 10 shows (left column) the hit count maps by frequency. In addition, we produce maps from horn pairs scanning the same path in the sky (see [Mennella et al. 2011](#), for details on the LFI focal plane arrangement). We have produced 30 GHz maps at HEALPix resolution  $N_{\text{side}} = 512$ , and 44 and 70 GHz maps at  $N_{\text{side}} = 1024$ . All maps are in the NESTED scheme, in Galactic coordinates, with units of thermodynamic kelvin. The baseline length in Madam was one minute<sup>3</sup>.

## 7.2. White noise covariance matrices

If we bin the pure white noise stream  $\mathbf{n}$  to a map using the pointing  $\mathbf{P}$ , we obtain a binned white noise map,

$$\mathbf{w} = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{n}. \quad (27)$$

This map is a theoretical concept because we do not have access to the radiometer white noise streams. Its covariance matrix, however, is important because it provides an estimate of both the white noise power in each pixel and white noise correlations between Stokes parameters at a given pixel.

This *white noise covariance matrix* (WNC) is computed as (Eq. (27))

$$\mathbf{C}_w = \langle \mathbf{w} \mathbf{w}^T \rangle = (\mathbf{P}^T \mathbf{P})^{-1} (\mathbf{P}^T \mathbf{C}_n \mathbf{P}) (\mathbf{P}^T \mathbf{P})^{-1}. \quad (28)$$

Here  $\mathbf{C}_n \equiv \langle \mathbf{n} \mathbf{n}^T \rangle$ ,  $\mathbf{C}_n$  is a matrix that operates in the TOD domain, and angle brackets denote the ensemble mean. Because the radiometers have independent white noise,  $\mathbf{C}_n$  is diagonal. We assume that each radiometer has a uniform white noise variance  $\sigma_{\text{WN}}^2$  (see Sect. 6), but that each radiometer has its own variance. The radiometer  $\sigma_{\text{WN}}$  values that we used in the WNC computation are reported in [Mennella et al. \(2011\)](#).

## 7.3. Half-ring jackknife noise maps

For noise estimation purposes we divided the time ordered data into two halves and produced jackknife maps as follows. Each pointing period lasts typically  $\approx 44$  min (median 43.5 min, standard deviation 10 min). Typically, during the first 4 min the pointing is unstable, so these data are not used for science. During the remaining stable 40 min, each horn scans a ring on the sky. This ring consists of scan circles. One full scan circle takes 1 min. Therefore, each ring has about 40 scan circles. We made half-ring jackknife maps  $\mathbf{j}_1$  (and  $\mathbf{j}_2$ ) with the same pipeline as described in Sect. 7.1, but using stable data only from the first or the second half of each pointing period. Specifically, this is implemented by marking the other half of each ring as a gap in the data. Madam knows that for any given pointing period the first-used scan circle/sample of any half-ring is far apart in time (typically 25 min) from the last used scan circle/sample of the previous half-ring.

At each pixel  $p$ , the jackknife maps  $\mathbf{j}_1$  and  $\mathbf{j}_2$  contain the same sky signal (as long no time-varying sources or moving objects cross  $p$  at the time of observation), since they result from the same scanning pattern on the sky. However, because of instrumental noise, the maps  $\mathbf{j}_1$  and  $\mathbf{j}_2$  are not identical.

We can estimate the sky signal+noise as

$$\mathbf{m}_{1+2}(p) = [\mathbf{j}_1(p) + \mathbf{j}_2(p)]/2, \quad (29)$$

<sup>3</sup> One minute baselines for 30 GHz, 44 GHz, and 70 GHz are 1950, 2792, and 4726 samples respectively.

and the noise in map  $\mathbf{m}_{1+2}$  as

$$\mathbf{n}_{1+2}(p) = [\mathbf{j}_1(p) - \mathbf{j}_2(p)]/2. \quad (30)$$

This noise map includes noise that is not correlated on timescales longer than 20 min. In particular,  $\mathbf{n}_{1+2}$  gives a good estimate of the white noise in  $\mathbf{m}_{1+2}$ .

However, we are interested in the noise level in the full map  $\mathbf{m}$ , (see Eq. (26)). To estimate this, we construct another noise map

$$\mathbf{n}_m(p) = \frac{\mathbf{j}_1(p) - \mathbf{j}_2(p)}{\mathbf{w}_{\text{hit}}(p)}, \quad (31)$$

with weights

$$\mathbf{w}_{\text{hit}}(p) = \sqrt{\mathbf{hit}_{\text{full}}(p) \left[ \frac{1}{\mathbf{hit}_1(p)} + \frac{1}{\mathbf{hit}_2(p)} \right]}. \quad (32)$$

Here  $\mathbf{hit}_{\text{full}}(p) = \mathbf{hit}_1(p) + \mathbf{hit}_2(p)$  is the hit count at pixel  $p$  in the full map  $\mathbf{m}$ , while  $\mathbf{hit}_1$  and  $\mathbf{hit}_2$  are the hit counts of  $\mathbf{j}_1$  and  $\mathbf{j}_2$ , respectively. The weight factor  $\mathbf{w}_{\text{hit}}(p)$  is equal to 2 only in those pixels where  $\mathbf{hit}_1(p) = \mathbf{hit}_2(p)$ . In a typical pixel,  $\mathbf{hit}_1(p)$  will differ slightly from  $\mathbf{hit}_2(p)$  and hence the weight factor is  $\mathbf{w}_{\text{hit}}(p) > 2$ .

Noise maps from half-ring jackknives are shown in the right-hand column of Fig. 10. A detailed comparison of the jackknife noise estimates and other noise estimates (WNC, noise Monte Carlo; see next section) are presented in the LFI instrument paper ([Mennella et al. 2011](#)).

