Linearized iterative least-squares (LIL): A parameter fitting algorithm for component separation in multifrequency CMB experiments such as Planck

Rishi Khatri[1]⋆

[1] *Max-Planck Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85740 Garching, Germany*

29 October 2014

ABSTRACT

We present an efficient algorithm for the least squares parameter fitting optimized for component separation in multi-frequency CMB experiments like Planck. We sidestep some of the problems associated with non-linear optimization by taking advantage of the quasi-linear nature of foreground model i.e. a subset of parameters in the model are non-linear and even they are expected to lie within a very narrow range from physical considerations. We demonstrate our algorithm, linearized iterative least-squares (LIL), on the publicly available Planck sky model FFP6 simulations and compare our result with the other algorithms, in particular the COMMANDER algorithm which is a Bayesian parameter fitting algorithm. We work at full Planck resolution and show that degrading the resolution of all channels to that of the lowest frequency channel is not necessary. We show that the performance of LIL in terms of residuals and noise is comparable to the existing methods. Finally we present results for the publicly available Planck data. Our algorithm is extremely fast, fitting 6 parameters to 7 lowest Planck channels at full resolution (50 million pixels) in less than 160 CPU-minutes (or few minutes running in parallel on few tens of cores). LIL is therefore easily scalable to future experiments which may have even higher resolution and more frequency channels requiring more complicated models with more parameters. We also naturally propagate the uncertainties in different parameters due to noise in the maps as well as degeneracies between the parameters to the final errors on the parameters using Fisher matrix. One indirect application of LIL could be a front-end for Bayesian parameter fitting to find the maximum of the likelihood to be used as the starting point for the Gibbs sampling. Our algorithm is also useful for exploring different models of foregrounds and the effect on noise and systematic properties of resulting CMB maps. We show that fitting a single most general model to all pixels is not optimal since not all components are present everywhere and argue, with the example of CO component, that the optimal approach should combine parameter fitting with model selection. LIL may also be useful in other astrophysical applications which satisfy the quasi-linearity criteria.

Key words: Cosmology: cosmic microwave background – theory – observations

1 INTRODUCTION

Multi-frequency CMB experiments such as WMAP (Bennett et al. 2013) and Planck (Planck Collaboration et al. 2013a) make it possible to separate the observed signal into CMB and foreground components by taking advantage of the fact that different components have different spectral properties. Component separation allows us to use a larger fraction of the sky for cosmological analysis. For high sensitivity experiments, such as Planck, even in the relatively clean parts of the sky the foregrounds would be above the noise level and we need component separation to take advantage of the full sensitivity of the experiment. Although most component sep-

aration methods assume quite simple foreground models, the foregrounds in reality are much more complicated. We are limited to simple models mostly because of the lack of spectral resolution in the CMB experiments. Planck has just 7 frequency bands from 30 GHz to 353 GHz where CMB dominates over a significant portion of the sky and an additional two bands at 545 GHz and 857 GHz which are dominated by the foregrounds over most of the sky. Foregrounds on the other hand are expected to have multiple components such as synchrotron (Haslam et al. 1981), free-free (Bennett et al. 2003), anomalous dust emission (Erickson 1957; Draine & Lazarian 1998a,b, 1999) and dust emission (Wheelock et al. 1994; Smoot et al. 1992; Sodroski et al. 1994; Fixsen et al. 1996; Schlegel et al. 1998; Finkbeiner et al. 1999). Each of these components is in fact made up of superposition of emission from different regions

along the line of sight with different temperatures, spectral indices and intensity. In addition there is also a contribution from the extra-galactic radio (Longair & Sunyaev 1972; Gervasi et al. 2008) and infrared backgrounds (Hauser et al. 1998).

The component separation methods must therefore make many simplifying assumptions about the foregrounds. Although the details of different methods differ, broadly they can be classified according to whether they assume that the emission laws vary over the sky or not. For example, the template fitting method Spectral Estimation Via Expectation Maximization (SEVEM) (Martínez-González et al. 2003) makes the assumption that the spectral parameters are constant over large regions of the sky so that a template subtraction can be performed. The Spectral Matching Independent Component Analysis (SMICA) (Cardoso et al. 2008) and its Bayesian incarnation (Vansyngel et al. 2014) construct and fit a template or model of foreground emission, CMB and noise to the channel maps in the spherical harmonic domain and also makes the assumption that the emission laws do not vary with the spherical harmonic mode numbers ℓ, m i.e. they are constant over the whole sky. COMMANDER Eriksen et al. (2006, 2008) uses a simplified parametric model motivated by the known physics about the foregrounds. The spectral parameters vary from pixel to pixel on a much lower resolution map and this method is therefore more flexible in modelling the varying emission laws on the sky and a little closer to the physical nature of the foregrounds. The Needlet Internal Linear Combination algorithm (NILC) Delabrouille et al. (2009) constructs an internal linear combination map in a wavelet space whose basis functions are needlets. The needlets are localized both in the spatial and harmonic space and this method lies somewhere in between the SEVEM and SMICA in terms of the basis. It fits the foregrounds as a function of both spatial location as well as angular scale. This method also assumes that the emission laws over large portions of the sky are constant. The foregrounds also provide useful information about the galactic and extragalactic physics and accurate estimation of the foregrounds components is therefore also important. The component separation methods referred to above, except for COMMANDER, do not separate different foreground components but only the sum of the foregrounds from the CMB.

It is also a reasonable question to ask if the assumptions about the foregrounds and the CMB can bias our CMB results. In particular can the foreground cleaning/component separation introduce anomalies in the resulting CMB maps or, what may be even worse, mask some of the primordial anomalies. It is therefore worthwhile to have different foreground removal methods which operate on different assumptions. But more importantly we need methods which rely on minimum amount assumptions about the nature of foregrounds and the CMB, especially with regards to the angular correlation structure of both the amplitudes and the emission laws, and only use what we physically know about the foregrounds.

The purpose of the present work is to present such a method. Our approach is parameter fitting in pixel space, similar to the COMMANDER, but we do so at full resolution relaxing the assumption that the spectral parameters are constant on degree scales. In particular, we find that there is no need to smooth the resolution of all maps to the lowest resolution map. In fact re-beaming the low resolution maps to higher resolution, as is done by SMICA and NILC, would be a better choice yielding CMB maps with an effective beam closer to that of the highest resolution channels. COMMANDER explores the full posterior distribution using Gibbs sampling Eriksen et al. (2006, 2008) and is therefore computationally very expensive with 100 s per pixel. COMMANDER simultaneously esti-

mates the angular power spectrum of the CMB also which we will not be concerned about in this paper. We however still need to produce high quality maps if we want to look for anomalies that go beyond the isotropic power spectrum.

Our approach is straightforward least-squares parameter fitting. This has been attempted before using non-linear optimization algorithms Brandt et al. (1994). The non-linear optimization algorithms however converge quite slowly and often to the local minimum which may be far from the solution we are interested in. We develop a new algorithm taking advantage of the fact that most of the parameters in the foreground model are linear, for example, the amplitudes of different components. Even the non-linear parameters lie within a narrow range. For example, the low frequency foregrounds originate in free-free and synchrotron emission and are expected to have spectral indices in the range -2 to -4. Similarly the spectral index of the dust component is expected to lie between 2 and 3 (see the next section for the exact definitions of spectral indices and the foreground model) and its temperature between $\sim$ 10 and 30 K. This suggests that we can Taylor expand the foreground model around a reasonable guess and solve the resulting linear problem. Since the Taylor series is a good approximation within a narrow range around the central value, we expect the linear approximation to be a very good one. This forms the basis of our algorithm presented in the next section. In section 2 we present our algorithm explaining the reasoning behind the different steps. In section 3 we apply our algorithm to the publicly available simulations of the Planck sky model (Delabrouille et al. 2013) called FFP6 simulations[1]. In section 4 we discuss a further refinement/extension of our algorithm by including model selection. In section 5 we apply our algorithm to the publicly available Planck maps at full resolution and present the resulting CMB and foreground component maps. Our results as well as the FORTRAN code will be made publicly available at `http://www.mpa-garching.mpg.de/~khatri/lilcmb`.

2 LINEARIZED ITERATIVE LEAST-SQUARES PARAMETER FITTING (LIL)

For definiteness we will work with Planck experiment. However, the algorithm is quite general and is applicable to future experiments with many more frequency channels compared to Planck (The COrE Collaboration 2011; Kogut et al. 2011; P. Andre et al. 2014) and also to more complicated foreground models. Following Planck Collaboration et al. (2013b) we fit the following 6-parameter model to the 7 Planck frequency channels from 30 GHz to 353 GHz. In the following we will refer to the observed value of data = signal + noise in a particular pixel (p) in frequency channel j as $d_j(p)$ and we number the frequency channels from lowest to highest i.e. $j \in \{1, \ldots, 7\}$ for $\nu \in \{30, \ldots, 353\}$ GHz respectively. Our parametric model is

$$s_\nu(p) = A_{CMB}(p) + f_\nu^{lf} A_{lf}(p) \left(\frac{\nu}{\nu_0^{lf}}\right)^{\beta_{lf}(p)} + f_\nu^{co} A_{co}(p)$$

$$+ f_\nu^d A_d(p) \frac{1}{\exp\left(\frac{h\nu}{k_B T_d}\right) - 1} \left(\frac{\nu}{\nu_0^d}\right)^{\beta_d(p)}, \qquad (1)$$

---

[1] `http://wiki.cosmos.esa.int/planckpla/index.php/Simulation_data`

where p is the pixel number in HEALPix[2] nested numbering scheme (Górski et al. 2005), s_ν is the total sky emission in units of K_{CMB} at frequency ν, $A_i(p)$ is the amplitude of the component $i \in \{CMB, co, lf, d\}$ where the abbreviations are for the CMB, CO line emission, low frequency emission including synchrotron, free-free and anomalous microwave emission, and dust emission respectively. The factors of f_ν^i convert the Rayleigh Jeans temperature (or K_{RJ} Km/s for CO) to the thermodynamic CMB units, K_{CMB}, and also include the color correction factors, ν_0^i are the reference frequencies with $\nu_0^{lf} = 30$ GHz and $\nu_0^d = 353$ GHz and β_i are the spectral indices. The f_ν^{co} in addition includes relative amplitudes of the CO lines in units of K_{RJ} Km/s in different frequency channels which we keep fixed following Planck Collaboration et al. (2013b) as (1:0.595:0.297) for 100 GHz , 217 GHz and 353 GHz channels respectively and zero for all other channels. We will assume the dust temperature to be constant, $T_d = 18$ K. In principle it is possible to include higher frequency channels at 545 GHz and 857 GHz and allow the dust temperature to vary. However, it is not expected to lead to a significant improvement in the foreground separation as far as CMB is concerned since the additional information provided by the higher frequency channels is absorbed in the additional complexity of the model as well as having to apply the simple dust emission model over a wider range of frequencies than is physically justified. Note that our definitions are slightly different from (Planck Collaboration et al. 2013b) to make the formulae a little simpler. We will also suppress the argument p in the following since we work with one pixel at a time, so the value of data in j^{th} frequency map is denoted by just d_j.

We want to do a Taylor series expansion of Eq. 1 around some initial guess for the parameters. Therefore the zeroth step for our algorithm is the following:

(0) Set the initial values of parameters for the current pixel as follows: $A_{CMB} = \min(d_4, d_5, d_6)$, $A_{lf} = (d_1 - A_{CMB})/f_{30\ GHz}^{lf}$, $A_d = (d_7 - A_{CMB})/f_{353\ GHz}^d$, where we ignore the color correction in f_ν^d and f_ν^{lf} which then just convert from Rayleigh-Jeans units to the thermodynamic units. Similarly using $d_1 - A_{CMB}$ and $d_2 - A_{CMB}$ as estimates of low frequency foreground in the lowest two channels we can solve for β_{lf} and using $d_7 - A_{CMB}$ and $d_6 - A_{CMB}$ as estimates of dust foregrounds we can solve for the initial β_d. Note that we do not need initial guess for CO amplitude since it is linear and sub-dominant compared to the other components.

(1) Do Taylor expansion around the current value of the parameters (or the initial guess if first iteration).

$$s_\nu \approx s_\nu' = A_{CMB} + f_\nu^{co} A_{co} + f_\nu^{lf} A_{lf} \left(\frac{\nu}{\nu_0^{lf}}\right)^{\beta_{lf}} \left[\delta_{lf} + \delta_\beta^{lf}\beta_{lf} \ln\left(\frac{\nu}{\nu_0^{lf}}\right)\right]$$
$$+ f_\nu^d A_d \frac{1}{\exp\left(\frac{h\nu}{k_B T_d}\right) - 1} \left(\frac{\nu}{\nu_0^d}\right)^{\beta_d} \left[\delta_d + \delta_\beta^d\beta_d \ln\left(\frac{\nu}{\nu_0^d}\right)\right] \quad (2)$$

In the above equation the $A_{lf}, \beta_{lf}, A_d, \beta_d$ are fixed from previous iteration. We thus have a 6 parameter linear model with the parameter vector $x = (A_{CMB}, A_{co}, \delta_{lf}, \delta_\beta^{lf}, \delta_d, \delta_\beta^d)$. We have defined δ_{lf} and δ_d as multiplicative corrections to the amplitude while the spectral indices $\delta_\beta^{lf}, \delta_\beta^d$ are fractional corrections so that the actual indices are $\beta_{lf}(1 + \delta_\beta^{lf}), \beta_d(1 + \delta_\beta^d)$ and we expect that $\delta_\beta^d, \delta_\beta^{lf} \lesssim 1$ so that the Taylor series is a good approximation.