## 7.4. Noise Monte Carlo simulations

To check the noise analysis, we produced Monte Carlo noise realizations on the ‘‘Louhi’’ supercomputer at ‘‘CSC-IT Center for Science’’ in Finland. The simulation takes as input estimates of the white noise  $\sigma_{\text{WN}}$ , knee frequency, and slope of the  $1/f$  noise estimated from the TOD for each radiometer ([Mennella et al. 2011](#)), as well as satellite pointing information. Flight pointing was reconstructed to machine accuracy using Planck Level-S simulation software [Reinecke et al. \(2006\)](#). For each frequency channel, we generated 101 Monte Carlo realizations of the noise, simulating white noise and correlated noise ( $1/f$ ) streams separately. Maps from these noise streams were produced with the map-making pipeline described in Sect. 7.1. For each simulated noise map, we computed the corresponding binned white noise maps defined in Eq. (27). The production of the binned white noise maps allows us to study the residual correlated noise, i.e., the difference between the total and binned white noise maps [Kurki-Suonio et al. \(2009\)](#). These Monte Carlo simulations were used to test and validate several approaches to noise estimation described in detail in the LFI instrument paper ([Mennella et al. 2011](#)).

## 8. Colour correction

The power measured by LFI can be expressed as

$$P = \frac{G}{2} \int g(\nu) \Delta T_{\text{RJ}}(\nu) d\nu, \quad (33)$$

where  $G$  is the overall gain,  $g(\nu)$  is the bandpass, and  $\Delta T_{\text{RJ}}$  is the Rayleigh-Jeans brightness temperature signal, in the case of LFI calibration procedure, due to the CMB dipole. At a given frequency  $\nu_0$ , the overall gain  $G$  is equal to  $2P/\Delta T_{\text{RJ}}(\nu_0)$ . For small

**Table 3.** Colour corrections for different input power-law spectral indices.

DETECTOR	SPECTRAL INDEX $\alpha$						
	-2.0	-1.0	0.0	1.0	2.0	3.0	4.0
<b>70 GHz</b>							
LFI18 .....	1.054	1.028	1.011	1.003	1.003	1.010	1.026
LFI19 .....	1.170	1.113	1.066	1.026	0.994	0.969	0.949
LFI20 .....	1.122	1.079	1.044	1.017	0.997	0.983	0.975
LFI21 .....	1.087	1.053	1.028	1.010	1.000	0.996	0.998
LFI22 .....	0.973	0.971	0.976	0.988	1.007	1.033	1.066
LFI23 .....	1.015	1.004	0.999	0.998	1.003	1.012	1.026
$\langle C \rangle_{70}$ .....	1.070	1.041	1.021	1.007	1.001	1.001	1.007
<b>44 GHz</b>							
LFI24 .....	1.028	1.015	1.007	1.002	1.000	1.003	1.009
LFI25 .....	1.039	1.024	1.013	1.005	1.000	0.999	1.000
LFI26 .....	1.050	1.032	1.017	1.007	1.000	0.997	0.997
$\langle C \rangle_{44}$ .....	1.039	1.024	1.012	1.004	1.000	0.999	1.002
<b>30 GHz</b>							
LFI27 .....	1.078	1.049	1.026	1.010	1.000	0.996	0.998
LFI28 .....	1.079	1.049	1.026	1.009	1.000	0.997	1.002
$\langle C \rangle_{30}$ .....	1.079	1.049	1.026	1.010	1.000	0.997	1.000

fluctuations around the mean CMB temperature  $T_0$ , the relation between intensity,  $I$ , Rayleigh-Jeans brightness temperature  $T_{\text{RJ}}$ , and thermodynamic temperature  $T$  is

$$\Delta I(\nu_0) = \frac{2k_B\nu_0^2}{c^2} \Delta T_{\text{RJ}}(\nu_0) = \left( \frac{\partial B(\nu_0, T)}{\partial T} \right)_{T_0} \Delta T. \quad (34)$$

The differential black-body spectrum is

$$\left( \frac{\partial B(\nu, T)}{\partial T} \right)_{T_0} = \frac{2k_B\nu^2}{c^2} e^{k\nu/k_B T_0} \left( \frac{h\nu/k_B T_0}{e^{h\nu/k_B T_0} - 1} \right)^2, \quad (35)$$

$$\equiv \frac{2k_B\nu^2}{c^2} \eta_{\Delta T}(\nu). \quad (36)$$

The function  $\eta_{\Delta T}(\nu)$  is the differential black-body spectrum in Rayleigh-Jeans units. With our definition of the overall gain  $G$ , the bandpasses are normalised such that

$$\int g(\nu) \eta_{\Delta T}(\nu) d\nu = \eta_{\Delta T}(\nu_0). \quad (37)$$

Calibration data provide a nominal brightness temperature  $\Delta \tilde{T}_{\text{RJ}} = (2/G)P$ ; however, this is only exact for a monochromatic response. For a non-zero bandwidth, a colour correction  $C(\alpha)$  is required to convert the brightness temperature for emission with a particular spectral index  $\alpha$  to that of the map:

$$C(\alpha) \Delta T_{\text{RJ}}(\nu_0) = \Delta \tilde{T}_{\text{RJ}} = \eta_{\Delta T}(\nu_0) \Delta \tilde{T}. \quad (38)$$

By definition, the colour correction is unity when the source observed has a CMB spectrum. Within each LFI band,  $g(\nu)$  is well-approximated by a power law with spectral index  $\alpha = 2 - (h\nu_0/k_B T)^2/6$ .

The general expression for the colour correction is

$$C(\alpha) = \left[ \frac{\eta_{\Delta T}(\nu_0)}{\int g(\nu) \eta_{\Delta T}(\nu) d\nu} \right] \int g(\nu) (\nu/\nu_0)^\alpha d\nu, \quad (39)$$

where we assumed a power-law spectrum with temperature spectral index  $\beta = \alpha - 2$ . The term in square brackets is unity with our normalisation for  $g(\nu)$ , but has been included to show that  $C(\alpha)$  depends only on the shape and not the amplitude of the bandpass. Thus  $C(\alpha)$  is independent of  $G$ .

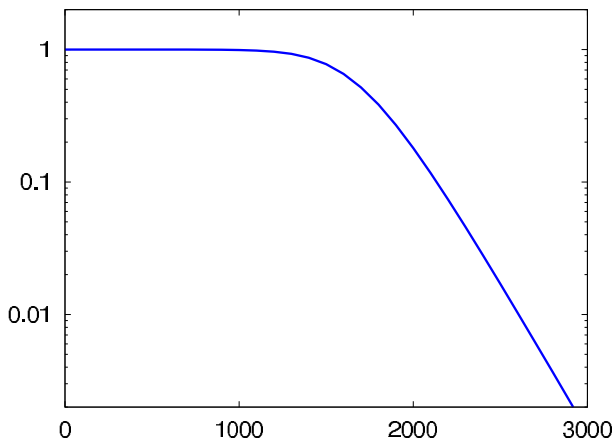
Each detector has a different bandpass, hence its own colour correction. We derive approximate colour corrections for band-averaged sky maps using bandpasses averaged over: (i) the two detectors in each radiometer; (ii) the two orthogonally-polarised radiometers behind each feed horn; and (iii) the several feed horns in each frequency band. In addition, although the bandpass is mainly defined by the front-end (Bersanelli et al. 2010), differences between back-end bandpasses on a single radiometer are measurable, e.g., in the form of  $\beta$ -dependent residuals in difference images.

Since the current sky maps have been produced, both for pairs of horns and for several horns in each band, with calibrated data combined with equal weights, we have used an unweighted average of all the contributing bandpasses for our band-averaged corrections. Using the bandpass models given in Zonca et al. (2009) derived from the pre-launch calibration campaign, we evaluate the integrals in Eq. (39) analytically for several spectral indices. The results are given in Table 3.

At the current stage of the mission and data analysis, uncertainties in the colour corrections are much smaller than those of the gains  $G$ ; however we aim to reduce the calibration error (using the orbital dipole) to below 0.2%. Two primary sources of error in  $C(\alpha)$  will then need to be considered. The first is related to uncertainties in the bandpass model (Leahy et al. 2010; Zonca et al. 2009). The second arises from the uneven sampling of individual sky pixels by the full set of detectors, which causes pixel-to-pixel variations in the colour correction.

## 9. CMB removal

This section was developed in common with HFI (Planck HFI Core Team 2011b) and is reported identically in both papers.



**Fig. 11.** Wiener-like filter function, plotted versus multipole, which was applied to produce the template for CMB removal.

In order to facilitate foreground studies with the frequency maps, a set of maps was constructed with an estimate of the CMB contribution subtracted from them. The steps undertaken in determining that estimate of the CMB map, subtracting it from the frequency maps, and characterising the errors in the subtraction are described below.

### 9.1. Masks

Point source masks were constructed from the source catalogues produced by the LFI pipeline for each of the LFI frequency channel maps. The algorithm used in the pipeline to detect the sources was a Mexican-hat wavelet filter. All sources detected with a signal-to-noise ratio greater than 5 were masked with a cut of radius  $3\sigma \approx 1.27$  FWHM of the effective beam. A similar process was applied to the HFI frequency maps (Planck HFI Core Team 2011b).

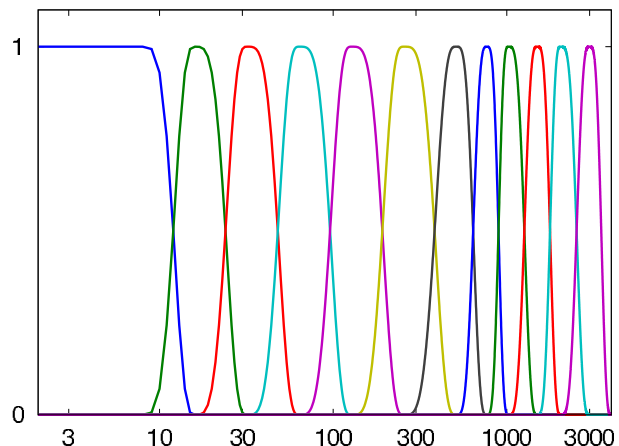
Galactic masks were constructed from the 30 GHz and 353 GHz frequency channel maps. An estimate of the CMB was subtracted from the maps in order not to bias the construction. The maps were smoothed to a common resolution of  $5^\circ$ . The pixels within each mask were chosen to be those with values above a threshold value. The threshold values were chosen to produce masks with the desired fraction of the sky remaining. The point source and Galactic masks were provided as additional inputs to the component separation algorithms.

### 9.2. Selection of the CMB template

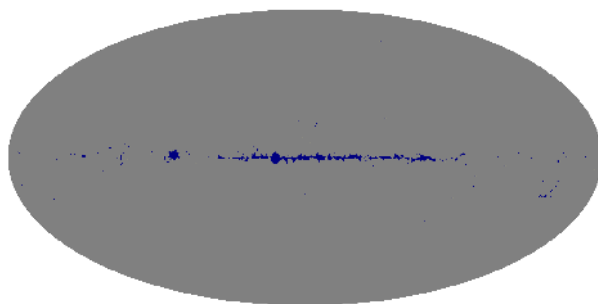
Six component separation or foreground removal algorithms were applied to the HFI and LFI frequency channel maps to produce CMB maps. They are, in alphabetical order:

- AltICA: internal linear combination (ILC) in the map domain;
- CCA: Bayesian component separation in the map domain;
- FastMEM: Bayesian component separation in the harmonic domain;
- Needlet ILC: ILC in the needlet (wavelet) domain;
- SEVEM: template fitting in map or wavelet domain;
- Wi-fit: template fitting in wavelet domain.

Details of these methods may be found in Leach et al. (2008). These six algorithms make different assumptions about the data, and may use different combinations of frequency channels used as input. Comparing results from these methods (see Fig. 14)



**Fig. 12.** The bandpass filters, plotted versus multipole, that define the spectral domains used in the NILC.



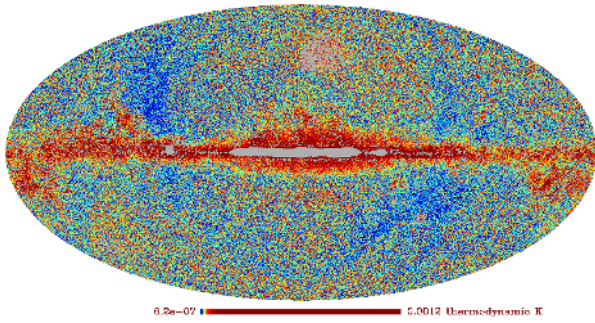
**Fig. 13.** Galactic mask used with NILC.

demonstrated the consistency of the CMB template and provided an estimate of the uncertainties in the reconstruction. A detailed comparison of the output of these methods, largely based on the CMB angular power spectrum, was used to select the CMB template that was removed from the frequency channel maps. The comparison was quantified using a jackknife procedure: each algorithm was applied to two additional sets of frequency maps made from the first half and second half of each pointing period. A residual map consisting of half the difference between the two reconstructed CMB maps was taken to be indicative of the noise level in the reconstruction from the full data set. The Needlet ILC (hereafter NILC) map was chosen as the CMB template because it had the lowest noise level at small scales.