(2) We now solve the linear least squares problem minimizing

$$\chi^2 = \sum_i \left(\frac{s_{\nu_i}' - d_i}{\sigma_i}\right)^2 \equiv \sum_i [(Mx)_i - d_i']^2, \quad (3)$$

where $d_i' \equiv d_i/\sigma_i$ and σ_i is the standard deviation of noise in channel i in the current pixel p. The problem is easily solved by numerous linear algebra techniques. We use LQ/QR factorization routines of the Intel Math Kernel Library. The solution of the least squares problem gives us a direction to move our current parameter vector as well as the amplitude of the step which is just the actual least squares solution. In practice we may decide not to take the full step but only move a fraction of the amplitude in the relevant direction. We will come back to this point below.

So far we have done nothing new. The steps (1) and (2) are in fact just the standard Taylor expansion to linear (or quadratic order) and a Gauss-Newton step. These two steps, with or without some modification, form the basis of most non-linear optimization algorithms (see e.g. Gill & Murray 1976; Press et al. 1992; Conn et al. 2000). From Eq. 2 it is clear why we do not arrive at the correct solution in just one step. Although this linear model is a good approximation, in fact the term multiplying the amplitude parameters δ_{lf} and δ_d have the wrong spectral indices since they also change in the step from their current values and vice-versa for the terms multiplying the spectral index parameters δ_β^{lf} and δ_β^d. This is in fact the only place the non-linearity of our model manifests itself. This also suggests a cure.

(3) Use the results of step (2) only to update the spectral indices $\beta_{lf} \to \beta_{lf}(1 + \delta_\beta^{lf})$ and $\beta_d \to \beta_d(1 + \delta_\beta^d)$. Fix the spectral indices and repeat step (2) for the smaller parameter vector $x' = (A_{CMB}, A_{co}, \delta_{lf}, \delta_d)$. Update the amplitudes with the new solution.

With the spectral indices fixed the problem is in fact linear and can be solved exactly. Thus we arrive at the minimum in a subspace of our full parameter space.

(4) Exit if we meet one of the exit criterion given below otherwise go back to step (1)

We are therefore searching for a minimum in the two non-linear directions, which are almost uncorrelated with each other. while always remaining at the global minimum in the subspace, defined by the component amplitudes, of the full parameter space.

Finally we can also improve our initial guess and make the initial amplitudes and indices of the model compatible with each other by solving the least-squares problem in the subspace of amplitudes (x') but keeping the initial guess for the indices.

(0') After step (0) solve the least squares problem for the amplitudes (x') similar to step (3) and update the amplitudes before proceeding to step (1).

It turns out that the combination of the linearized model, Eq. 2, combined with additional sub-iteration step (3) makes the algorithm extremely efficient in finding the minimum of the χ^2. We have in fact a significant advantage over a general non-linear problem. The linear model, Eq. 2, is a good approximation over almost the entire range of allowed parameter space since the spectral indices cannot vary by a large amount and therefore our minimum lies within the range of validity of the expansion. This is not true in general and the Taylor expansion at linear or quadratic order for a general non-linear problem would probably be valid in a small region of parameter space which may not include the minimum we are after. It is of course very difficult to prove that we have indeed
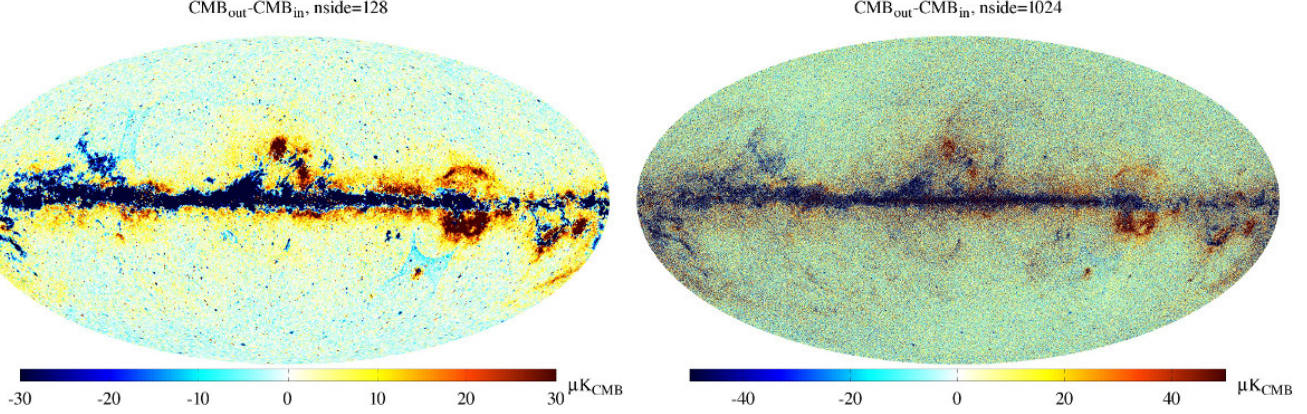
[2] http://healpix.sourceforge.net

Figure 1. CMB Residuals for the FFP6 simulation, (CMB$_{out}$-CMB$_{in}$). The residuals are less than few μK at high latitudes at nside=128. There is also no visible bias at high latitudes and the residuals are consistent with noise. Note that our color scale is different from that used by (Planck Collaboration et al. 2013b). A monopole and dipole calculated at latitudes $|b| > 30°$ has been subtracted from the residual maps.

found the global minimum and the only way to test the quality of results is to use simulations which we do in section 3.

2.1 Additional constraints and exit criterion

We need to add additional constraints to the above algorithm motivated by the physical nature of the foregrounds. In other words we want to solve the least-squares problem subject to certain constraints so that we remain in the physically relevant parameter space in the presence of noise.

(i) We want to restrict the allowed range of the spectral indices β_{lf}, β_d. We impose a somewhat loose bound in the regions where foregrounds have high amplitude and a tighter bound in the regions with low foregrounds. This is an attempt to emulate a prior in the Bayesian sense. If a foreground component is detected with high signal to noise the bounds have no effect and we can allow a wider freedom for the spectral index to vary. The bounds we use are as follows:

$$d_1 - d_2 > 10^{-3}\ \mathrm{K_{CMB}} \rightarrow -5 \leqslant \beta_{lf} \leqslant -1$$
$$d_1 - d_2 \leqslant 10^{-3}\ \mathrm{K_{CMB}} \rightarrow -3.5 \leqslant \beta_{lf} \leqslant -2.5$$
$$d_7 - d_6 > 10^{-2}\ \mathrm{K_{CMB}} \rightarrow 1 \leqslant \beta_{lf} \leqslant 5$$
$$d_7 - d_6 \leqslant 10^{-2}\ \mathrm{K_{CMB}} \rightarrow 2 \leqslant \beta_{lf} \leqslant 3 \qquad (4)$$

It may happen, usually where foregrounds have too small S/N, that there is no local minimum in the direction of spectral indices, where the derivative of χ^2 vanishes, within the bounds. We will still have a minimum value for the χ^2 within the bounds which we can use as the best value of our parameters usually at the one of the boundaries. In that case it is not possible to use the Fisher matrix to estimate the uncertainties and in particular it will in general not be positive definite. In such situations, since foregrounds are anyway small and therefore their influence on the errors on the CMB component should be minimal, we calculate the covariance after removing the columns/rows corresponding to the non-linear parameter (spectral index) which hit the boundary or which is causing the Fisher matrix to be non-positive definite.

(ii) In the regions of low or non-existent foregrounds there are less components in reality than present in the model. Therefore we will be fitting most of the parameters to noise. Also the lowest frequency channels in Planck have much higher noise and poorer resolution and they may pull down the S/N of the final component maps. To

avoid this we explicitly fit the low frequency foregrounds only to the lowest 4 frequency channels and dust to the highest 4 frequency channels. We do this by setting following conditions

$$d_1 - d_2 \leqslant 10^{-3}\ \mathrm{K_{CMB}} \rightarrow f_\nu^{lf}|_{\nu \geqslant 143\ \mathrm{GHz}} = 0$$
$$d_7 - d_6 \leqslant 10^{-2}\ \mathrm{K_{CMB}} \rightarrow f_\nu^{d}|_{\nu \leqslant 70\ \mathrm{GHz}} = 0 \qquad (5)$$

(iii) It is possible that the amplitudes would tend to go negative during the iterations. We take this also as a sign that the foregrounds are negligible and are just being fitted to noise. When this happens for the dust and low frequency amplitudes, we multiply the existing amplitude by a factor of 0.1 instead of updating to negative value. For the CO amplitude, if the CO contribution to a channel signal goes below 1% of noise level in all of the three channels in which the CO contribution is non-zero, we remove CO as a foregrounds component for the next 5 iterations. If the final parameters with minimum χ^2 were fitted without CO then we remove the corresponding columns/rows also from the Fisher matrix.

Finally we have the following exit criterion out of the iteration loop:

(i) Exit if $\chi^2 < 0.1$ or change in χ^2 in the previous 2 iterations is less than 10^{-3}.
(ii) The number of iterations has exceeded 100.

On exit the output parameters correspond to the iteration which had the minimum χ^2 among all the iterations. This is necessary since as we mentioned earlier, the global minimum of the non-linear problem is not necessarily the physical solution that we desire if it lies outside the physical constraints imposed by us. Within the physical constraints, the minimum χ^2 encountered may be a saddle point or may even have non-vanishing first derivative, therefore the final converged value of the χ^2 may not be a minimum but an asymptote flying off to outside the allowed parameter region or even to infinity. Our prescription gives us the best values of parameters within the physical constraints imposed by us.

3 VALIDATION OF THE ALGORITHM ON THE PLANCK SKY MODEL FFP6 SIMULATIONS

We know apply our algorithm to the publicly available FFP6 simulations of the Planck sky model (PSM) (Delabrouille et al. 2013). In particular we use the nominal survey full sky signal maps. These
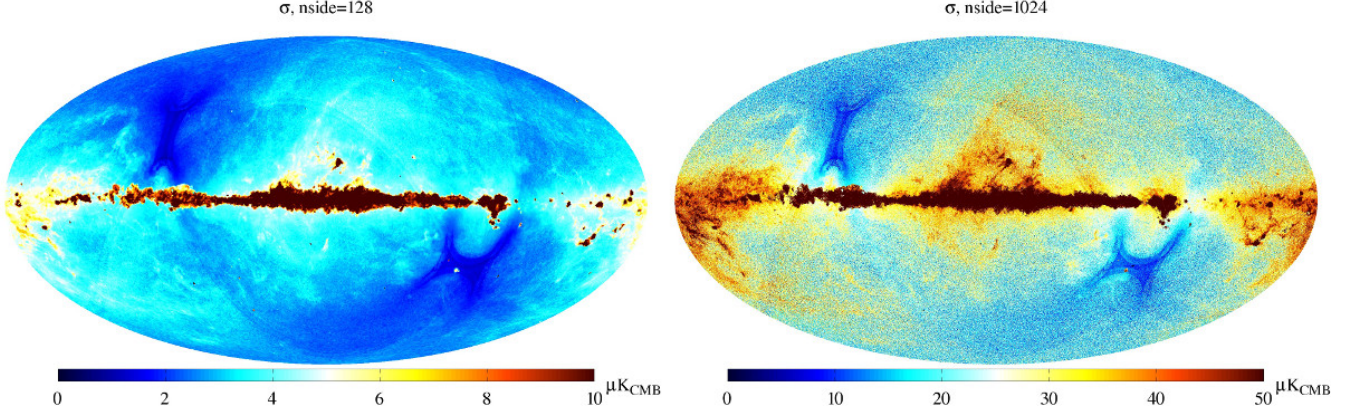
Figure 2. The standard deviation marginalised over the foreground components. The errors roughly follow the pattern of foregrounds on the sky correctly modelling the uncertainties in foreground subtraction.
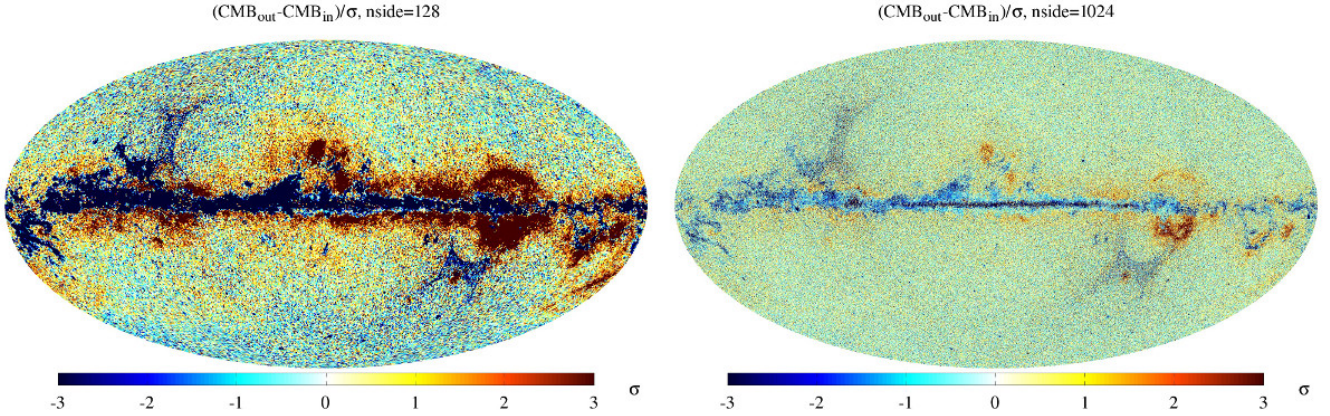