The CMB template was removed from the frequency channel maps after application of a filter in the spherical harmonic domain. The filter has a transfer function made of two factors. The first corresponds to the Gaussian beam of the channel to be cleaned; the second is a transfer function attenuating the multipoles of the CMB template that have low signal-to-noise ratio. It is designed in Wiener-like fashion, being close to unity up to multipoles around  $\ell = 1000$ , then dropping smoothly to zero with a cut-off frequency around  $\ell = 1700$  (see Fig. 11). All angular frequencies above  $\ell = 3900$  are completely suppressed. This procedure was adopted to avoid doing more harm than good to the small scales of the frequency channel maps where the signal-to-noise ratio of the CMB is low.

### 9.3. Description of Needlet ILC

The NILC map was produced using the ILC method in the “needlet” domain. Needlets are spherical wavelets that allow localisation both in multipole and sky direction. The input maps



**Fig. 14.** Estimate of the rms error in the CMB subtraction. The map is histogram-equalised to bring out the details.

are decomposed into twelve overlapping multipole domains (called “scales”), using the bandpass filters shown in Fig. 12 and further decomposed into regions of the sky. Independent ILCs are applied in each sky region at each needlet scale. Large regions are used at large scales, while smaller regions are used at fine scales.

The NILC template was produced from all six HFI channels, using the tight Galactic mask shown in Fig. 13, which covers 99.36% of the sky. Additional areas are excluded on a per-channel basis to mask point sources. Future inclusion of the LFI channels will improve cleaning of low-frequency foregrounds such as synchrotron emission from the CMB template. Before applying NILC, pixels missing due to point source and Galactic masking are filled in by a “diffusive inpainting” technique, which consists of replacing each missing pixel by the average of its neighbours and iterating to convergence. This is similar to solving the heat diffusion equation in the masked areas with boundary conditions given by the available pixel values at the borders of the mask. All maps are re-beamed to a common resolution of  $5'$ . Re-beaming blows up the noise in the less resolved channels, but that effect is automatically taken into account by the ILC filter.

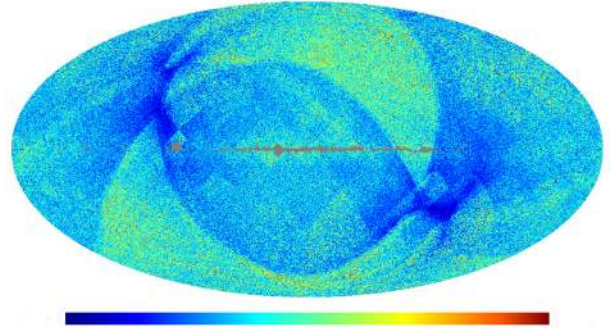
The CMB template obtained after NILC processing is filtered to have the “Wiener beam” shown in Fig. 11. The ILC coefficients are saved to be applied to the jackknife maps for performance evaluation as described in Sect. 9.4.2

#### 9.4. Uncertainties in the CMB removal

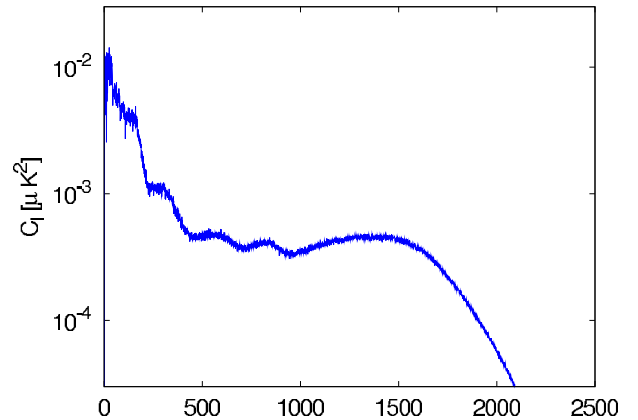
The uncertainties in the CMB removal have been gauged in two ways, firstly by comparing the CMB maps produced by the different algorithms and secondly by applying the NILC coefficients to jackknife maps.

##### 9.4.1. Dispersion of the CMB maps produced by the various algorithms

The methods that were used to produce the estimates of the CMB are diverse. They work by applying different algorithms (ILC, template fitting, or Bayesian parameter estimation) in a variety of domains (pixel space, Needlet/wavelet space, or spherical harmonic coefficients). Each method carries out its optimisation in a different way and thus will respond to the foregrounds differently. Dispersion in the CMB rendition by different methods provides an estimate of the uncertainties in the determination of the CMB, and thus in the subtraction process. The rms difference between the NILC map and the other CMB estimates is shown in Fig. 14. As expected, the uncertainties are largest in the Galactic



**Fig. 15.** Local rms of the noise (estimated by jackknife) in the NILC CMB map. The colour scale is from 0 to  $30 \mu\text{K}$  per pixel at resolution  $N_{\text{side}} = 2048$ .



**Fig. 16.** Angular spectrum in  $\mu\text{K}^2$  of the noise (estimated by jackknife) in the NILC CMB map. It corresponds to  $11 \mu\text{K}$  per pixel.

plane where the foregrounds to remove are strongest, and smallest around the Ecliptic poles where the noise levels are lowest.

##### 9.4.2. CMB map uncertainties estimated by applying NILC filtering of jackknives

The cleanliness of the CMB template produced by the NILC filter can be estimated using jackknives. We apply the NILC filter to the maps built from the first and last halves of the ring set. The power distribution of the half-difference of the results provides us with a reliable estimate of the power of the noise in the NILC CMB template, (while previous results correspond to applying the NILC filter to the half-sum maps from which they can be derived).

The jackknives allow estimates of the relative contributions of sky signal and noise to the total data power. Assume that the data are in the form  $X = S + N$  where  $S$  is the sky signal and  $N$  is the noise, independent of  $S$ . The total data power  $\text{Var}(X)$  decomposes as  $\text{Var}(X) = \text{Var}(S) + \text{Var}(N)$ . One can obtain  $\text{Var}(N)$  by applying the NILC filter to half difference maps, and  $\text{Var}(S)$  follows from  $\text{Var}(X) - \text{Var}(N)$ . This procedure can be applied in pixel space, in harmonic space, or in pixel space *after* the maps have been bandpass-filtered, as described next.

We first used pixel space jackknifing to estimate the spatial distribution of noise. Figure 15 shows a map of the local rms of the noise. We applied the NILC filter to a half-difference map and we display the square root of its smoothed squared values, effectively resulting in an estimate of the local noise rms. Using the same approach, we obtain an estimate of the angular spectrum of the noise in the NILC map, shown in Fig. 16.

That spectrum corresponds to an rms  $[(1/4\pi) \sum_{\ell}(2\ell + 1)C_{\ell}]^{1/2}$  of  $11 \mu\text{K}$  per pixel. The “features” in the shape of the noise angular spectrum at large scale are a consequence of the needlet-based filtering (such features would not appear in a pixel-based ILC map). Recall that the coefficients of an ILC map are adjusted to minimize the total contamination by both foregrounds *and* noise. The strength of foregrounds relative to noise being larger at coarse scales, the needlet-based ILC tends to let more noise in, with the benefit of better foreground rejection.

The half-difference maps offer simple access to the power distribution of the residual noise in the estimated CMB template. However, it is more difficult to evaluate other residual contamination, since all fixed sky emissions cancel in half difference maps. Any such large-scale contamination is barely visible in the CMB template, since it is dominated by the CMB itself. However, contamination is more conspicuous if one looks at intermediate scales. Figure 17 shows the local power of the CMB template after it is bandpassed to retain only multipoles in the range  $\ell = 500 \pm 200$ . This smooth version of the square of a bandpassed map clearly shows where the errors in the component separation become large and so complicate some specific science analyses.

## 10. Infrastructure overview

To organize the large number of data processing codes and data products, the DPC employs the *Planck* Integrated Data and Information System (IDIS). This allows flexible development of the processing pipeline, while ensuring complete traceability and reproducibility of data products. For this, the most relevant components of IDIS are the Data Management Component (DMC) and the Process Coordinator (ProC), developed at the MPA Planck Analysis Centre (MPAC) at the Max-Planck-Institute for Astrophysics in Garching. Access to both components as well as to the *Planck* document and software management system is controlled via another IDIS component: the Federation Layer developed and maintained at the ESA ESTEC Research and Scientific Support Department (RSSD).

Here we describe the essential features of the IDIS data processing components and their use at the DPC. A more detailed description of these components and their capabilities will be given in a future paper.

### 10.1. Data management component

The DMC organizes the storage and access to DPC data products. To combine optimal performance in data I/O with the data management capabilities of modern databases, scientific data are stored in files, while metadata identifying them are stored in a database. The data files can only be modified in synchronization with the database, preventing concurrent access to data objects via locking mechanisms. The DMC software supports several database management systems of various complexity; the LFI DPC operates an Oracle 10g database, which ensures good performance and stability.

The DMC provides a uniform Application Programming Interface (API) for Fortran, C, C++, and Java, hiding all specific database operations from the user, who is therefore not required to have database experience. DMC data types are defined in the Data Definition Layer (DDL), which describes data and metadata structures. The DDL supports inheritance of data types (e.g., a data type `polarized_map` can be inherited from a data type `map`) as well as association of data types (i.e., one data type containing a reference to another).

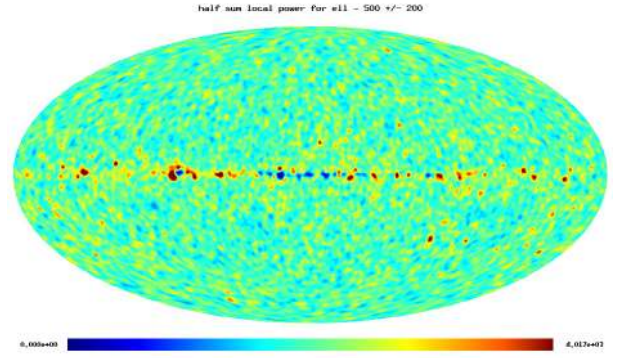


Fig. 17. Local power of the NILC CMB template in the range  $\ell = 500 \pm 200$ .