Figure 3. The normalized error defined as ratio of residuals to the standard deviation estimate. If the error estimates are correct then this map should resemble Gaussian noise. The features at the Ecliptic poles are better visible in this map owing to extremely low noise and shows that there are systematic errors not included in our error estimates.

maps include estimates of the variance in each pixel which we use as estimates of the noise variance in our least squares fitting. The low frequency instrument (LFI) maps in the lowest three channels are at lower resolution of HEALPix nside 1024 compared to the nside 2048 for the high frequency instrument (HFI). We upgrade the LFI maps to 2048 also scaling the variance. Since the LFI maps are anyway more noisy compared to the HFI maps, this upgradation does not affect the CMB or the high frequency components. To get the best low frequency components the component separation can be repeated by degrading all maps to, for example, nside=512.

Since Planck measures only the change in signal across the sky and not the absolute sky brightness, there are relative offsets between different channels. Therefore estimates of monopole and dipole should be subtracted from the maps to prevent systematic errors. In the FFP6 simulations we find that the offsets are small enough that they do not matter. For the actual Planck data we will however subtract the best fit values of the monopoles and dipoles provided by Planck Collaboration et al. (2013b).

The residual map, i.e. difference (CMB$_{out}$-CMB$_{in}$) between the output from LIL (CMB$_{out}$) and the input CMB map in FFP6 simulation (CMB$_{in}$) is shown in Figure 1. We show the differences for two HEALPix resolution at nside 1024 and 128. The residual at nside 128 can be compared to that from the other algorithms used by the Planck collaboration (Fig. 7 in Planck Collaboration et al.

(2013b)). The residuals are about the same level as other algorithms and of order of few μK at high latitudes. There is a systematic feature around the Ecliptic poles where because of the Planck scanning strategy the depth of the survey and the noise levels change abruptly. We discuss this and the small systematic bias near the galactic plane where the residuals are preferentially positive in detail section 4. We note that similar systematic effects and biases are also present Planck analysis of FFP6 simulations (Planck Collaboration et al. 2013b). In particular the feature near the Ecliptic poles can be seen in the COMMANDER residual maps.

3.1 Error estimation

We estimate errors on the final components using Fisher matrix. As mentioner earlier, for pixels to which the model without the CO component was fitted we delete the corresponding columns/rows from the Fisher matrix. Also for pixels which did not converge to a minimum in the direction of one or both of the spectral indices those columns/rows are also deleted. We will see that this prescription gives us a reasonable estimate for the errors. The Fisher matrix is calculated at full resolution, nside=2048. For lower resolutions we combine the errors in the smaller pixels in each larger pixel assuming that the errors/noise are uncorrelated by first degrading the variance map to new nside, which averages the variance, and
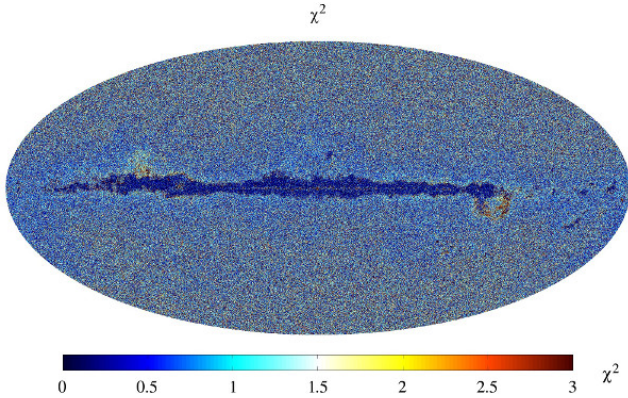
χ^2

Figure 4. The χ^2 map for the least squares fit performed by LIL. For 1 or 2 degrees of freedom that we have with or without the CO component, the average χ^2 is expected to be between 1 and 2.
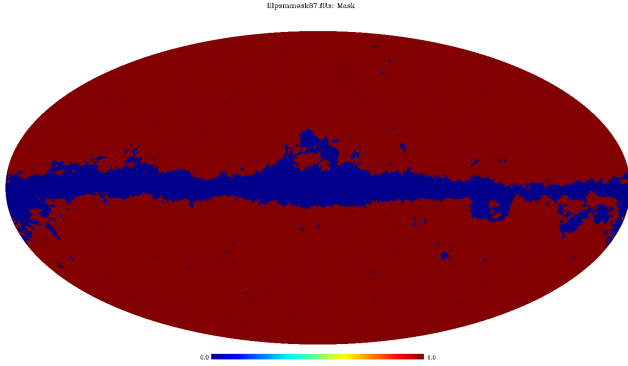
lilpsrmmask87.fits: Mask

Figure 5. The mask used in further validation of the results from FFP6 simulation yielding a sky fraction usable for CMB of 86.4%.

then dividing the variance in each pixel by $\text{npix}_{2048}/\text{npix}_{\text{nside}} = 2048^2/\text{nside}^2$, where $\text{npix}_{\text{nside}} \propto \text{nside}^2$ is the total number of pixels in the map. This is an approximation since the beams of all frequency channels are larger than the size of any pixel. We show the standard deviation error and normalized error map in Figures 2 and 3. The normalized error map $((\text{CMB}_{\text{out}}-\text{CMB}_{\text{in}})/\sigma)$ is just the ratio of the residual map to the standard deviation $(\sigma(p))$ map.

The standard deviation σ roughly follows the distribution of foregrounds on the sky. This is because there are uncertainties associated with the foreground subtraction and which we marginalise over. The normalized error is an estimate of how good and compatible our CMB signal and error estimates are. For perfect reconstruction of the mean signal and errors this map should resemble Gaussian noise. We see that this is so except in the galactic plane where our simple model does not capture all the complexities of the foregrounds. We have also neglected the noise correlations between pixels and therefore in the degraded maps the noise is slightly underestimated. In particular the features near the Ecliptic poles are more prominent in the normalized errors map indicating that there are systematic uncertainties in these extremely low noise regions which are not included in our error estimates. The goodness of fit can be quantified by the χ^2, and the χ^2 map is shown in Fig. 4. This can be compared with the χ^2 map of COMMANDER (Planck Collaboration et al. 2013b). One difference is that we get low χ^2 values in the galactic plane in contrast to COMMANDER, implying that our

model is a good fit to the data in the galactic plane. This difference from COMMANDER could be because of the tight prior they impose on the spectral indices while we allow a wider range of spectral indices in the high foreground regions.

3.2 Mask

To better analyze the quality of our component separation for the CMB as well as the foreground components we need to mask out the worst regions of the sky. We use the dust amplitude at 353 GHz, low frequency amplitude at 30 GHz and the standard deviation maps as estimated by LIL to construct the mask. We smooth the maps with a 30 arcmin full width half maximum (FWHM) Gaussian beam and threshold the dust amplitude at 2 MJ/Sr, low frequency amplitude at 600 μK and standard deviation estimate at 75 μK resulting in masking 13.6% of the sky. We also mask pixels with the $\chi^2 \geqslant 10$. This scheme also masks the brightest point sources as can be seen in Fig. 5 and we do not use any additional point source mask for the results from FFP6 simulations.

3.3 Validation

If our model was a good description of the Planck sky then the χ^2 will follow the χ^2-distribution with 1 or 2 degrees for freedom for the pixels with and without the CO components respectively. The probability density function of χ^2 is shown in Fig. 6 for the pixels fitted with and without the CO component and compared respectively along with the theoretically expected χ^2-distributions with 1 and 2 degrees of freedom labeled $\chi^2(1), \chi^2(2)$. LIL distributions match the theoretical curves quite well despite the fact that our foreground model is much simpler compared to the actual foregrounds. The deviations we see are similar in nature to those observed in the Bayesian codes (Eriksen et al. 2006).

The χ^2 on average is a little larger than the theoretical values. We should however expect this since not all the components fitted in our model are present everywhere in the sky. The obvious example is the CO component which is present at detectable levels only near the galactic plane. Even the low frequency and dust components are not present everywhere on the sky. The low frequency components are below the noise levels in a major fraction of the sky. The dust component also is very weak in the 217 GHz and lower frequency channels in a good fraction of the sky and so the spectral index of even the dust component cannot always be determined with any accuracy. We are therefore fitting a model with more parameters than are required by the data and this shows up in the higher values of χ^2. Our hypothesis is supported by the fact that the χ^2 distribution inside the mask for pixels where there is detectable CO is much closer to the theoretical curve. This reasoning also explains why the algorithms such as SMICA, NILC and SEVEM, which have an equivalent implicit model with smaller number of parameters, do so well. The above arguments also suggest that fitting the most general model to the data may not be the optimal approach. What we should really do it fit many models to the data and choose from the models the one which fits best given the number of parameter and the degrees of freedom. We will do exactly this for the CO component in section 4.

We show in Fig. 7 the probability distribution (PDF) of normalized error for the CMB for different resolutions. At high resolution the PDF is close to a Gaussian. The deviations from the Gaussian are of similar magnitude to that of (Eriksen et al. 2006). The reason for the under-estimation of errors as we decrease the resolution is because we have calculated the errors ignoring the noise
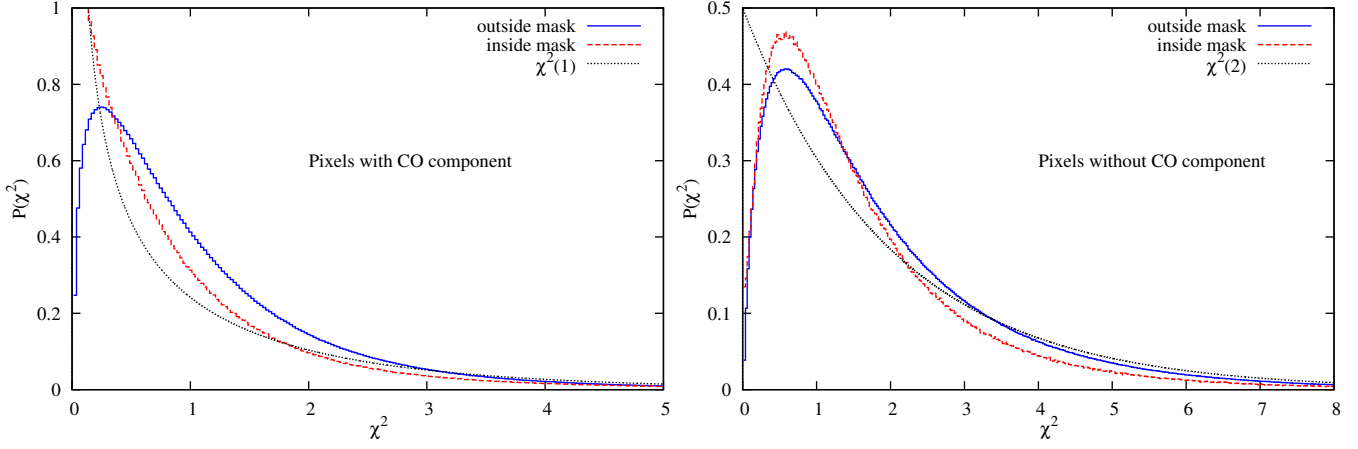
Figure 6. The χ^2 probability density function obtained by LIL and compared with the theoretically expected χ^2 distributions with 1 and 2 degrees of freedom. The LIL distribution matches well with the theoretical curves for large χ^2 values and inside the mask in the left figure where all the modelled components are present.
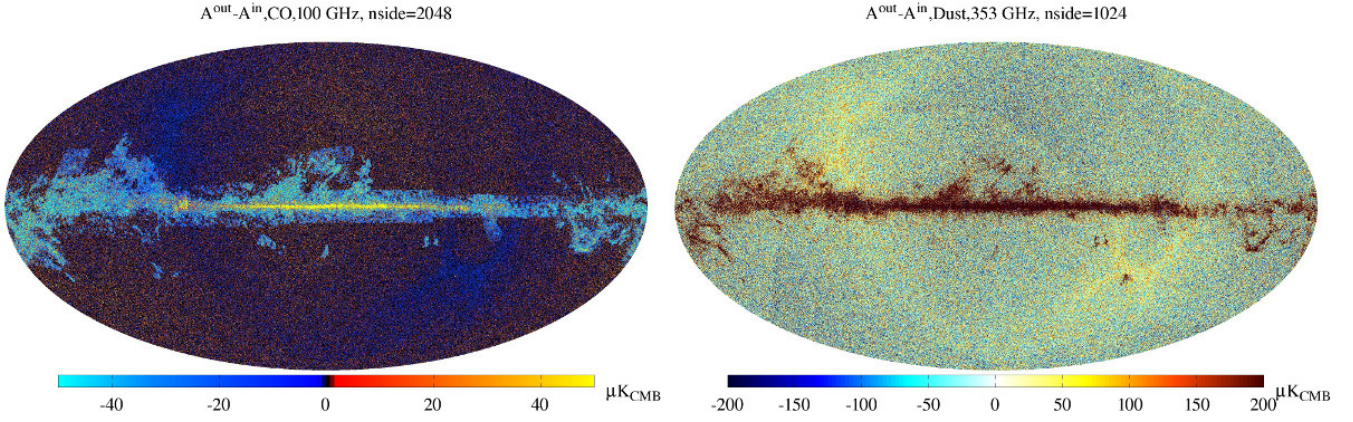