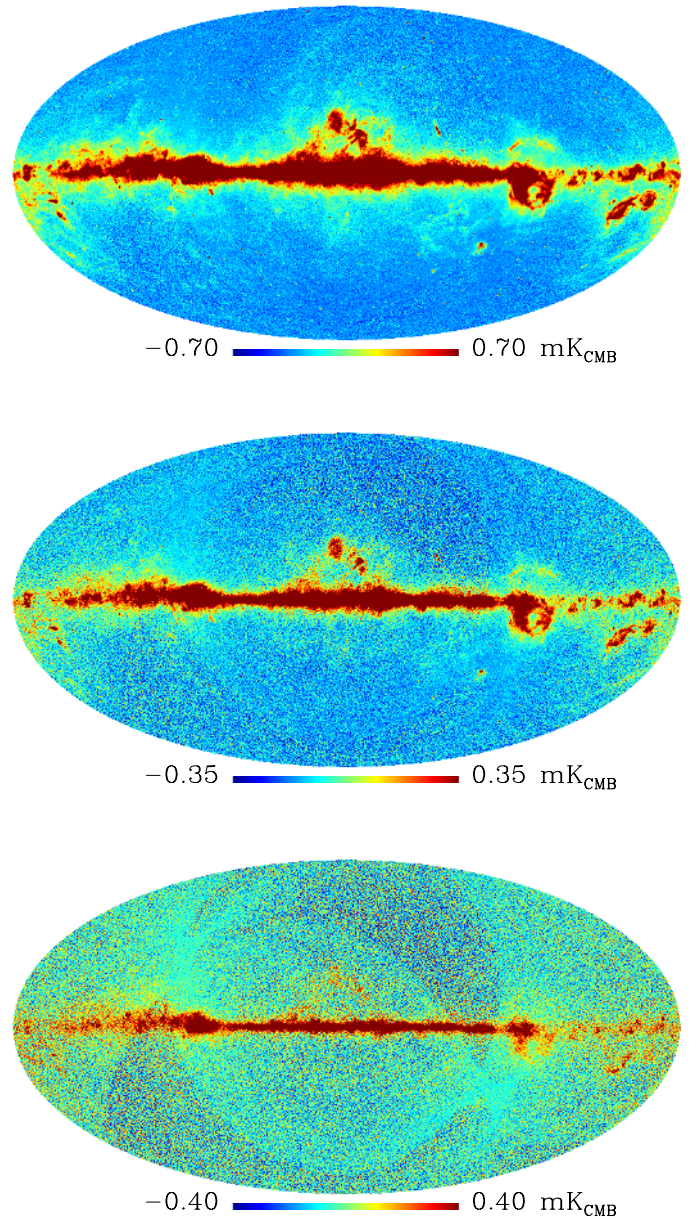


Fig. 18. CMB-removed channel maps. *From top to bottom*, 30, 44, and 70 GHz. The main galactic structures are clearly visible, as well as scanning strategy signatures at 44 and 70 GHz.



In addition to the API, the DMC provides a Graphical User Interface (GUI), which supports user queries of the database and retrieval of information on the data. The GUI offers the user the ability to list the (meta-)data of specific objects and also to visualize the data in a simple way (although data can also be exported to other powerful visualizing tools). The GUI also allows the display of history information on data objects, permitting the user to browse intermediate data products used in generating those objects, and the controlled deletion of data, observing dependencies of data types and maintaining the history information for all remaining data. For this, the DMC relies on additional metadata on the processing history of the data objects, which are generated by the Process Coordinator workflow engine (ProC).

### 10.2. Pipeline management – the ProC workflow engine

The ProC is a generic engine to construct, verify, and execute computational workflows. It comprises computing modules and data flows between them. The modules can be written in any programming language, provided they conform to simple I/O format requirements described in an XML module description file. These interface files specify the input and output objects, as well as the parameters of the individual programs, in terms of DMC data types as described above.

The ProC provides a pipeline editor to support graphical construction of data processing workflows. It allows users to arrange and connect computing modules of a workflow in a clearly structured manner, and at the same time to configure the parameters of the algorithms used. It provides control structures for data flow, for data object I/O and consistent parameter definition.

The execution of workflows is controlled by a forward chaining algorithm, which ensures that modules are executed as soon as all necessary data products and parameters are known. If the same version of a module has been executed with identical inputs and parameters, the ProC will skip the execution and use the data product from the earlier execution for further processing. The ProC maintains control of pipeline execution also on massive parallel computing environments. In the LFI implementation, the ProC communicates with the PBS (Portable Batch System) scheduling system to send jobs to the DPC cluster and to log their execution status.

The ProC logs workflow executions on log files, which can also contain logging messages of the executed modules. Additionally, it creates so-called *Pipeline-Run* and *Module-Run* objects in the DMC, which are used to recover the generation history of data products (including versions of processing modules via MD5-sums). Besides the GUI, the ProC can also be executed from the command line.

At the LFI DPC, the ProC is used to execute the official pipeline producing *Planck* data products.

## 11. Discussion and conclusions

We have described the status of the pipeline as it stands at the time of the ERCSC release and submission of the *Planck* early papers. All the algorithms run during this process have been verified, validated, and tested before launch and the start of operations using realistic simulations. This allowed us to begin analyzing the data as soon as they were acquired from the first day of operations. The entire Level 1 pipeline suffered no significant problems, and all of the data were transformed efficiently from telemetry packets to timelines. At present, the Level 2 pipeline is capable of providing relative calibration to an overall statistical accuracy in the range 0.05–0.1% and absolute calibration

at around the 1% level. The beams are accurately characterised down to  $-10$  dB. We expect to improve many aspects in the near future. Concerning the calibration, our intention is to reach the levels determined by the stability of the instrument. For the beam reconstruction, our aim is to improve the characterisation of the far side lobes and to refine the entire beam reconstruction pipeline, with particular attention to polarization measurements.

*Acknowledgements.* *Planck* is too large a project to allow full acknowledgement of all contributions by individuals, institutions, industries, and funding agencies. The main entities involved in the mission operations are as follows. The European Space Agency operates the satellite via its Mission Operations Centre located at ESOC (Darmstadt, Germany) and coordinates scientific operations via the Planck Science Office located at ESAC (Madrid, Spain). Two Consortia, comprising around 50 scientific institutes within Europe, the USA, and Canada, and funded by agencies from the participating countries, developed the scientific instruments LFI and HFI, and continue to operate them via Instrument Operations Teams located in Trieste (Italy) and Orsay (France). The Consortia are also responsible for scientific processing of the acquired data. The Consortia are led by the Principal Investigators: J.L. Puget in France for HFI (funded principally by CNES and CNRS/INSU-IN2P3) and N. Mandolesi in Italy for LFI (funded principally via ASI). NASA US Planck Project, based at J.P.L. and involving scientists at many US institutions, contributes significantly to the efforts of these two Consortia. The author list for this paper has been selected by the Planck Science Team, and is composed of individuals from all of the above entities who have made multi-year contributions to the development of the mission. It does not pretend to be inclusive of all contributions. The *Planck*-LFI project is developed by an International Consortium lead by Italy and involving Canada, Finland, Germany, Norway, Spain, Switzerland, UK, USA. The Italian contribution to *Planck* is supported by the Italian Space Agency (ASI) and INAF. This work was supported by the Academy of Finland grants 121703 and 121962. We thank the DEISA Consortium (<http://www.deisa.eu>), co-funded through the EU FP6 project RI-031513 and the FP7 project RI-222919, for support within the DEISA Virtual Community Support Initiative. We thank CSC – IT Center for Science Ltd (Finland) for computational resources. We acknowledge financial support provided by the Spanish Ministerio de Ciencia e Innovación through the Plan Nacional del Espacio y Plan Nacional de Astronomía y Astrofísica. We acknowledge The Max Planck Institute for Astrophysics Planck Analysis Centre (MPAC) is funded by the Space Agency of the German Aerospace Center (DLR) under grant 500P0901 with resources of the German Federal Ministry of Economics and Technology, and by the Max Planck Society. This work has made use of the Planck satellite simulation package (Level-S), which is assembled by the Max Planck Institute for Astrophysics Planck Analysis Centre (MPAC) Reinecke et al. (2006). We acknowledge financial support provided by the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Some of the results in this paper have been derived using the HEALPix package Górski et al. (2005). 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.rssd.esa.int/index.php?project=PLANCK&page=Planck\\_Collaboration](http://www.rssd.esa.int/index.php?project=PLANCK&page=Planck_Collaboration).