Figure 8. Residuals in the CO and dust components. The CO residual away from the galactic plane follows the hit count map decided by the Planck scanning strategy. A monopole and dipole calculated at latitudes $|b| > 30°$ has been subtracted from the residual maps. Note that we are using a different color scheme for the CO to better highlight the non-zero pixels.
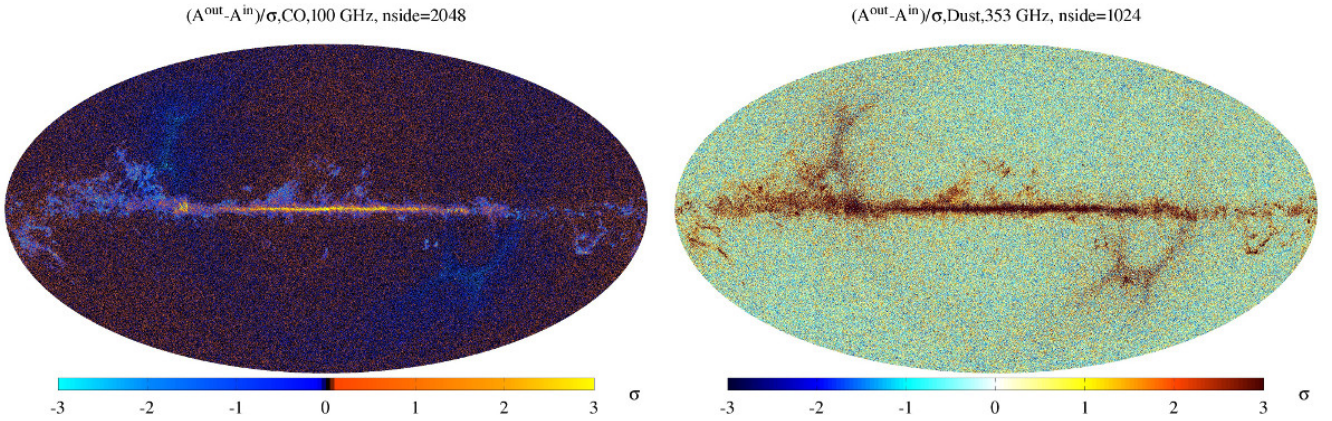
Figure 9. The normalized errors in the CO and dust components. The difference in the CO plot from Fig. 8 around the Galactic plane is because in this plot we only include pixels for which we fitted for the CO component while Fig. 8 includes all Pixels. In particular CO in the rectangular strips present in the FFP6 simulation is not detected as the signal is too low.
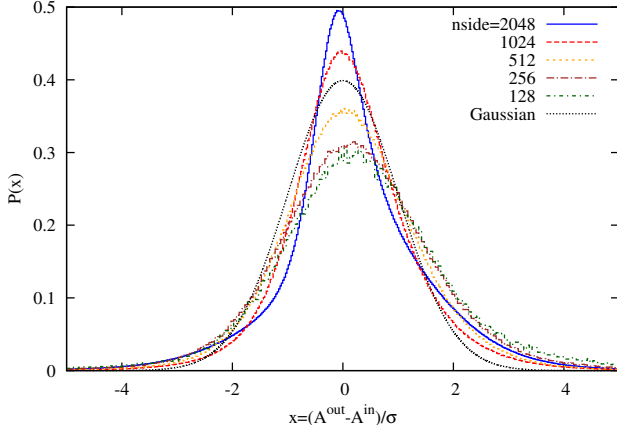
Figure 7. The probability density function of normalized error (Fig. 3) $(A_{CMB}^{in} - A_{CMB}^{out})/\sigma_{CMB}$ for different resolutions. For high resolution the curves are close to the expected Gaussian but the error is underestimated as we go to the lower resolutions because we ignored the correlations in the noise. We however note that the distribution approaches a Gaussian as the average pixel size approaches the resolution of the highest frequency channels. For the foregrounds results therefore we will use nside=1024 for dust and nside=512 for the low frequency component.

correlations between pixels. The average size of the pixels is closer to the actual resolution of the Planck HFI channels for nside = 512 and 1024 and these distributions are closest to the Gaussian. This fact gives us some confidence that we are recovering correctly the best fit values as well as the error bars.

We show in Figs. 8 and 9 the residuals and the normalized error for the two high frequency foreground components, the CO and the dust amplitudes, at the corresponding reference frequencies. Figure 10 shows the same quantities for the low frequency component of our foreground model. Away from the Galactic plane, the residuals are consistent with noise. The CO residuals follow the Planck hit count map and is consistent with the CO component just fitting noise away from the galactic plane. The true CO emission in the FFP6 simulation is zero away from the galactic plane. Note that a large fraction of the pixels are fitted without the CO component (54% pixels outside the mask). These pixels have the CO amplitude set to zero and do not have an error on the CO amplitude. For the CO residual map we have used all pixels while for the CO normalized error map only the pixels with the CO component are non-zero. This accounts for the difference in structure of the two maps since the residual map includes differences for pixels which had a non-zero value in the original FFP6 simulation but are not detectable by LIL in the Planck data. These pixels are mostly in form of the rectangular bands around the Galactic plane in the residual map, Fig. 8. We have also used a modified color scheme to better highlight the non-zero pixels against the background of pixels with CO amplitude fixed to zero. Also since most pixels with the CO component away from the galactic plane are surrounded by the pixels without the CO component, there is no consistent way to degrade the resolution of the error map and for the CO component we present results at the full resolution, nside=2048.

Figure 11 shows the PDF of normalized errors on the foreground amplitudes within and outside our mask. Outside the mask the PDF is close to Gaussian and we do not see as significant an underestimation of errors for the dust component as reported in Planck Collaboration et al. (2013b). Note that our plots are at a higher resolution compared to Planck Collaboration et al. (2013b),

nside=1024 for the high frequency components and nside=512 for the low frequency component. Inside the mask the deviations from the Gaussian are more significant indicating more complex foregrounds compared to what we have modelled.

Finally to validate the full foreground model including the spectral indices, we show the residuals for sum of all foregrounds in the 7 Planck frequency channels used in our analysis in Figs. A2 and A1 in the Appendix. The foreground residuals are very low for the 70 GHz channel of Planck Low Frequency Instrument (LFI) and 100 GHz, 143 GHz and 217 GHz channels of the Planck High Frequency Instrument (HFI) thus affirming the accuracy of spectral indices inferred by LIL.

4 LIL-MS: PARAMETER FITTING WITH MODEL SELECTION

We have so far emulated the approach of COMMANDER and shown that our results are consistent. We can improve over this approach of fitting parameters to a given model. One obvious improvement is suggested by looking at Figs. 8 and 9 for the CO component. The CO component is detectable in the Planck data only in the galactic plane and small regions around it, which show up as small scale features in the maps. The large scale features are identical to the Planck hit count maps and are the result of fitting the CO component to noise in most of the sky where no CO is present. In LIL algorithm we did set the CO component to zero in 51% of all pixels (54% if only considering pixels outside the mask) where the amplitude tended to go negative. These pixels are evenly distributed over the sky and there are still enough pixels with the CO component all over the sky that this does not help with the large scale systematics.

Since the CO is present only in the high foreground regions which are anyway masked for CMB analysis, the simplest solution to get the best CMB maps would be to ignore the CO component altogether. This would give incorrect results where the CO component is present but would not affect the cosmology if these regions are masked. A more sophisticated approach, which does not assume that CO is present only in the high foreground regions, would be to fit models both with and without the CO component and select the model which fits best given the number of parameters and degrees of freedom. In the Bayesian approach this amounts to the comparison of Bayesian evidence. For the least-squares parameter fitting the equivalent is the comparison of the χ^2 of the two models.

We therefore run LIL with and without the CO component. The difference in χ^2 between the two models again has a χ^2 distribution with the degree of freedom to one (i.e. difference in the degrees of freedom between the two models) (Stuart et al. 2004, see e.g.). We therefore accept the model with the CO component only if this model gives an improvement of $\Delta\chi^2 \geqslant 2.7$ over the model without the CO component. This corresponds to a 10% probability that we will accept a model with the CO component when there is no CO component present. When following this criteria the CO component is now absent in 79.6% of the pixels (84% if only considering pixels outside the mask) which is a considerable improvement over the single model approach. We will refer to this approach as LIL-MS from now on. There is a small improvement in the χ^2 distributions in LIL-MS compared to LIL (see section 5 for a plot of LIL-MS χ^2 distributions).

The CMB residuals are shown in Fig. 12 and the corresponding normalized errors in Fig. 13. The Ecliptic systematic features are invisible in the residual maps and are only visible in the normalized error map because the noise is underestimated in these ex-
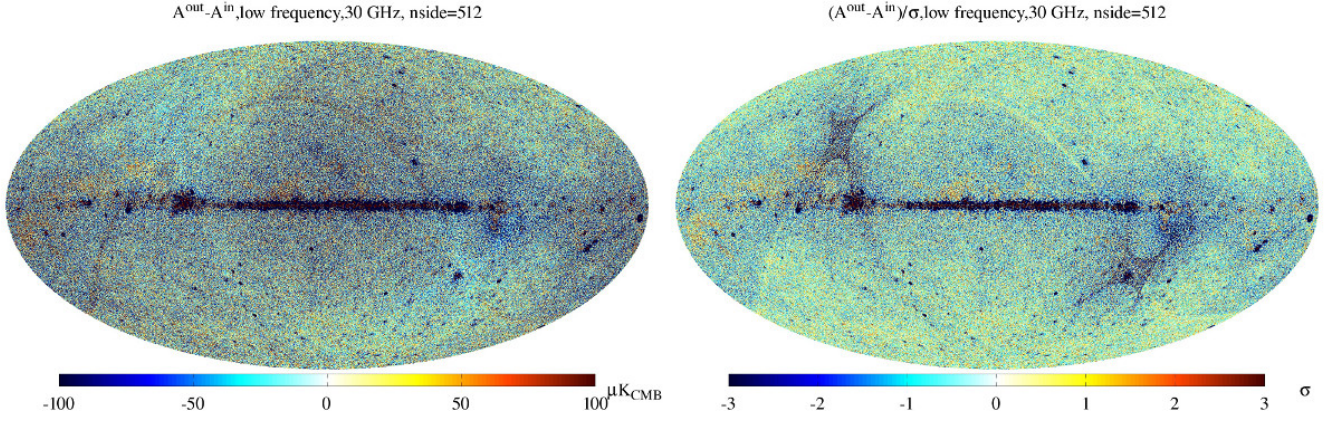
Figure 10. The residual for the low frequency component at 30 GHz and the normalized error. Away from the galactic plane the residuals and errors are consistent with Gaussian noise.
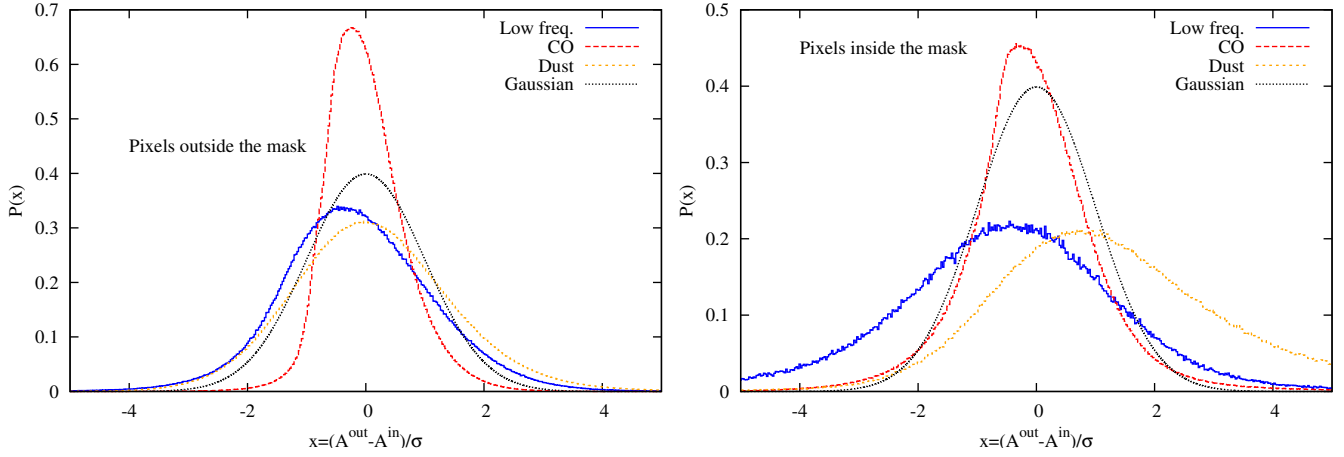

Figure 11. The probability density function of normalized error (Figs. 9 and 10) $(A_{in} - A_{out})/\sigma$ for different foreground component amplitudes in our foreground model. Outside the mask, the distributions are closed to Gaussian except for the CO for which there is a significant overestimation of error. Note that the CO plots are for pixels at nside=2048 while the plots for dust and low frequency components were made after degrading the residual and error maps to nside=1024 and 512 respectively.
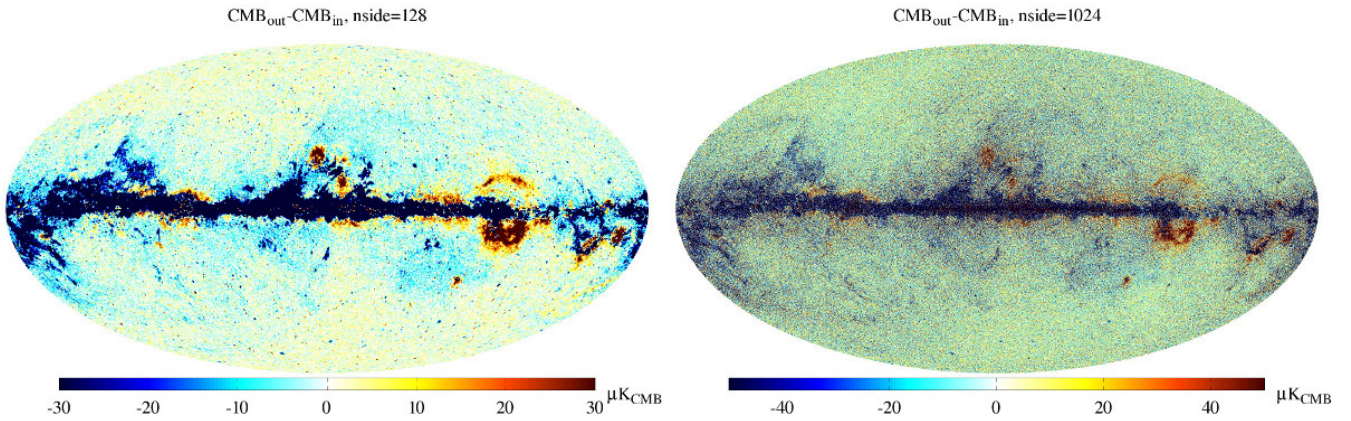
Figure 12. CMB Residuals for the FFP6 simulation, (CMB$_{out}$-CMB$_{in}$) for LIL-MS. There is marked improvement compared to LIL near the galactic plane and the Ecliptic features are also weaker and less sharp.
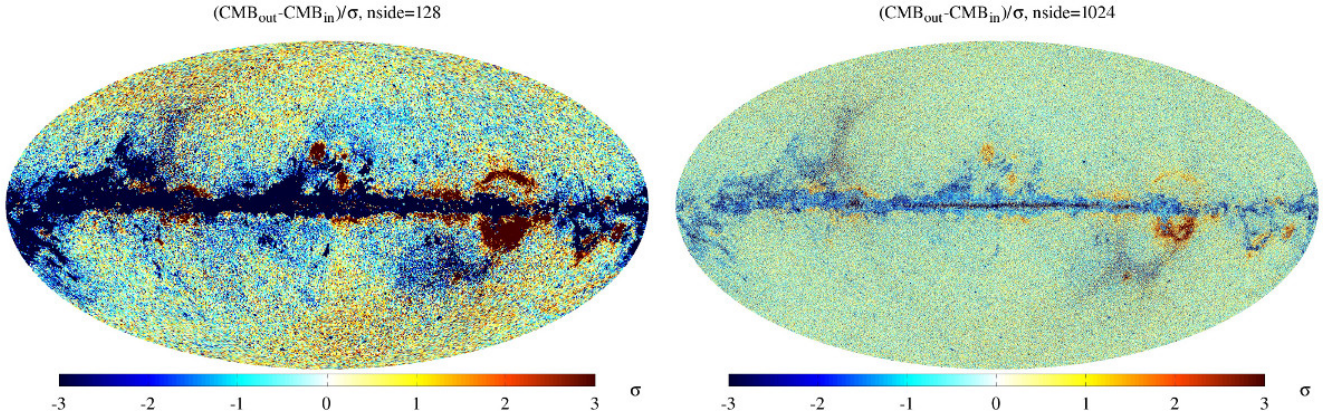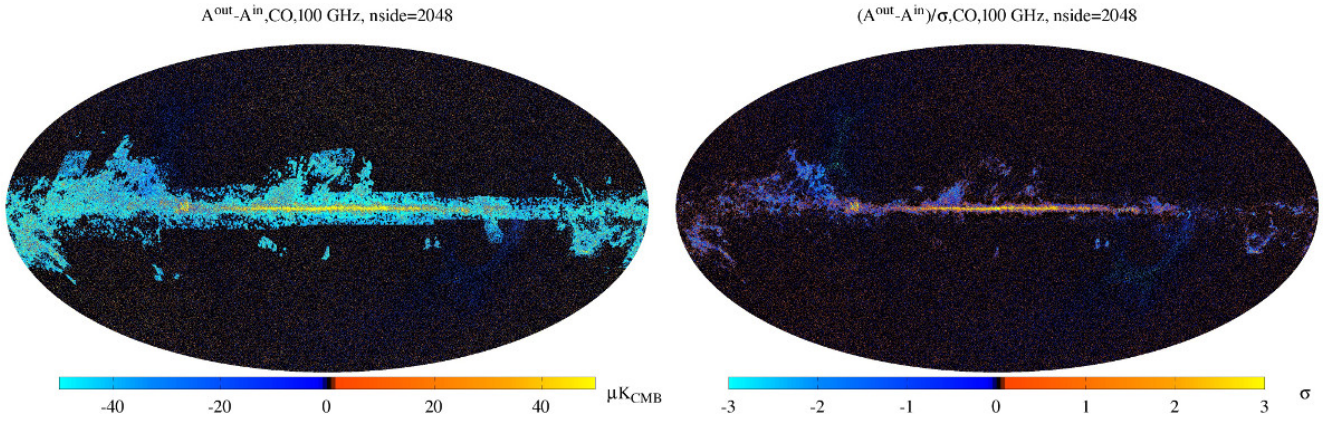
Figure 13. The normalized error for LIL-MS.

Figure 14. The residuals and normalized error for the CO component in LIL-MS. There is a marked improvement outside the Galaxy in the residuals compared to Figs. 8 and 9.
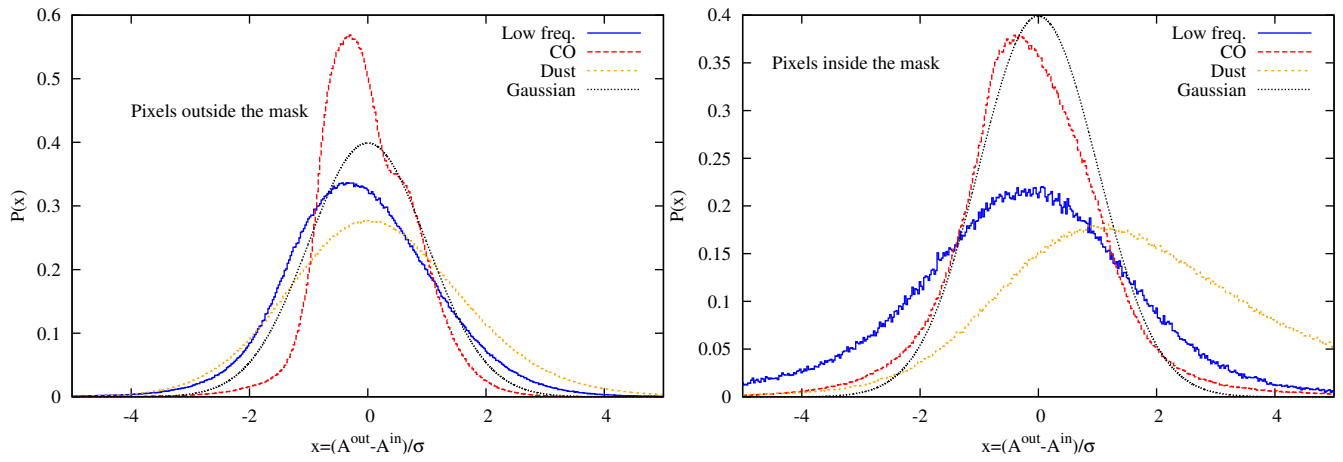

Figure 15. The probability density function of normalized error $(A_{in} - A_{out})/\sigma$ for different foreground component amplitudes for LIL-MS. Outside the mask, the distributions are closed to Gaussian except for the CO for which there is a significant overestimation of error. There is however an improvement compared to the LIL. Note that the CO plots are for pixels at nside=2048 while the plots for dust and low frequency components were made after degrading the residual and error maps to nside=1024 and 512 respectively.
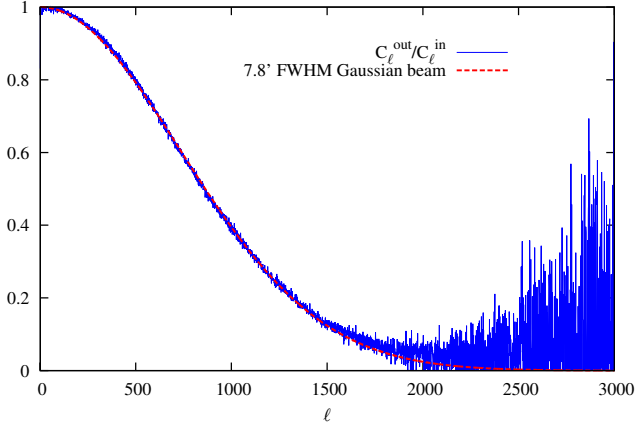
Figure 16. The effective beam of LIL-MS CMB map as estimated by the ratio of output to input angular power spectra corrected for the effect of mask and noise is close to Gaussian with FWHM of 7.8′.
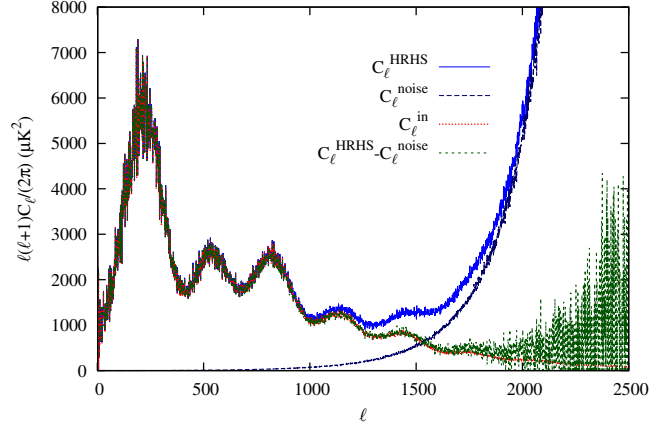

Figure 17. The angular power spectra corrected for the effect of mask and beam. Compared to similar figure in Planck Collaboration et al. (2013b) for Commander-Ruler our noise is slight larger because of the non-linearity of our model even on small scales.

tremely low noise regions. The systematic bias towards the positive residuals around the galactic plane has also gone away and the residuals everywhere except the highest foreground regions, which would be masked, are consistent with Gaussian noise.