## References

- Ashdown, M. A. J., Baccigalupi, C., Balbi, A., et al. 2007a, *A&A*, 471, 361
- Ashdown, M. A. J., Baccigalupi, C., Balbi, A., et al. 2007b, *A&A*, 467, 761
- Ashdown, M. A. J., Baccigalupi, C., Bartlett, J. G., et al. 2009, *A&A*, 493, 753
- Bersanelli, M., Mandolesi, N., Butler, R. C., et al. 2010, *A&A*, 520, A4
- Burigana, C., Natoli, P., Vittorio, N., Mandolesi, N., & Bersanelli, M. 2001, *Exp. Astron.*, 12, 87
- Cappellini, B., Maino, D., Albeti, G., et al. 2003, *A&A*, 409, 375
- de Gasperis, G., Balbi, A., Cabella, P., Natoli, P., & Vittorio, N. 2005, *A&A*, 436, 1159
- Frailis, M., Maris, M., Zacchei, A., et al. 2009, *J. Instr.*, 4, 2021
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759
- Hinshaw, G., Spergel, D. N., Verde, L., et al. 2003, *ApJS*, 148, 135
- Hoffman, Y., & Ribak, E. 1991, *ApJ*, 380, L5
- Keihänen, E., Keskitalo, R., Kurki-Suonio, H., Poutanen, T., & Sirviö, A. 2010, *A&A*, 510, A57
- Keihänen, E., Kurki-Suonio, H., & Poutanen, T. 2005, *MNRAS*, 360, 390
- Kurki-Suonio, H., Keihänen, E., Keskitalo, R., et al. 2009, *A&A*, 506, 1511
- Lamarre, J., Puget, J., Ade, P. A. R., et al. 2010, *A&A*, 520, A9
- Leach, S. M., Cardoso, J., Baccigalupi, C., et al. 2008, *A&A*, 491, 597
- Leahy, J. P., Bersanelli, M., D’Arcangelo, O., et al. 2010, *A&A*, 520, A8