The residuals and standard deviation for the CO component is shown in Fig. 14. There is a marked improvement compared to LIL. In particular outside the mask/Galaxy 84% of the pixels have no detected CO component which is much closer to reality. We show the PDFs for the foreground amplitudes for LIL-MS in Fig. 15. There is an improvement in the CO component which has become closer to a Gaussian but not much change in the other components. Note that the dust component is very lopsided for the pixels inside the mask because the monopole that was subtracted in doing this calculation was calculated from pixels at latitude $|b| > 30°$. The biases for the dust component are different at high and low latitudes since the low frequency component is very small at $|b| > 30°$ where there is still considerable dust component. This suggests that we should extend the model selection to the low frequency component also. In particular if the model comparison favors absence of low frequency component in part of the sky then we could remove the noisy lowest two LFI channels completely from the fit and fit a 4-parameter model without the LFI component to the remaining five channels. We will leave such an extension of our model selection approach for future work. Since the systematic effect we have removed is quite small, there is not much change in the PDFs for the CMB component and we do not show them again.

4.1 Resolution and noise in CMB maps

The FFP6 simulations also provide half-ring maps in addition to the full survey maps and we use these to estimate the resolution and noise in our CMB maps. We have processed the half-ring frequency maps with LIL-MS which produces the corresponding two CMB half-ring maps. The half the difference between the two CMB half-ring maps (HRHD) then gives an estimate of the noise in the half ring half sum (HRHS) or the average of the two CMB half-ring maps. To get an estimate of the resolution, we calculate the pseudo-power spectrum ($\hat{C}_\ell$) of the HRHD and HRHS maps on masked sky. We create the mask as described in section 3.2 but lower the thresholds so that 30% of the sky is masked and also apodize the mask with a 30′ Gaussian beam. We then deconvolve the mask by solving the linear convolution equation for C_ℓ (see Hivon et al. 2002,

for a derivation)

$$\hat{C}_{\ell_1} = \sum_{\ell_2} M_{\ell_1 \ell_2} C_{\ell_2} \qquad (6)$$

where C_{ℓ_2} is the full sky power spectrum and

$$M_{\ell_1 \ell_2} = \frac{2\ell_2 + 1}{4\pi} \sum_{\ell_3} (2\ell_3 + 1) W_{\ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2, \qquad (7)$$

W_ℓ is the power spectrum of the mask and the term in brackets is the Wigner-3j symbol. This equation is strictly applicable to only ensemble averages but it is good enough for our purpose to test the quality of our algorithm. Subtracting the HRHD or noise power spectrum C_ℓ^{noise} from HRHS power spectrum C_ℓ^{HRHS} gives an estimate of the CMB power spectrum which we call C_ℓ^{out}. We apply the same mask and deconvolution procedure to the input CMB map. The input CMB map for the FFP6 simulation was made with a 4′ Gaussian beam with spherical harmonic transform $b(l)$. We therefore divide the power spectrum of the input CMB map by $b(l)^2$ to get the input CMB power spectrum C_ℓ^{in}. The ratio of the output and the input power spectrum then gives an estimate of the beam function of our CMB map. This ratio is shown in Fig. 16 and is well approximated by a Gaussian beam of FWHM 7.8′. This is close to the resolution achieved by Commander-Ruler of 7.4′ (Planck Collaboration et al. 2013b). We plot the input, HRHS and noise power spectrum corrected for the effect of the respective beams in Fig. 17. The noise in our CMB map is a little worse compared to Commander-Ruler. The reason for it is probably the fact that we did not change the resolution of all frequency maps to a common one. Therefore the high noise low resolution channels get higher weight than they are entitled to during the least squares fitting, i.e. the relative noise in the low resolution channels is underestimated compared to the higher resolution channels. This effect is aggravated for us compared to Commander-Ruler since we fit a non-linear model at full resolution while RULER fits a linear model. We leave detailed investigation and possible improvements to mitigate the noise and improve the resolution for future work.
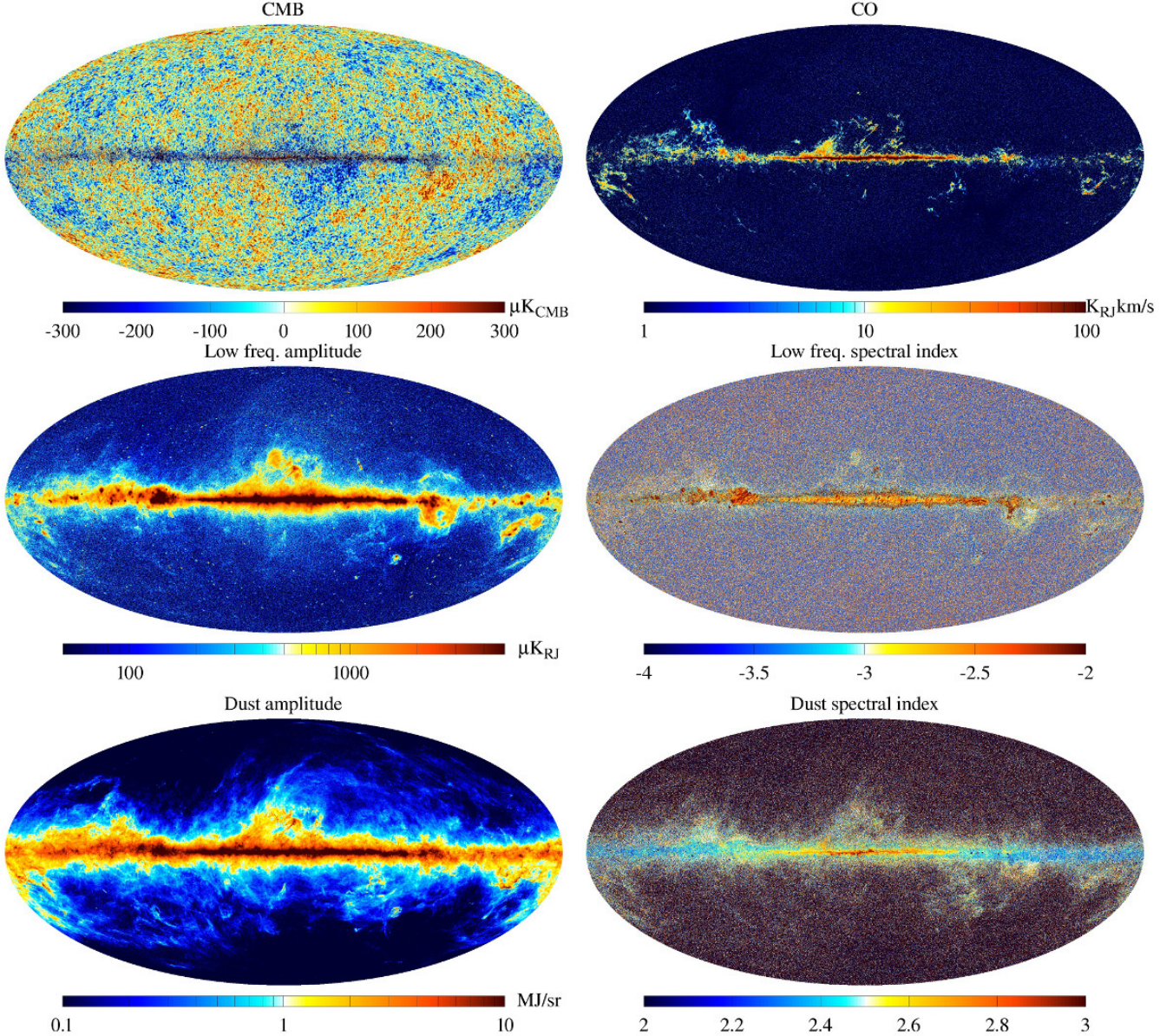
Figure 18. The CMB and foreground parameter maps from LIL-MS applied to Planck data.

5   APPLICATION TO PLANCK DATA AND COMPARISON WITH OFFICIAL PLANCK MAPS

We now apply LIL-MS to the Planck first data release. The maps for the 6 parameters recovered by LIL-MS are shown in Fig. 18. The CO amplitude is negligible away from the galactic plane and model selection plays an important role in recovering the true behavior of CO. In particular the CO component was removed from 84% of the pixels outside a 12.3% mask on the highest foreground regions. The low frequency component parameters are consistent with those from Commander. The low frequency amplitude is mostly determined by 30 GHz channel which has high signal to noise and this is apparent in the amplitude map. To constrain the spectral index we need high signal to noise in more than one channel. Since the foreground amplitude drops quite a bit from 30 GHz to 44 GHz, the spectral index is well constrained in much smaller region and there-

fore looks much noisier at high galactic latitudes. The dust amplitude is very well constrained and fine features such as streams and filaments are well recovered over most of the sky. Same constraints as low frequency index apply also to the dust spectral index and it is well constrained in regions of high amplitude but becomes noisier and unconstrained as we go to the higher latitudes. Our CMB map has some residual small scale noise in a narrow ridge in the galactic plane compared to the CMB maps released by the Planck collaboration. This is probably because we did not adjust the resolution of channel maps to a common resolution before parameter fitting. We will discuss a possible solution to this in the conclusions section.

We show in Fig. 19 the standard deviation estimated using the Fisher matrix. Comparison of these maps gives useful information about the degeneracies between the parameters. The CMB is influenced by both the dust and the low frequency components and against a background that follows the Planck hit count map, the in-
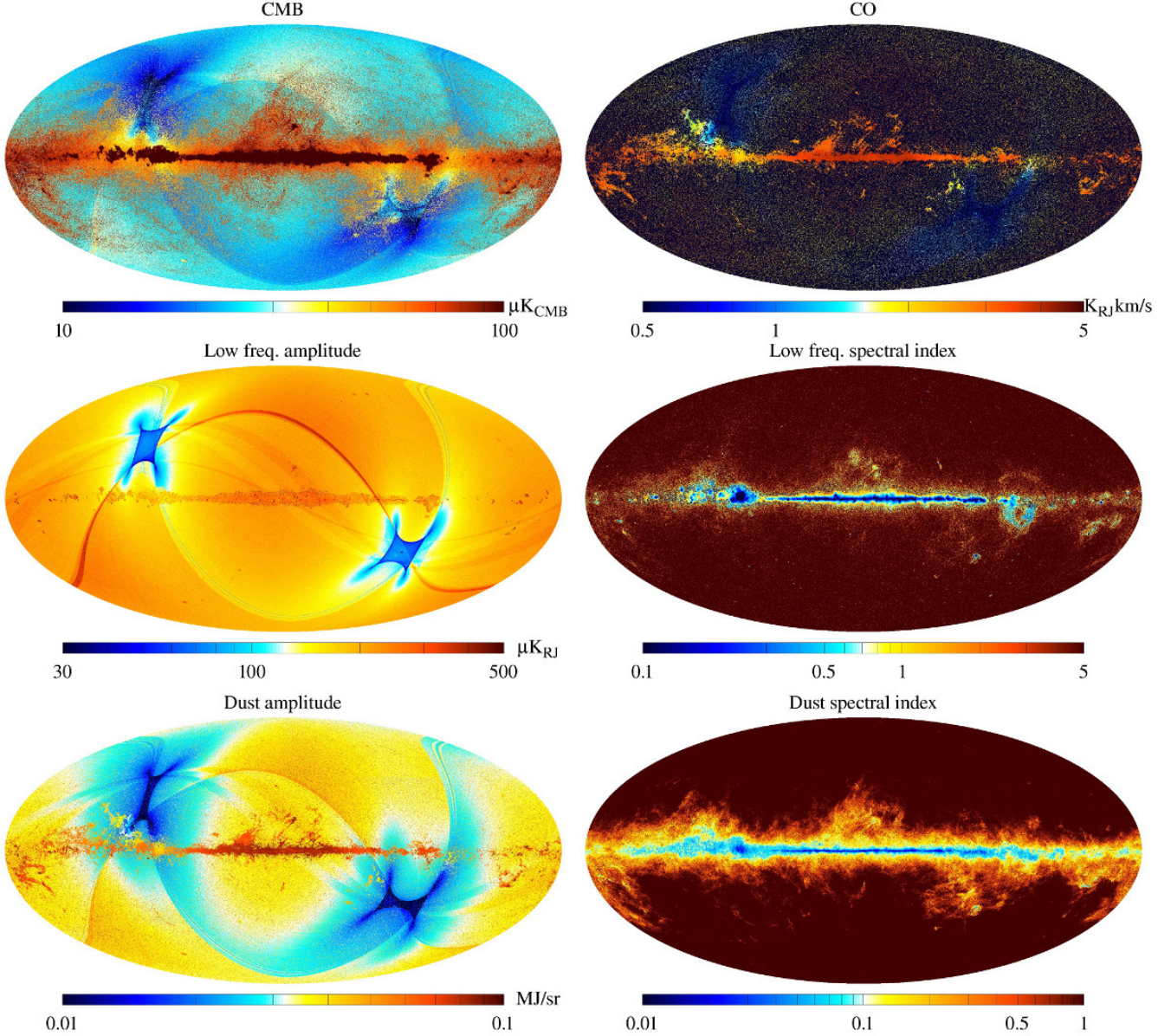
Figure 19. The CMB and foreground parameter standard deviation maps estimated by LIL-MS using Fisher matrix.

fluence of low and high frequency foregrounds is clearly visible. Comparison of CO map with dust map shows that these two are very degenerate while there is not much degeneracy between the low frequency foregrounds with the dust and CO. The spectral index is well constrained in the high signal regions, where there is good signal to noise for the corresponding foreground component in at least two of the channels. The errors on the spectral indices therefore follow the morphology of the respective foreground components.

We show the χ^2 map in Fig. 20 and compare the χ^2 distribution with expected distributions with one or two degrees of freedom in Fig. 21. The tails agree very well for the χ^2 distribution with one degree of freedom for pixels with the CO component selected by our model selection criteria and that of χ^2 distribution with two degrees of freedom for pixels where the CO component is not present.

Overall the distributions are close to what we expected from simulations.

5.1 Comparison with the Planck collaboration results

We show in Fig. 23 the difference between our CMB maps and the CMB maps from different algorithms released by the Planck collaboration (Planck Collaboration et al. 2013b). We also show for comparison the difference between the COMMANDER-RULER(CR) and SMICA/SEVEM maps on the same color scale. The agreement between our maps and other algorithms is as good as the agreement in-between the methods used by the Planck collaboration. All maps are downgraded to HEALPix nside=128 so they can be compared with similar maps in Planck Collaboration et al. (2013b). We also show the comparison between the foreground amplitudes recovered by us and those from the CR in Fig. 22 for CO and Fig. 24
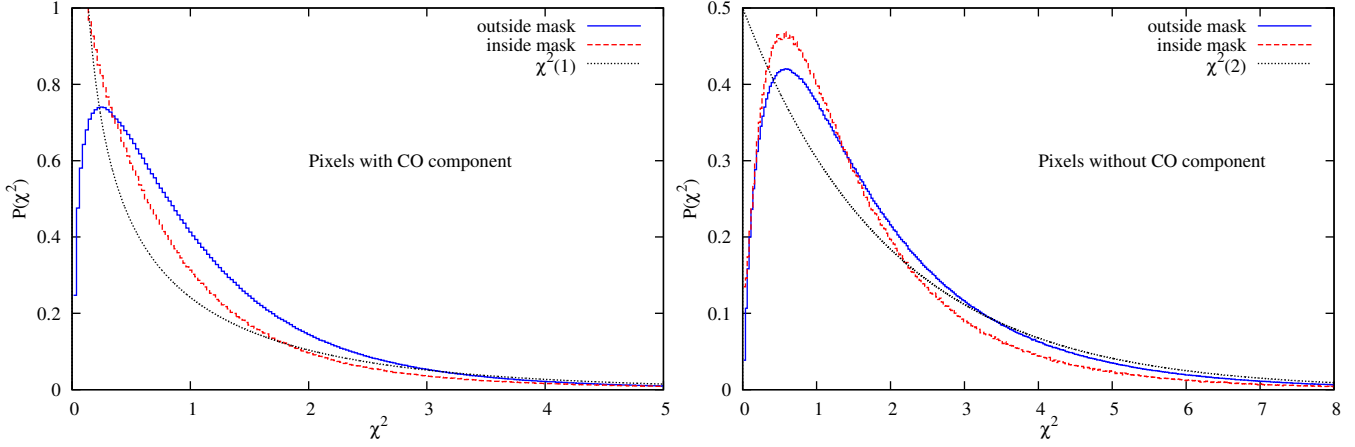
Figure 21. The χ^2 probability density function obtained by LIL-MS for Planck maps and compared with the theoretically expected χ^2 distributions with 1 and 2 degrees of freedom. The distributions are close to what we predicted from the FFP6 simulations. The tails for real data has even better agreement with the theoretical distributions than what we had for simulations.
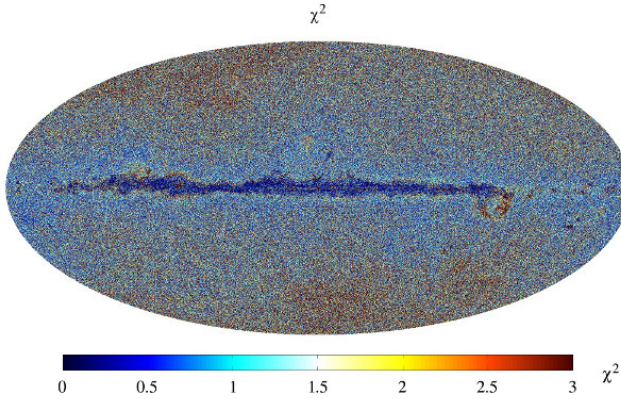

Figure 20. The χ^2 map for the least squares fit performed by LIL-MS. For 1 or 2 degrees of freedom that we have with or without the CO component, the average χ^2 is expected to be between 1 and 2.
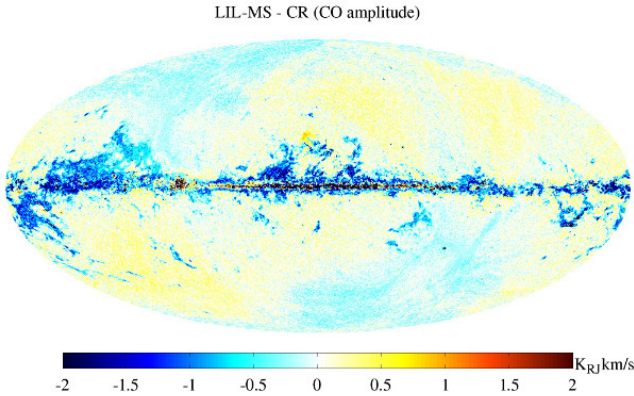

Figure 22. Difference between LIL-MS CO amplitude at 100 GHz and the CR map degraded to nside=128. A monopole and dipole calculated at latitudes $|b| > 30°$ has been subtracted.

for the low frequency and dust amplitudes. The agreement is again quite good away from the galactic plane. It is not possible to directly compare the spectral indices since for COMMANDER algorithm they are calculated at much lower resolution and are mostly driven

by the prior over most of the sky away from the Galactic plane. In the Galactic plane we see that our maps are broadly consistent with those from COMMANDER by comparing our Fig. 18 with the corresponding figure in Planck Collaboration et al. (2013b).

6   CONCLUSIONS

The main aim of this paper is to present a least squares parameter fitting algorithm, LIL, optimized for component separation in the CMB sky. Our algorithm is extremely efficient, fitting 6 parameter model to 7 frequency channels for 50 million pixels in 160 CPU-minutes or a few minutes running in parallel on few tens of cores. We have also argued for an extension to the algorithm, LIL-MS, by including model selection for the components such as CO and perhaps also for the low frequency component, which we know from observations are only present in detectable amount over a fraction of the sky. In particular we have shown that parameter fitting can be done at high resolution and in particular there is no need to smooth all channels to the resolution of the lowest channels. The main aim of developing such a parameter fitting algorithm is to try to get maximum information about the foregrounds and the CMB from data while making least amount of assumptions. In particular our assumptions and models are motivated by the prior knowledge about the foregrounds from other observations. Our method is still not optimal and there are several improvements which can be done based on the results we have obtained so far.

(i) Difference in resolution between different channels can be a source of high frequency noise and we see this in the CMB map. In addition the resolution of our final CMB map is close to that of 143 GHz channel. SMICA and NILC (Planck Collaboration et al. 2013b) achieve a higher resolution by re-beaming all channel maps to 5′. In principle we see no reason why we cannot do that also for LIL-MS. Re-beaming to highest resolution channel will lead to increase in noise in the lower resolution channels. Thus during parameter fitting, the low resolution data will get the correct additional down-weighting and the final maps would be closer to the resolution of the highest resolution channels, especially for the CMB. We expect that this should also take care of the residual high frequency noise we see in our CMB map around the Galactic plane.
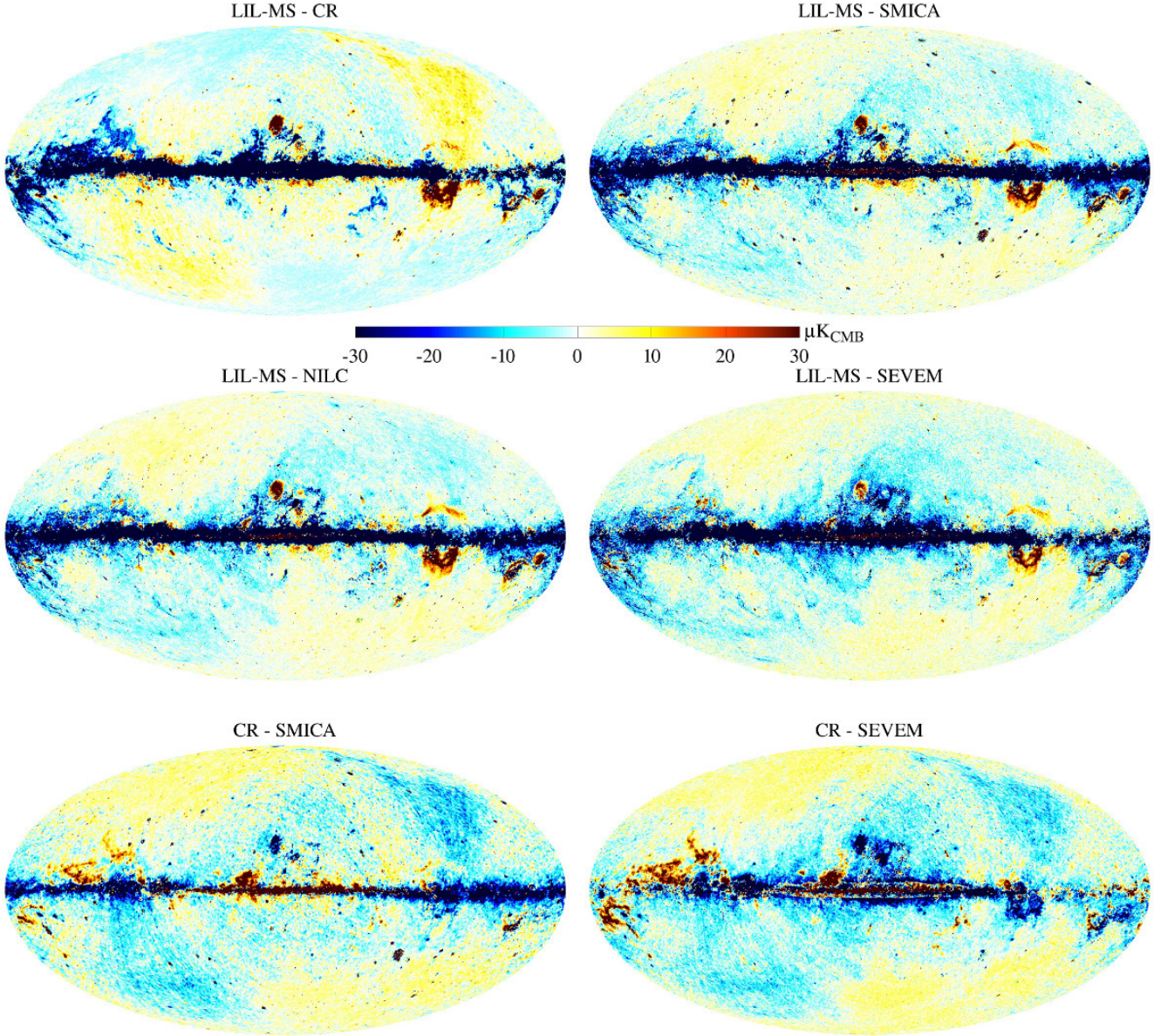
Figure 23. Difference between LIL-MS CMB map and maps released by the Planck collaboration using different algorithms. All maps have been downgraded to nside=128. We also show for comparison the difference between the Commander-RULER and SMICA/SEVEM maps on the same color scale. A monopole and dipole calculated at latitudes $|b| > 30°$ has been subtracted.

(ii) The low frequency component, like the CO component, is important only on a fraction of the sky. This suggests to extend the model selection to include the low frequency component. In particular in parts of the sky where a model with a low frequency component is disfavored, the two lowest frequency LFI channels can simply be omitted from the fit.

(iii) If we are interested in low frequency component, then doing a fit at the nside=2048 does not really make sense. For the low frequency component therefore a dedicated analysis can be done by re-beaming to a lower resolution close to that of the 70 GHz channel.

(iv) We have assumed that the noise between different pixels in uncorrelated at nside=2048. This is not true since for most channels the beam size is much bigger than the pixel size. A better treatment

of noise is therefore needed when doing cosmological analysis. We note that Commander also ignores correlations between the pixels.

We leave study of the above issues as well as cosmological results for a future publication. Nevertheless we have shown that our simple approach to component separation using the least squares fitting works quite well for multi-frequency experiments like Planck. An advantage of parameter fitting over other methods (see however Vansyngel et al. 2014) is that the errors we get take into account the uncertainties in the foregrounds and that it allows for considerable flexibility to the foregrounds to vary over the sky. We have shown our results to be consistent with the published results by the Planck collaboration. We hope to overcome the shortcomings in our algorithm concerning the resolution and noise compared to the existing methods in the near future. Our algorithm
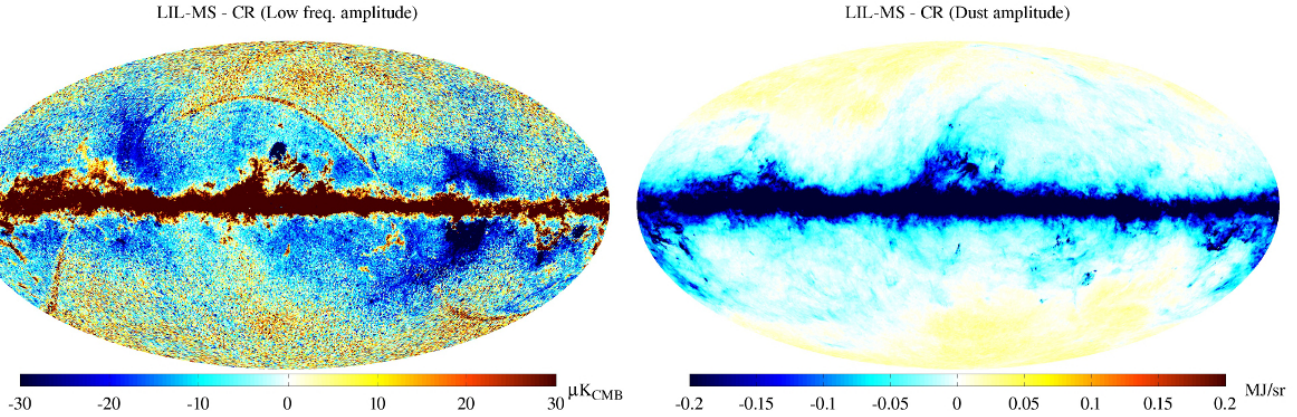
Figure 24. Difference between LIL-MS low frequency and dust amplitudes at 30 GHz and 353 GHz respectively and the corresponding CR maps degraded to nside=128. A monopole and dipole calculated at latitudes $|b| > 30°$ has been subtracted.

can be also be extended to include polarization, since the only requirement for it to be applicable is that the foreground model be almost-linear. As recent results from the Bicep2 show (Ade et al. 2014; Planck Collaboration et al. 2014) we are entering a regime in the CMB experiments where the signal is buried in the foregrounds and assumptions about the foregrounds can have a big influence on the interpretation of the experimental data. Finally, LIL-MS can be easily modified to separate out the *y*-type distortion (Zeldovich & Sunyaev 1969) by simply replacing the CO component by the *y*-type distortion component. We will present the results for the *y*-distortions in a separate publication.

REFERENCES

Ade P. A. R., et al., 2014, Physical Review Letters, 112, 241101 (arXiv:1403.3985)

Bennett C. L., et al., 2003, ApJS, 148, 97 (arXiv:astro-ph/0302208)

Bennett C. L., et al., 2013, ApJS, 208, 20 (arXiv:1212.5225)

Brandt W. N., Lawrence C. R., Readhead A. C. S., Pakianathan J. N., Fiola T. M., 1994, ApJ, 424, 1

Cardoso J.-F., Martin M., Delabrouille J., Betoule M., Patanchon G., 2008, preprint, (arXiv:0803.1814)

Conn A. R., Gould N. I. M., Toint P. L., 2000, Trust-Region Methods. Society for Industrial and Applied Mathematics, Philadelphia

Delabrouille J., Cardoso J.-F., Le Jeune M., Betoule M., Fay G., Guilloux F., 2009, A&A, 493, 835 (arXiv:0807.0773)

Delabrouille J., et al., 2013, A&A, 553, A96 (arXiv:1207.3675)

Draine B. T., Lazarian A., 1998a, ApJL, 494, L19 (arXiv:astro-ph/9710152)

Draine B. T., Lazarian A., 1998b, ApJ, 508, 157 (arXiv:astro-ph/9802239)

Draine B. T., Lazarian A., 1999, ApJ, 512, 740 (arXiv:astro-ph/9807009)

Erickson W. C., 1957, ApJ, 126, 480

Eriksen H. K., et al., 2006, ApJ, 641, 665 (arXiv:astro-ph/0508268)

Eriksen H. K., Jewell J. B., Dickinson C., Banday A. J., Górski K. M., Lawrence C. R., 2008, ApJ, 676, 10 (arXiv:0709.1058)

Finkbeiner D. P., Davis M., Schlegel D. J., 1999, ApJ, 524, 867 (arXiv:astro-ph/9905128)

Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, ApJ, 473, 576 (arXiv:astro-ph/9605054)

Gervasi M., Tartari A., Zannoni M., Boella G., Sironi G., 2008, ApJ, 682, 223 (arXiv:0803.4138)

Gill P., Murray W., 1976, in Watson G., ed., Lecture Notes in Mathematics, Vol. 506, Numerical Analysis. Springer, Berlin, pp 134–147, doi:10.1007/BFb0080120

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759 (arXiv:astro-ph/0409513)

Haslam C. G. T., Klein U., Salter C. J., Stoffel H., Wilson W. E., Cleary M. N., Cooke D. J., Thomasson P., 1981, A&A, 100, 209

Hauser M. G., et al., 1998, ApJ, 508, 25 (arXiv:astro-ph/9806167)

Hivon E., Górski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, ApJ, 567, 2 (arXiv:astro-ph/0105302)

Kogut A., et al., 2011, JCAP, 7, 25

Longair M. S., Sunyaev R. A., 1972, Soviet Physics Uspekhi, 14, 569

Martínez-González E., Diego J. M., Vielva P., Silk J., 2003, MNRAS, 345, 1101 (arXiv:astro-ph/0302094)

P. Andre et al. 2014, JCAP, 2, 6 (arXiv:1310.1554)

Planck Collaboration et al., 2013a, preprint, (arXiv:1303.5062)

Planck Collaboration et al., 2013b, preprint, (arXiv:1303.5072)

Planck Collaboration et al., 2014, preprint, (arXiv:1409.5738)

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes. Cambridge University Press, Cambridge

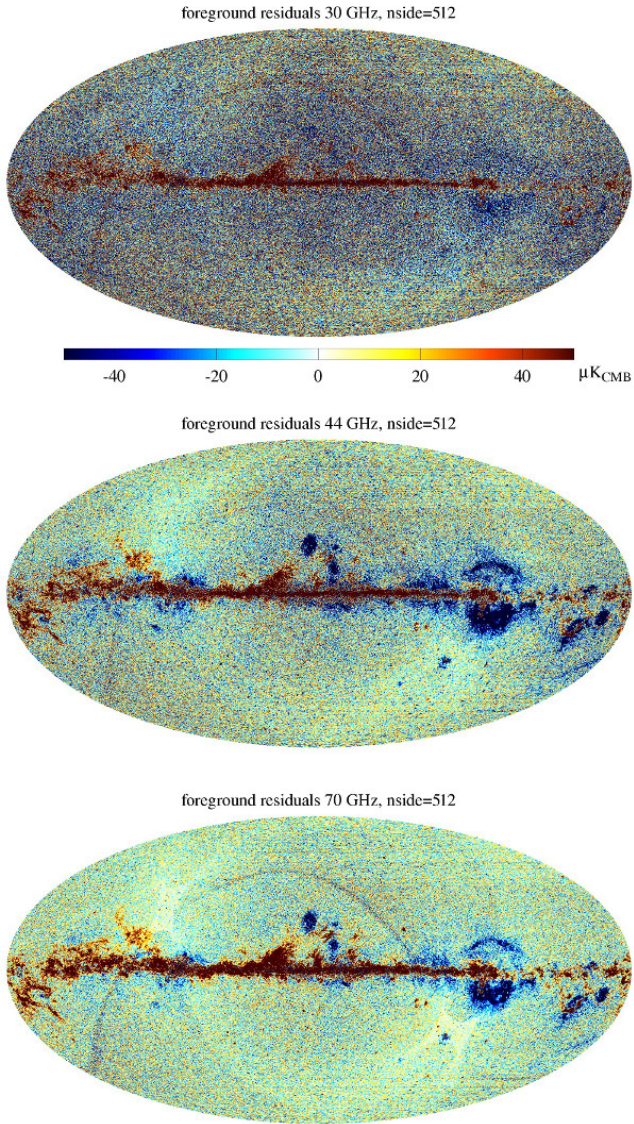Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525 (arXiv:astro-ph/9710327)

Figure A1. Residuals for the sum of all foregrounds for the LFI channels.

Smoot G. F., et al., 1992, ApJL, 396, L1

Sodroski T. J., et al., 1994, ApJ, 428, 638

Stuart A., Ord J. K., Arnold S., 2004, Kendall's Advanced Theory of Statistics, vol 2A. John Wiley & Sons Ltd, Sussex

The COrE Collaboration 2011, preprint, (`arXiv:1102.2181`)

Vansyngel F., Wandelt B. D., Cardoso J.-F., Benabed K., 2014, preprint, (`arXiv:1409.0858`)

Wheelock S. L., et al., 1994, NASA STI/Recon Technical Report N, 95, 22539

Zeldovich Y. B., Sunyaev R. A., 1969, ApSS, 4, 301

APPENDIX A: TOTAL FOREGROUND RESIDUALS IN ALL THE SEVEN PLANCK CHANNELS USED IN COMPONENT SEPARATION

We show in Figs. A1 and A2 the residuals for the sum of all foreground components in order to validate the recovered values of the spectral indices.
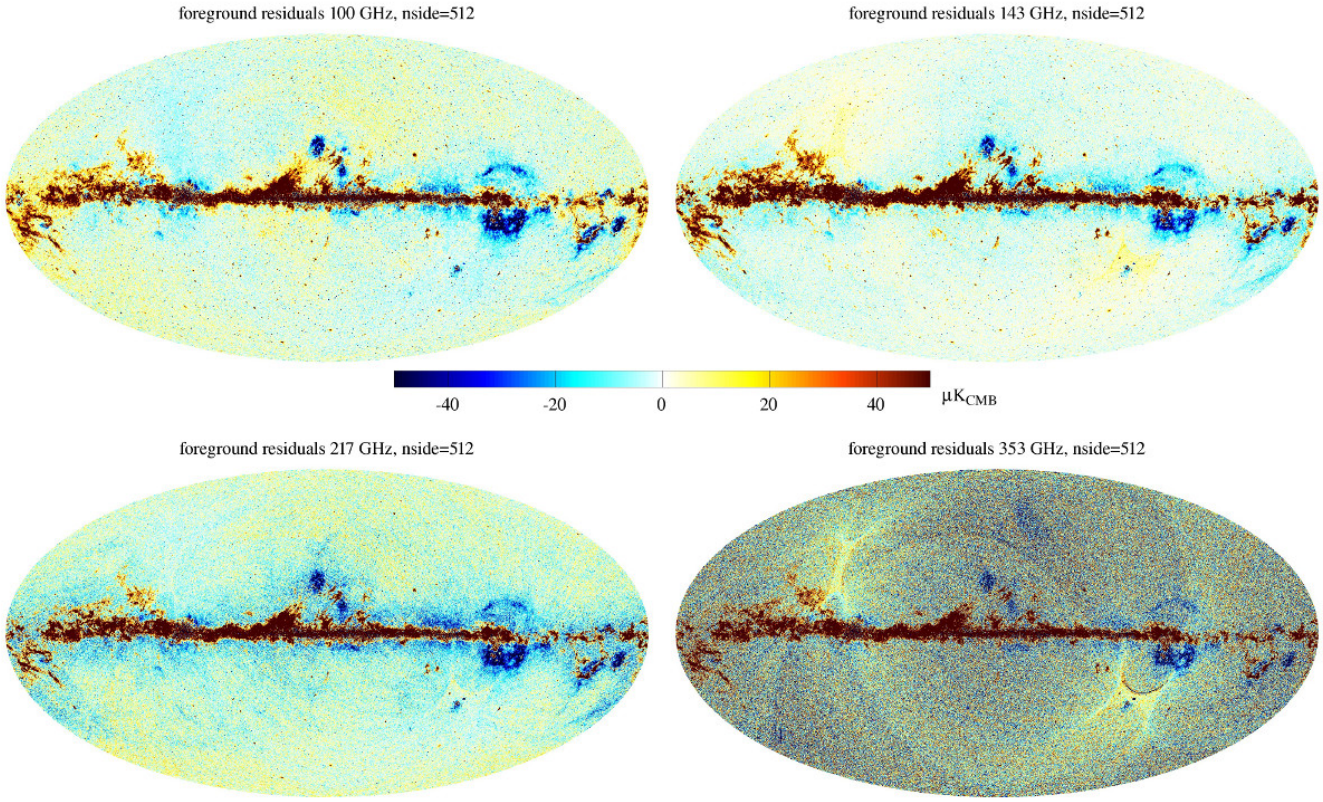
Figure A2. Residuals for the sum of all foregrounds for the HFI channels.