- Maino, D., Burigana, C., Górski, K. M., Mandolesi, N., & Bersanelli, M. 2002, A&A, 387, 356
- Mandolesi, N., Bersanelli, M., Butler, R. C., et al. 2010, A&A, 520, A3
- Maris, M., Tomasi, M., Galeotta, S., et al. 2009, J. Instr., 4, 2018
- Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., & Picaud, S. 2006, A&A, 453, 635
- Massardi, M. 2006, in CMB and Physics of the Early Universe
- Meinhold, P., Leonardi, R., Aja, B., et al. 2009, J. Instr., 4, 2009
- Mennella, A., Bersanelli, M., Seiffert, M., et al. 2003, A&A, 410, 1089
- Mennella, A., Bersanelli, M., Butler, R. C., et al. 2010, A&A, 520, A5
- Mennella, A., Butler, R. C., Curto, A., et al. 2011, A&A, 536, A3
- Natoli, P., de Gasperis, G., Gheller, C., & Vittorio, N. 2001, A&A, 372, 346
- Pasian, F., & Gispert, R. 2000, *Astrophys. Lett. Comm.*, 37, 247
- Planck Collaboration 2011a, A&A, 536, A1
- Planck Collaboration 2011b, A&A, 536, A2
- Planck Collaboration 2011c, A&A, 536, A7
- Planck Collaboration 2011d, A&A, 536, A8
- Planck Collaboration 2011e, A&A, 536, A9
- Planck Collaboration 2011f, A&A, 536, A10
- Planck Collaboration 2011g, A&A, 536, A11
- Planck Collaboration 2011h, A&A, 536, A12
- Planck Collaboration 2011i, A&A, 536, A13
- Planck Collaboration 2011j, A&A, 536, A14
- Planck Collaboration 2011k, A&A, 536, A15
- Planck Collaboration 2011l, A&A, 536, A16
- Planck Collaboration 2011m, A&A, 536, A17
- Planck Collaboration 2011n, A&A, 536, A18
- Planck Collaboration 2011o, A&A, 536, A19
- Planck Collaboration 2011p, A&A, 536, A20
- Planck Collaboration 2011q, A&A, 536, A21
- Planck Collaboration 2011r, A&A, 536, A22
- Planck Collaboration 2011s, A&A, 536, A23
- Planck Collaboration 2011t, A&A, 536, A24
- Planck Collaboration 2011u, A&A, 536, A25
- Planck Collaboration 2011v, The Explanatory Supplement to the Planck Early Release Compact Source Catalogue (ESA)
- Planck Collaboration 2011w, A&A, 536, A26
- Planck HFI Core Team 2011a, A&A, 536, A4
- Planck HFI Core Team 2011b, A&A, 536, A6
- Prunet, S., Ade, P. A. R., Bock, J. J., et al. 2001 [arXiv:astro-ph/0101073]
- Reinecke, M., Dolag, K., Hell, R., Bartelmann, M., & EnBlin, T. A. 2006, A&A, 445, 373
- Rosset, C., Tristram, M., Ponthieu, N., et al. 2010, A&A, 520, A13
- Tauber, J. A., Mandolesi, N., Puget, J., et al. 2010, A&A, 520, A1
- Zacchei, A., Frailis, M., Maris, M., et al. 2009, J. Instr., 4, 2019
- Zacchei, A., Maino, D., Baccigalupi, C., et al. 2011, A&A, 536, A5
- Zonca, A., Franceschet, C., Battaglia, P., et al. 2009, J. Instr., 4, 2010
- <sup>14</sup> Department of Physics and Astronomy, University of Southern California, Los Angeles, California, USA
- <sup>15</sup> Department of Physics, Gustaf Hällströmin katu 2a, University of Helsinki, Helsinki, Finland
- <sup>16</sup> Department of Physics, University of California, Berkeley, California, USA
- <sup>17</sup> Department of Physics, University of California, One Shields Avenue, Davis, California, USA
- <sup>18</sup> Department of Physics, University of California, Santa Barbara, California, USA
- <sup>19</sup> Department of Physics, University of Oxford, 1 Keble Road, Oxford, UK
- <sup>20</sup> Dipartimento di Fisica G. Galilei, Università degli Studi di Padova, via Marzolo 8, 35131 Padova, Italy
- <sup>21</sup> Dipartimento di Fisica, Università La Sapienza, P. le A. Moro 2, Roma, Italy
- <sup>22</sup> Dipartimento di Fisica, Università degli Studi di Milano, via Celoria, 16, Milano, Italy
- <sup>23</sup> Dipartimento di Fisica, Università degli Studi di Trieste, via A. Valerio 2, Trieste, Italy
- <sup>24</sup> Dipartimento di Fisica, Università di Ferrara, via Saragat 1, 44122 Ferrara, Italy
- <sup>25</sup> Dipartimento di Fisica, Università di Roma Tor Vergata, via della Ricerca Scientifica, 1, Roma, Italy
- <sup>26</sup> Dipartimento di Matematica, Università di Roma Tor Vergata, via della Ricerca Scientifica, 1, Roma, Italy
- <sup>27</sup> Dpto. Astrofísica, Universidad de La Laguna (ULL), 38206 La Laguna, Tenerife, Spain
- <sup>28</sup> European Space Agency, ESAC, Planck Science Office, Camino bajo del Castillo, s/n, Urbanización Villafraanca del Castillo, Villanueva de la Cañada, Madrid, Spain
- <sup>29</sup> European Space Agency, ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
- <sup>30</sup> Haverford College Astronomy Department, 370 Lancaster Avenue, Haverford, Pennsylvania, USA
- <sup>31</sup> Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland
- <sup>32</sup> INAF – Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, Firenze, Italy
- <sup>33</sup> INAF – Osservatorio Astrofisico di Catania, via S. Sofia 78, Catania, Italy
- <sup>34</sup> INAF – Osservatorio Astronomico di Padova, Vicolo dell’Osservatorio 5, Padova, Italy
- <sup>35</sup> INAF – Osservatorio Astronomico di Roma, via di Frascati 33, Monte Porzio Catone, Italy
- <sup>36</sup> INAF – Osservatorio Astronomico di Trieste, via G.B. Tiepolo 11, Trieste, Italy
- <sup>37</sup> INAF/IASF Bologna, via Gobetti 101, Bologna, Italy
- <sup>38</sup> INAF/IASF Milano, via E. Bassini 15, Milano, Italy
- <sup>39</sup> INRIA, Laboratoire de Recherche en Informatique, Université Paris-Sud 11, Bâtiment 490, 91405 Orsay Cedex, France
- <sup>40</sup> ISDC Data Centre for Astrophysics, University of Geneva, ch. d’Ecogia 16, Versoix, Switzerland
- <sup>41</sup> Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA
- <sup>42</sup> Institut d’Astrophysique Spatiale, CNRS (UMR8617) Université Paris-Sud 11, Bâtiment 121, Orsay, France
- <sup>43</sup> Institut d’Astrophysique de Paris, CNRS UMR7095, Université Pierre & Marie Curie, 98 bis boulevard Arago, Paris, France
- <sup>44</sup> Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
- <sup>45</sup> Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo, Norway
- <sup>46</sup> Instituto de Astrofísica de Canarias, C/Vía Láctea s/n, La Laguna, Tenerife, Spain
- <sup>47</sup> Instituto de Física de Cantabria (CSIC-Universidad de Cantabria), Avda. de los Castros s/n, Santander, Spain
- <sup>48</sup> Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California, USA
- <sup>1</sup> Aalto University Metsähovi Radio Observatory, Metsähovintie 114, 02540 Kylmälä, Finland
- <sup>2</sup> Agenzia Spaziale Italiana Science Data Center, c/o ESRIN, via Galileo Galilei, Frascati, Italy
- <sup>3</sup> Agenzia Spaziale Italiana, Viale Liegi 26, Roma, Italy
- <sup>4</sup> Astroparticule et Cosmologie, CNRS (UMR7164), Université Denis Diderot Paris 7, Bâtiment Condorcet, 10 rue A. Domon et Léonie Duquet, Paris, France
- <sup>5</sup> Australia Telescope National Facility, CSIRO, PO Box 76, Epping, NSW 1710, Australia
- <sup>6</sup> CNR - ISTI, Area della Ricerca, via G. Moruzzi 1, Pisa, Italy
- <sup>7</sup> CNRS, IRAP, 9 Av. colonel Roche, BP 44346, 31028 Toulouse Cedex 4, France
- <sup>8</sup> California Institute of Technology, Pasadena, California, USA
- <sup>9</sup> Centre of Mathematics for Applications, University of Oslo, Blindern, Oslo, Norway
- <sup>10</sup> DTU Space, National Space Institute, Juliane Mariesvej 30, Copenhagen, Denmark
- <sup>11</sup> Departamento de Física, Universidad de Oviedo, Avda. Calvo Sotelo s/n, Oviedo, Spain
- <sup>12</sup> Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Avda. de los Castros s/n, Santander, Spain
- <sup>13</sup> Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia, Canada

- <sup>49</sup> Jodrell Bank Centre for Astrophysics, Alan Turing Building, School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK
- <sup>50</sup> LERMA, CNRS, Observatoire de Paris, 61 Avenue de l'Observatoire, Paris, France
- <sup>51</sup> Lawrence Berkeley National Laboratory, Berkeley, California, USA
- <sup>52</sup> Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
- <sup>53</sup> MilliLab, VTT Technical Research Centre of Finland, Tietotie 3, Espoo, Finland
- <sup>54</sup> SISSA, Astrophysics Sector, via Bonomea 265, 34136 Trieste, Italy
- <sup>55</sup> Space Sciences Laboratory, University of California, Berkeley, California, USA
- <sup>56</sup> Spitzer Science Center, 1200 E. California Blvd., Pasadena, California, USA
- <sup>57</sup> Université de Toulouse, UPS-OMP, IRAP, 31028 Toulouse Cedex 4, France
- <sup>58</sup> Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland