

A full sky, low foreground, high resolution CMB map from WMAP

J. Delabrouille¹, J.-F. Cardoso^{1,2}, M. Le Jeune¹, M. Betoule¹, G. Fay^{1,3}, and F. Guilloux^{1,4,5}

¹ Laboratoire AstroParticule et Cosmologie (APC), CNRS: UMR 7164, Université Paris Diderot – Paris 7, 10 rue A. Domon et L. Duquet, 75205 Paris Cedex 13, France
e-mail: delabrouille@apc.univ-paris7.fr

² Laboratoire de Traitement et Communication de l'Information (LTCI), CNRS: UMR 5141, Télécom ParisTech, 46 rue Barrault, 75634 Paris Cedex, France

³ Laboratoire Paul Painlevé, CNRS: UMR 8524, Université Lille 1, 59655 Villeneuve d'Ascq Cedex, France

⁴ Modélisation Aléatoire de Paris X (MODAL'X), Université Paris Ouest – Nanterre-la-Défense, 200 avenue de la République, 92001 Nanterre Cedex, France

⁵ Laboratoire de Probabilités et Modèles Aléatoires (LPMA), CNRS: UMR 7599, Université Paris Diderot – Paris 7, 175 rue du Chevaleret, 75013 Paris, France

Received 3 July 2008 / Accepted 18 October 2008

ABSTRACT

Context. The WMAP satellite has made available high quality maps of the sky in five frequency bands ranging from 22 to 94 GHz, with the main scientific objective of studying the anisotropies of the Cosmic Microwave Background (CMB). These maps, however, contain a mixture of emission from various astrophysical origin, superimposed on CMB emission.

Aims. The objective of the present work is to make a high resolution CMB map in which contamination by such galactic and extragalactic foreground emissions, as well as by instrumental noise, is as low as possible.

Methods. The method used is an implementation of a constrained linear combination of the channels with minimum error variance, and of Wiener filtering, on a frame of spherical wavelets called needlets, allowing localised filtering in both pixel space and harmonic space.

Results. We obtain a low contamination low noise CMB map at the resolution of the WMAP W channel, which can be used for a range of scientific studies. We obtain also a Wiener-filtered version with minimal integrated error.

Conclusions. The resulting CMB maps offer significantly better rejection of galactic foreground than previous CMB maps from WMAP data. They can be considered as the most precise full-sky CMB temperature maps to date.

Key words. cosmology: cosmic microwave background – methods: data analysis

1. Introduction

The WMAP satellite is one of the most successful experiments dedicated to mapping the Cosmic Microwave Background (CMB). The all-sky maps obtained in the WMAP five frequency bands, in temperature and polarisation, offer the best data set to date for making a sensitive all-sky map of the CMB anisotropies.

The CMB, however, is not the only source of emission at WMAP frequencies. Diffuse galactic emission from several processes contaminates the maps with an amplitude roughly proportional to the cosecant of the galactic latitude, compromising the observation of the CMB close to the galactic plane. In addition, a background of radio and infrared compact sources, galactic or extra-galactic, contributes to the total emission even at high galactic latitudes. Component separation consists of separating one or more of these sources of emission from the others in the data.

One of the main objectives of CMB experiments is the measurement of the CMB angular power spectrum C_ℓ which, with the assumption of statistical isotropy, describes the second order distribution of the CMB and can be estimated on a fraction of the sky. For this reason, many ground-based and balloon-borne experiments have concentrated their observations in “clean” regions of the sky, where galactic emission is low enough to negligibly impact the observations. For power spectrum estimation from full-sky observations, a safe approach consists of masking

regions at low galactic latitude, and estimating power spectra on the cleanest regions of the sky. The impact of extragalactic point sources (evenly spread on the sky) on power spectrum estimates can be evaluated and corrected for using ancillary data (catalogues of known point sources and priors on the statistical distribution of sources).

Besides the power spectrum, the CMB map itself is interesting for several additional purposes:

- as a CMB template, to be subtracted from millimetre-wave observations when the scientific focus is on other emissions, or to be used for the calibration of other instruments;
- to assess the statistical isotropy of the CMB and check the homogeneity and isotropy of the Universe on the largest scales;
- to search for signatures of non-trivial topology, such as that of a multi-connected universe (Aurich et al. 2006; Caillerie et al. 2007; Niarchou & Jaffe 2007);
- to search for correlations of the CMB map with other emissions (Nolta et al. 2004; Fosalba & Gaztañaga 2004; Cabré et al. 2006; Cao et al. 2006; Pietrobon et al. 2006; McEwen et al. 2007; Rassat et al. 2007);
- to search for signatures of a non Gaussian distribution in the CMB (Mukherjee & Wang 2004; Vielva et al. 2004; Cayón et al. 2005; Cabella et al. 2006; McEwen et al. 2006b;

Tojeiro et al. 2006; McEwen et al. 2006a; Wiaux et al. 2007; Yadav & Wandelt 2007).

Several CMB maps obtained with the WMAP data are available for such research projects. The WMAP team has released part-sky foreground-reduced maps in the Q , V and W bands, and maps obtained by an Internal Linear Combination (ILC) of all WMAP channels (Hinshaw et al. 2007; Gold et al. 2008). Tegmark et al. (2003) have produced CMB maps with WMAP one year data, and subsequently with three year data, based on a harmonic space ILC method. Eriksen et al. (2004) presented an alternate version of the ILC CMB map at 1 degree resolution on one year data. Eriksen et al. (2007) used a Gibbs sampler to draw 100 realisations of the CMB under the posterior likelihood of a model of CMB and foregrounds; their estimated CMB map is the average of these realisations for three year data. On three year data again, Park et al. (2007) use an ILC technique on 400 different pixel ensembles, selected by the spectral index of the foreground emission as estimated by the WMAP team using a Maximum Entropy Method (MEM). More authors have addressed component separation on WMAP three year data and produced versions of a “clean” CMB map (Maino et al. 2007; Saha et al. 2007; Bonaldi et al. 2007). More recently, Kim et al. (2008) produced a CMB map from WMAP five year data.

All available maps suffer from limitations, some of which result from specific choices in the way the CMB map is produced. Several of these maps, for instance, do not fully exploit the resolution of the original observations. Some focus on cleaning the CMB of foreground at high galactic latitude, and are significantly contaminated by foreground in the galactic plane. Some are not full sky CMB maps. Finally, not all of the available maps have well characterised noise and effective beam. All these limitations impact on their usefulness for accurate CMB science.

In this paper, we address the problem of making a CMB temperature map which has the following properties:

- full sky coverage;
- being as close as possible to the true CMB (minimum variance of the error) everywhere on the sky, and on all scales;
- having the best resolution possible;
- having a well-characterised beam and noise.

In the following, we review these requirements and their impact on a CMB cleaning strategy (Sect. 2). We then review and compare available maps in Sect. 3. In Sect. 4, we describe and explain our ILC needlet method. The approach is tested and validated on realistic simulated data sets (Sect. 5) before applying it to WMAP data (Sect. 6). We then compare our CMB maps to the other existing maps, discuss the results, and conclude.

This paper considers only *temperature* maps.

2. General considerations

2.1. Requirements

We start with a review of the requirements above, and on the implication on the method to be used.

2.1.1. Full sky?

WMAP data are full sky. We wish to devise a method which allows recovery of an estimate of CMB emission everywhere, including in the galactic plane, and even (as much as possible) in the galactic centre as well as in pixels strongly contaminated by compact sources.

The large scale correlation properties of the CMB make it possible to estimate the CMB emission even in unobserved regions, by some kind of interpolation. This is what is also obtained with the Gibbs sampling technique of Eriksen et al. (2007). Equivalent in spirit although quite different in implementation is the use of an “inpainting” method such as that of Elad et al. (2005) and Abrial et al. (2008). Such interpolation methods alone are not fully satisfactory, as they discard information. In particular they do not allow the recovery of small scale CMB features in the mask. This is obvious, for instance, in the Gibbs-sampling average map of Eriksen et al. (2007).

At the opposite extreme, one may try to separate components in the galactic plane independently of what is done at higher galactic latitudes, since levels and properties of foreground emission depend strongly on sky direction. Hinshaw et al. (2007) and Tegmark et al. (2003) divide the sky into several independent regions, perform component separation independently in these regions, and then make a composite map by stitching together these independent solutions. Such approaches discard information (zone-to-zone correlations) and require careful treatment at the zone borders to avoid discontinuities and ringing.

A good method should perform well on both counts: localised processing and full exploitation of large scale correlations of the CMB and of galactic foreground emissions. This can be achieved with a spherical wavelet or *needlet* analysis of the maps (using the tools developed in Marinucci et al. 2008 and Guilloux et al. 2008), which is our approach in the present work.

2.1.2. Minimum variance?

Recovery of a CMB map can be conducted following various objectives, quantified by different “figures of merit”. In this work, we choose, as do most authors, to minimise the variance of the difference between true and recovered CMB (this is the *overall* error; it includes additive noise, foreground emissions, and even multiplicative errors affecting the CMB itself).

This choice alone does not fully characterise the method to be used. The best theoretical solution also depends on the model of the data. An overview of existing methods can be found in the review by Delabrouille & Cardoso (2007).

In this paper, contrarily to other approaches which rely heavily on the structure of the data as described by a model, for instance a noisy linear mixture of independent components (Delabrouille et al. 2003), we assume as little as possible about the foreground emissions and the noise. In fact, we assume nothing except the following:

- the WMAP data are well calibrated, and have a known beam in each channel;
- the instrumental noise in all WMAP maps is close to being Gaussian and uncorrelated; its pixel-dependent level is approximately known;
- the CMB anisotropy emission law is known to be the derivative with respect to temperature of a $T = 2.725$ K black-body;
- to first order, the template of CMB anisotropies is well represented by a Gaussian stationary random field, the spectrum of which is given by the WMAP best fit (as will be seen later, this last assumption is needed only to derive the Wiener filter; it is not necessary for our needlet ILC map).

These assumptions lead us to consider an “Internal Linear Combination” (ILC) method, followed by a Wiener filter to minimise the error integrated over all scales.

2.1.3. Best resolution?

The WMAP data comes in five frequency channels with varying resolution. To make the best of the data, we need a method which uses the smallest scale information from the *W* band, and information from additional bands (*V*, then *Q*, then *Ka*, and finally *K*) at increasingly larger scales. Multi-scale tools are well suited for this purpose.

The “best possible resolution” is not a well defined concept (and not necessarily the resolution of the *W* band). Indeed, there is a conflict between best possible resolution and minimum variance, as one can smooth or deconvolve a map arbitrarily in harmonic space, reducing or increasing the total noise variance in the process. Here, we make a map at the resolution of the *W* channel over the full sky, leaving open the option to filter this map if needed to reduce the noise – or deconvolve it for better angular resolution. Note that additional global filtering or deconvolution does not change the signal to noise per mode (only the integrated S/N).

The minimum variance map is obtained with a Wiener filter, which smoothes the map depending on the signal to noise ratio. As the noise and the contaminants are inhomogeneous, there is a strong motivation for the smoothing to depend on the location on the sky (the optimal solution to the resolution-variance trade-off depending on the contamination level, which is local). If we relax the constraint about beam homogeneity, again, spherical needlets offer a natural way to obtain such location-dependent smoothing. In the present, however, in order to preserve the constancy of the effective beam over the sky, we implement the Wiener filter in harmonic space.

2.1.4. Accurate characterisation?

A fully accurate characterisation of the beam and noise is not straightforward, in particular because of the limited knowledge about the original frequency maps, which automatically propagates into the final CMB map. This work makes several approximations about beams and noise. Beams are assumed to be symmetric and therefore described by the b_ℓ transfer functions provided by the WMAP team. The instrumental noise is assumed uncorrelated, although non stationary. Analyses and Monte-Carlo simulations are used to characterise the residual noise of the final map, as well as to estimate the contribution of residual foregrounds, and biases if any. This is detailed further below.

2.1.5. Noise

Throughout this paper, the term “noise” refers to all sources of additive error, i.e. instrumental noise *and* foreground emissions.

2.2. Evaluation and comparison of reconstructed CMB temperature

We briefly discuss here the tools used for characterising and comparing CMB temperature maps.

2.2.1. Map description

A pixelised map is fully characterised by the specification of:

- a set of temperature values y_p in a number of pixels (here indexed by p);

- the effective beam at each pixel p , which in the most general case is a function $b_{p,p'}$;
- the noise n_p , the statistical properties of which, in the Gaussian case, are fully described by a noise covariance matrix $N_{p,p'}$.

The map value y_p is then linked to the true signal value s_p by:

$$y_p = \sum_{p'} b_{p,p'} s_{p'} + n_p. \quad (1)$$

The full characterisation of a given CMB map requires the specification of the additive noise n_p and of the response $b_{p,p'}$. When the beam is stationary over the sky and symmetric, which we assume in this work, it is fully specified by the coefficients b_ℓ of the expansion of the beam in Legendre polynomials.

2.2.2. Assumptions

Throughout this paper, the beam is assumed symmetric. Although this is an approximation, most pixels of the WMAP map are “visited” by any particular detector through a wide range of intersecting scans. The average beam in that pixel then is an average of the physical beam over many orientations, which makes the symmetry assumption reasonable.

In addition, in the absence of any specific localised processing, the beam is assumed to be invariant over the sky.

With the above two assumptions, the effect of beam convolution is best represented in harmonic space, with a multiplicative coefficient b_ℓ , independent of m , applied to the harmonic coefficients $a_{\ell m}$ of the map. We assume perfect beam knowledge as well as perfect calibration, so that no multiplicative uncertainty is attached to the map description (the beam integral, approximated as $\sum_{p'} b_{p,p'}$, is equal to unity independently of p , or, equivalently, the value of b_ℓ for $\ell = 0$ is assumed to be exactly unity).

The noise n_p of the original WMAP maps, for each frequency channel and each differencing assembly, is assumed uncorrelated from pixel to pixel, i.e. $N_{p,p'} = \langle n_p n_{p'} \rangle = \delta_{pp'} \sigma_p^2$. The variance σ_p^2 is pixel dependent because of uneven sky coverage. Noise is non-stationary, but assumed to be Gaussian distributed.

2.2.3. Comparison of maps at different resolution

The comparison of CMB maps is meaningful only if the maps are at the same resolution. As long as the beam transfer function does not vanish at any useful ℓ (which is always the case for Gaussian beams), the resolution of any map can be changed to anything else by harmonic transform and multiplication by the ratio of the beam transfer functions. In the present work, we approximate WMAP beams by their symmetric fit (i.e. azimuthally averaged beams), which is justified by the large range of scanning directions at any point in the sky.

This convolution – deconvolution property is widely used throughout this paper.

2.2.4. Masking

We define “tapered” regions of the sky for map comparison at varying galactic latitude. In particular, we define a Low Galactic Latitude (LGL) region and a complementary High Galactic Latitude (HGL) region. The LGL region, used to evaluate results in the galactic plane, completely excludes all data above 30 degrees galactic latitude (and below -30 degrees), and has a

15 degree transition zone with a cosine square shape. All pixels at absolute galactic latitudes below 15 degrees are kept with a coefficient of 1. The HGL region is the complementary, i.e. $HGL = 1 - LGL$. These “tapered” regions allow the computation of local power spectra with negligible spectral leakage of large scale power into small scales.

2.2.5. Power spectra comparison

For a given beam (i.e. multiplicative response as a function of ℓ), the comparison of the total map power as a function of ℓ (i.e. of the power spectra of the maps) is a direct figure of merit. The lower the power spectrum, the better the map.

Indeed, for fixed b_ℓ , two distinct maps contain the exact same contribution from CMB, but different noise levels. CMB and noise being assumed to be uncorrelated, power spectra are the sum of a CMB term ($b_\ell^2 C_\ell$) and a noise term (N_ℓ). The CMB term being the same for both spectra, any difference in total power spectrum can be interpreted as a difference in noise level.

Power spectra are computed independently for different regions of the sky (e.g. inside or outside the galactic plane). To minimise aliasing due to sharp cuts, we use masks with smooth transitions, defining the LGL and HGL regions described above.

The power spectrum of a map is evaluated as:

$$\widehat{C}_\ell = \frac{1}{(2\ell + 1)\alpha} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2; \quad (2)$$

α is a normalising factor computed as the average value of the squared masking coefficients¹.

Power spectra estimated directly in this way from a masked sky are unreliable for modes corresponding to angular sizes larger than the typical size of the zone of the sky retained for computation.

2.3. Methods

We now give a brief introduction to the two main methods used in this paper (ILC and Wiener filtering). Many other methods exist for CMB cleaning (or component separation in general), which assume varying degrees of prior knowledge about sky emission, and model the data in different ways. These methods are not discussed nor used in this paper. For a review, see [Delabrouille & Cardoso \(2007\)](#).

2.3.1. The ILC

The data are modelled as

$$\mathbf{x} = \mathbf{a}s + \mathbf{n} \quad (3)$$

where \mathbf{x} is the vector of observations (e.g. five maps), \mathbf{a} is the response to the CMB for all observations (e.g. a vector with 5 entries equal to 1 if WMAP data only are considered) and \mathbf{n} is the noise. Here it is assumed that all observations are at the same resolution.

The ILC provides an estimator \hat{s}_{ILC} of s as follows:

$$\hat{s}_{\text{ILC}} = \frac{\mathbf{a}^t \widehat{\mathbf{R}}^{-1} \mathbf{x}}{\mathbf{a}^t \widehat{\mathbf{R}}^{-1} \mathbf{a}} \quad (4)$$

¹ The masking coefficient is simply 1 in regions kept for power spectrum computations, 0 in regions masked, and between 0 and 1 in the transitions.

where $\widehat{\mathbf{R}}$ is the empirical covariance matrix of the observations (e.g., a 5×5 matrix when 5 channels are considered). The biggest concern with the ILC (and the source of all problems of bias) is that this empirical covariance matrix has to be estimated on the data themselves. This will be discussed further in Sect. 4.

Note that the ILC solution of Eq. (4) is the linear filter which minimises the total variance of the output map, under the condition that the filter has unit response to the signal of interest (the signal with the emission law given by vector \mathbf{a}).

The details of the method for its implementation in the context of this work are further discussed in Sect. 4.

2.3.2. Wiener filtering

Given a single CMB map of known beam (assumed to be constant over the sky), it is possible to minimise the contamination by noise and foreground by (one-dimensional) Wiener filtering. The data is modelled as $x = s + n$, where now x is a single map, s the true CMB and n the noise. The Wiener filter gives to individual “modes” a weight proportional to the fraction of signal power in that mode, i.e.

$$\hat{s}_{\ell m} = \frac{b_\ell^2 C_\ell}{b_\ell^2 C_\ell + N_\ell} x_{\ell m} \quad (5)$$

where $b_\ell^2 C_\ell$ and N_ℓ are the power spectra of the (smoothed) CMB and of the noise (including smoothed foreground or foreground residuals) respectively, and $x_{\ell m}$ is the original noisy CMB map. It should be noted that if the CMB and the noise are uncorrelated, then $b_\ell^2 C_\ell + N_\ell = X_\ell$ is the power spectrum of the map $x_{\ell m}$, and the Wiener filter can be estimated directly using only a prior on the CMB power spectrum (assuming C_ℓ is known), and estimating X_ℓ on the map itself.

Wiener filtering in harmonic space minimises the variance of the error in the map if signal and noise are Gaussian and stationary. For non-stationary contaminants, the Wiener filter (5) is still meaningful, but is no longer optimal. Here, the Wiener filtering is applied after the ILC. Foreground contamination, strongly reduced by the ILC, is no longer dominant. Instead, instrumental noise is the major source of error on all scales where the Wiener filter is useful (departs from 1). A harmonic space implementation of the Wiener filter disregards this non-stationarity, applying the same coefficient to a given scale, whether it is in a region of deeper integration or not.

Better efficiency may be obtained by an implementation in another domain than the harmonic space (e.g. needlets), but this is not investigated further in this paper, as the consequence is a non stationary equivalent beam.

3. Evaluation and comparison of available maps

Before describing how to make yet another CMB map from WMAP data, we review the existing maps and evaluate in what respects they can be improved.

We start with a discussion of the existing methods, identifying for each of them the strengths and weaknesses of the approach, and their foreseeable consequences. Available CMB maps obtained from WMAP data are compared in terms of resolution, of the estimated contamination by foreground emission, and of noise level. In the absence of an absolute reference, discrepancies between available maps are also evaluated. This comparison permits an estimate of typical uncertainties, and to outline the “difficult regions” for CMB cleaning (which, unsurprisingly, are mostly located close to the galactic plane). We also

look specifically for residuals of galactic contamination by comparing the power spectrum of the reconstructed CMB map at high and at low galactic latitudes. Significant discrepancies between the two are interpreted as indicative of a residue of foreground emission.

3.1. Available maps

3.1.1. The WMAP ILC

The ILC maps obtained by the WMAP team (denoted as WILC1, WILC3 and WILC5 hereafter, depending on whether they are obtained using one year, three year or five year data) are described in [Bennett et al. \(2003\)](#), [Hinshaw et al. \(2007\)](#), and [Gold et al. \(2008\)](#) respectively.

For the three year and five year maps, the original frequency maps are smoothed to a common resolution of one degree. The sky is subdivided into 12 regions. One large region covers most of the sky at moderate to high galactic latitudes. The rest of the sky, concentrated around the galactic plane, is divided into 11 regions of varying galactic emission properties (amplitude and colour). An internal linear combination (ILC) is performed independently in each of these regions. A full sky composite map is obtained by co-adding the maps of the individual regions (with a ≈ 1 degree transition between the zones to avoid sharp edge effects). Finally, a bias (inevitable consequence of empirical CMB-foreground correlation) is estimated by Monte-Carlo simulations, and subtracted from the composite map, to yield the final CMB map.

The three year and five year ILC maps differ from the one year map in several respects. The most significant is the recognition of the existence of a bias, and the attempt at correcting it using simulations. The limitations of the maps include their resolution (one degree), and the use of small regions in the ILC, which is bound to cause more bias than necessary on large scales (comparable to patch sizes)², as well as edge effects. This results in discontinuities between regions, obvious for instance in the estimated bias map shown in [Hinshaw et al. \(2007\)](#).

In the method used by the WMAP team, the coefficients of the linear combination used over most of the sky (region 0, which corresponds to the largest part of the sky at high galactic latitudes and a few low galactic latitude patches, and region 1, in the galactic plane but away from the galactic center) are set using only a small subset of the data inside the Kp2 cut (where the galactic emission is the strongest). This choice favours the rejection of galactic contamination, at the price of sub-optimal weighting of the observations in regions where the error is dominated by noise. It also assumes that the emission laws and relative power of the different foregrounds are the same in these regions, which is a strong (and probably wrong) assumption.

Furthermore, the ILC weights are set by minimising the variance of the map at one degree resolution. Modes at higher ℓ get very sub-optimal weighting, as they do not contribute significantly to the total variance of the one degree map. The *K* and *Ka* bands, in particular, contribute respectively about 0.156 and -0.086 to region 0 (the largest one) for WILC3. As a consequence, the final ILC map cannot be meaningfully deconvolved to better resolution than about 1 degree (because this would amplify dramatically small scale noise coming from the lowest frequency channels).

Finally, there is also an unsatisfactory degree of arbitrariness in the choice of the regions, which depend on priors about foreground emission, and are somewhat elongated across the galactic plane for no particular reason. Although none of these choices is unreasonable, the impact of this arbitrariness on the final map is difficult to evaluate.

For all these reasons, the WILC maps leave considerable margin for improvement. We aim, in particular, at obtaining a CMB map with better angular resolution, and with a better handling of non-stationarity and scale dependence of the contamination (foregrounds and noise).

3.1.2. The WMAP foreground-reduced maps

For temperature power spectrum analysis, the WMAP team has used part-sky foreground-reduced maps. The processing for foreground removal for the three year and five year releases is described in [Hinshaw et al. \(2007\)](#). Model templates for galactic emission are fitted to the Q, V and W WMAP channels outside of the Kp2 mask. A linear combination of synchrotron, free-free and dust, based on this fit, is then subtracted from the full sky Q, V, W maps.

In this procedure, a first galactic template, supposed to correspond to a linear combination of synchrotron and free-free emission, is obtained from the difference between the *K* and *Ka* bands. This template is produced at one degree resolution. An additional free-free template is obtained from H α emission ([Finkbeiner 2003](#)) corrected for dust extinction ([Bennett et al. 2003](#)). A dust template is obtained from model 8 of [Finkbeiner et al. \(1999\)](#). “Clean” Q, V and W maps are obtained by decorrelation of these templates from the original Q, V and W observations.

The main limitation of this approach is that the model used is insufficient to guarantee a good fit of the total foreground emission simultaneously inside and outside of the Kp2 mask. As a consequence, the maps produced are heavily contaminated by foreground emissions in the galactic plane, the priority being given to higher galactic latitudes, with the objective of obtaining a part-sky high quality map on which high multipole CMB power spectra could be estimated reliably.

In addition, the maps are likely to depend significantly on the prior model assumed. Here, the WMAP team chooses dust model number 8 of [Finkbeiner et al. \(1999\)](#), and also ignores the plausible existence of anomalous dust emission. The exact impact of these a priori decisions is difficult to evaluate. Anomalous dust emission will come into the foreground reduced maps chiefly as a contaminant in the K-Ka synchrotron template, and hence lead to an erroneous extrapolation of this template to higher frequencies. Assumptions about the emission laws of the template correlated to H-alpha, and of the dust template, if inaccurate, will also result in inaccurate subtraction of free-free and dust. These errors in the estimate of foregrounds are bound to leave, after template subtraction, foreground residuals in the foreground reduced maps.

As an additional drawback, we note that the method generates correlated noise in the foreground-reduced maps, originating either from K and Ka channel noise or from a background of radio sources (see, e.g., [Huffenberger et al. 2007](#); and [Wright et al. 2008](#), for discussions on radio sources in WMAP data and their impact on the analysis of CMB observations with WMAP). Finally, on supra degree scales, the K and Ka bands, which are the most sensitive ones, are used only to subtract foreground emissions, whereas in the cleanest regions of the sky they would be more usefully used to estimate the CMB emission.

² Small patches contain few modes, hence empirical correlations are stronger, as discussed in Appendix A.

For all these reasons, WMAP foreground-reduced maps are not good CMB maps according to the criteria listed in the introduction.

3.1.3. The ILC by Eriksen et al.

Eriksen et al. (2004) have obtained a CMB map at 1 degree resolution with another implementation of the ILC. The map, denoted here EILC1, uses only one year data.

An interesting remark from Eriksen et al. is that the amount of residual dust is high in the ILC maps – the method being able to subtract only about half of the dust present in the W band. At scales larger than 1 degree, this lack of performance is likely to be due to the part of dust emission uncorrelated to low frequency galactic foreground. On the smallest scales, the situation is worse, as the low frequency WMAP channels do not have the resolution to help remove small scale dust emission from W band observations.

For this reason, in the present work, we improve on dust removal by using, as an additional measurement, the IRIS 100 micron dust template obtained from a combination of DIRBE and IRAS maps (Miville-Deschênes & Lagache 2005). As compared to the map of Eriksen et al., we also aim at better angular resolution – and, obviously, better sensitivity, achieved by using five year data sets.

3.1.4. The Gibbs-sampling map by Eriksen et al.

Recently, Eriksen et al. (2007) have produced a low resolution (3 degree) CMB map from WMAP three year data, using a Gibbs sampling technique to explore the likelihood of a parametric model of CMB and foreground emissions. A CMB map is obtained as the average of 100 sample CMB maps each drawn at random according to the posterior distribution of the model parameters given the observations.

The free parameters in the model are spherical harmonic coefficients $a_{\ell m}$ of the CMB map, the CMB harmonic power spectrum C_ℓ , monopole and dipole amplitudes in each WMAP band, the amplitude $a(\nu)$ of a dust template in each band, and amplitudes $f(p)$ and spectral indices $\beta(p)$ of a low-frequency foreground component, for each map pixel p .

The model is constrained by fixing the dust template at 94 GHz according to Finkbeiner et al. (1999), by a prior on the low-frequency foreground spectral index, assumed to be close to that of synchrotron radiation (-3 ± 0.3), and by the constraint that the monopole and dipole coefficients are orthogonal to the (noise-weighted) pixel-averaged foreground spectrum.

In spite of a good fit of the assumed model to the data at high galactic latitudes, there are some strong limitations to the resulting CMB map, and hence to its usage:

- the result of the sampling is obtained assuming a parametric model of foreground emission. There is no possible way of validating the systematic errors due to mismodelling, except marginalising over all possible model skies. This would require a Monte-Carlo simulation which takes into account all uncertainties in the modelling, not only values of the parameters for a given parametric model, but also the choice of the parameter set to be used to model the foreground emissions (varying the dust template according to uncertainties, assuming a different foreground model, etc.). This is not presently available;
- the resulting map is at 3 degree resolution and HEALPix $n_{\text{side}} = 64$, considerably worse than WMAP can do;

- the data sets are cut with the Kp2 mask. Although a CMB is recovered in the mask by the average of the sample maps, the effective resolution inside the mask is lower than in the rest of the sky. In some sense, the Gibbs sampling technique (as implemented here) allows us to recover in the mask what is predictable from the outside map (assuming stationarity of the CMB anisotropy field). It allows only for a limited prediction of the CMB signal in the masked zone.

For all of these reasons, the Gibbs-sampling map of Eriksen et al. (hereafter EGS3) is not a good “best CMB” map according to our criteria.

3.1.5. The ILC by Park et al.

Park et al. (2007) provide their own version (hereafter the PILC3 map) of a one degree resolution CMB map obtained by an ILC on WMAP three year data. The originality of their approach lies in the fact that they cut the sky into 400 pixel ensembles, selected from a prior on their spectral properties. The 400 ensembles are defined from 20×20 spectral index bins (20 for K-V spectral index, and 20 for V-W). This approach is motivated by the fact that ILC weights are expected to vary with varying foreground properties.

There is a weak point to this approach. The authors use, to define their pixel “bins”, the MEM solution derived by the WMAP team. If the MEM solution is wrong for a given pixel, that pixel will automatically be classified in the wrong pixel ensemble, and be weighted using the weights of the wrong population of pixels. To some extent then, this binning forces the result of the ILC to match the prior assumptions given by the MEM results. In turn, the MEM solution uses as a prior the result of the WMAP ILC, which is subtracted from the WMAP observations prior to using the MEM method to separate galactic foreground emissions.

As a consequence, the connection of the CMB map of Park et al. to the original WMAP data is far from direct. The map is bound to bear the signature of any arbitrary choice made before, in particular the choice of WMAP ILC regions, the 3-component model for galactic emission, and the MEM priors. For instance, discontinuities at the boundaries between the 12 regions of the WMAP ILC are clearly visible in the map of K-V spectral index used by the authors, as well as their group index (see Figs. 3a and 4a of their paper).

Park et al. then investigate the error in their reconstructed map by Monte-Carlo simulations. However, they use as an input galactic emission template the very model obtained by the MEM. This means that in the simulations, the spectral index maps are “exact”. Therefore, the simulations accurately investigate the errors only if the MEM solution is correct, which is not likely to be the case – at least not to the level of precision required to produce a CMB map useful for precision cosmology.

Our method described in Sect. 4 uses as little prior information as possible, and aims at better angular resolution than 1 degree.

3.1.6. The “clean” map of Tegmark et al.

The approach of Tegmark et al. (2003), on both one year and three year data, is the only work to date which aims at producing a CMB map with both full sky coverage and best possible resolution.

Tegmark et al. (2003) have performed a foreground analysis of the WMAP one year maps, producing two high resolution

Table 1. Available CMB maps.

Name	resolution	data used	Reference	URL
WILC1	1°	1-yr	Bennett et al. (2003)	http://lambda.gsfc.nasa.gov/product/map/dr1/imaps_ILC.cfm
TILC1	W channel	1-yr	Tegmark et al. (2003)	http://space.mit.edu/home/tegmark/wmap/cleaned_map.fits
EILC1	1°	1-yr	Eriksen et al. (2004)	http://www.astro.uio.no/~hke/cmbdata/WMAP_ILC_lagrange.fits
WILC3	1°	3-yr	Hinshaw et al. (2007)	http://lambda.gsfc.nasa.gov/product/map/dr2/ilc_map_get.cfm
EGS3	3°	3-yr	Eriksen et al. (2007)	http://www.astro.uio.no/~hke/gibbs_data/cmb_mean_stddev_WMAP3_n64_3deg.fits
PILC3	1°	3-yr	Park et al. (2007)	http://newton.kias.re.kr/~park/CMB/SILC400/SILC400_bc.fits
TILC3	W channel	3-yr		http://space.mit.edu/home/tegmark/wmap/cleaned3yr_map.fits
WILC5	1°	5-yr	Gold et al. (2008)	http://lambda.gsfc.nasa.gov/product/map/dr3/ilc_map_get.cfm
KILC5	1°	5-yr	Kim et al. (2008)	http://www.nbi.dk/~jkim/hilc/

maps of the CMB. The first one, the “clean” map (hereafter TILC1), is obtained by a variant of the ILC in which weights are allowed to vary as a function of ℓ , and are computed independently in nine independent regions. The second, the “Wiener” map, is a Wiener-filtered version of the same map, in which the Wiener filter is applied in harmonic space, but independently in each zone.

This work by Tegmark et al. is an early attempt at finding linear combinations of the WMAP data which vary both in harmonic space and in pixel space. The pixel variation of the weights is made using zones which are defined according to the level of contamination by foreground emission, as computed from WMAP map differences W-V, V-Q, Q-K and K-Ka. The authors do not specify exactly how the frequency range is divided into ℓ -bands. While the text seems to indicate that weights are computed independently for each ℓ , figures hint that the weights are actually band-averaged, in 50 logarithmic bands subdividing the multipole range. It makes a difference for the ILC bias. For an ILC implemented ℓ by ℓ , the number of modes at a given ℓ is $2\ell + 1$, and thus remains small for most of the useful range of scales, yielding significant bias (of the order of 10% at $\ell = 25$, and 1% at $\ell = 250$, for a full sky ILC; for a part sky ILC, the bias is multiplied by a factor of the order of the inverse of the sky fraction). For band averages, the effect would be less dramatic, because of the higher number of modes per individual ILC.

In the end, the authors obtain a CMB “clean” map with a “beam corresponding to the highest-resolution map band”, i.e. the beam of the W band.

The original paper describes the work done on the one year WMAP data. However maps for three year data are available on Max Tegmark’s web site (see Table 1). We use the three year map (TILC3) for comparison with our own solution.

Although the approach of Tegmark and collaborators is quite good at high galactic latitudes, we can see in Fig. 1 that it performs poorly in the galactic plane. Also, the authors have not removed detected point sources from the WMAP data before making the ILC. As a result, their CMB map contains obvious point source residuals, for instance around galactic longitude 305° and latitude 57° , where a 5 mK peak can be seen.

Our method, although bearing some similarity to that of Tegmark and collaborators, aims at significantly improving the error characterisation, as well as the quality of foreground cleaning in the galactic plane.

3.1.7. The “Wiener” map of Tegmark et al.

In addition to their TILC map, Tegmark et al. (2003) publish a Wiener map (hereafter TW map), obtained from the TILC map by independent Wiener filtering in the 9 regions. This results in reduced integrated error in all regions, at the price of

pixel-dependent extra smoothing. The consequence of this filtering is an effective zone-dependent beam.

Because of this extra smoothing, it is difficult to compare the TW map with other maps. The most meaningful figure of merit for the Wiener map, though, is the actual power of the error (output map minus true CMB). This is unavailable for any useful up-to-date real data set. Additional discussion about Wiener-filtered maps is deferred to Sect. 6.

3.1.8. The ILC map of Kim et al.

More recently, Kim et al. (2008) have made a CMB map from WMAP five year data, using a “harmonic” ILC method (KILC5 hereafter). Their method performs an ILC in the pixel domain but with pixel-dependent weights. The ILC weights are not constant over predefined zones on the sky but are computed as smooth weight maps defined in terms of a harmonic decomposition (hence the qualification of the method). More specifically, the weight maps are determined by minimising the total output CMB map variance with the constraint that these maps have no multipoles higher than ℓ_{cutoff} . For stability reasons, the KILC5 map is obtained with $\ell_{\text{cutoff}} = 7$. Prior to computing ILC weights, all maps are deconvolved from their beam (effectively blowing up noise on small scales, in particular for the lowest frequency channels). Then, the channels are combined using map modes for $\ell < 300$.

With the above choices, the reconstructed map cannot be good on small scales. As the authors notice themselves, using small scale modes results in minimisation of noise rather than foreground (and, obviously, rejecting the low-frequency observations, which are the noisiest on small scales after deconvolution from the beam). Better results could probably be obtained by estimating weight maps for different bands of ℓ . In essence, this is what our needlet ILC method permits us to achieve.

Limiting the number of modes of the weight maps to $\ell \leq 7$, apparently for reasons of singularity of the system to be solved, results in spatial coherence of the weights on scales of about 35 degrees. The galactic ridge, however, is about 1 degree thick. Hence, the spatial variability of the ILC weights achieved by Kim et al. (2008) is not quite adapted to the actual scale of foreground variation. The needlet ILC method presented in our paper, as will be seen later on, solves this issue in a very natural way.

3.1.9. Other maps

Other authors have performed various foreground cleaning processes in the WMAP observations, producing a number of CMB maps for several different models of the foreground emission. Bonaldi et al. (2007) perform component separation on WMAP data using the CCA method described in

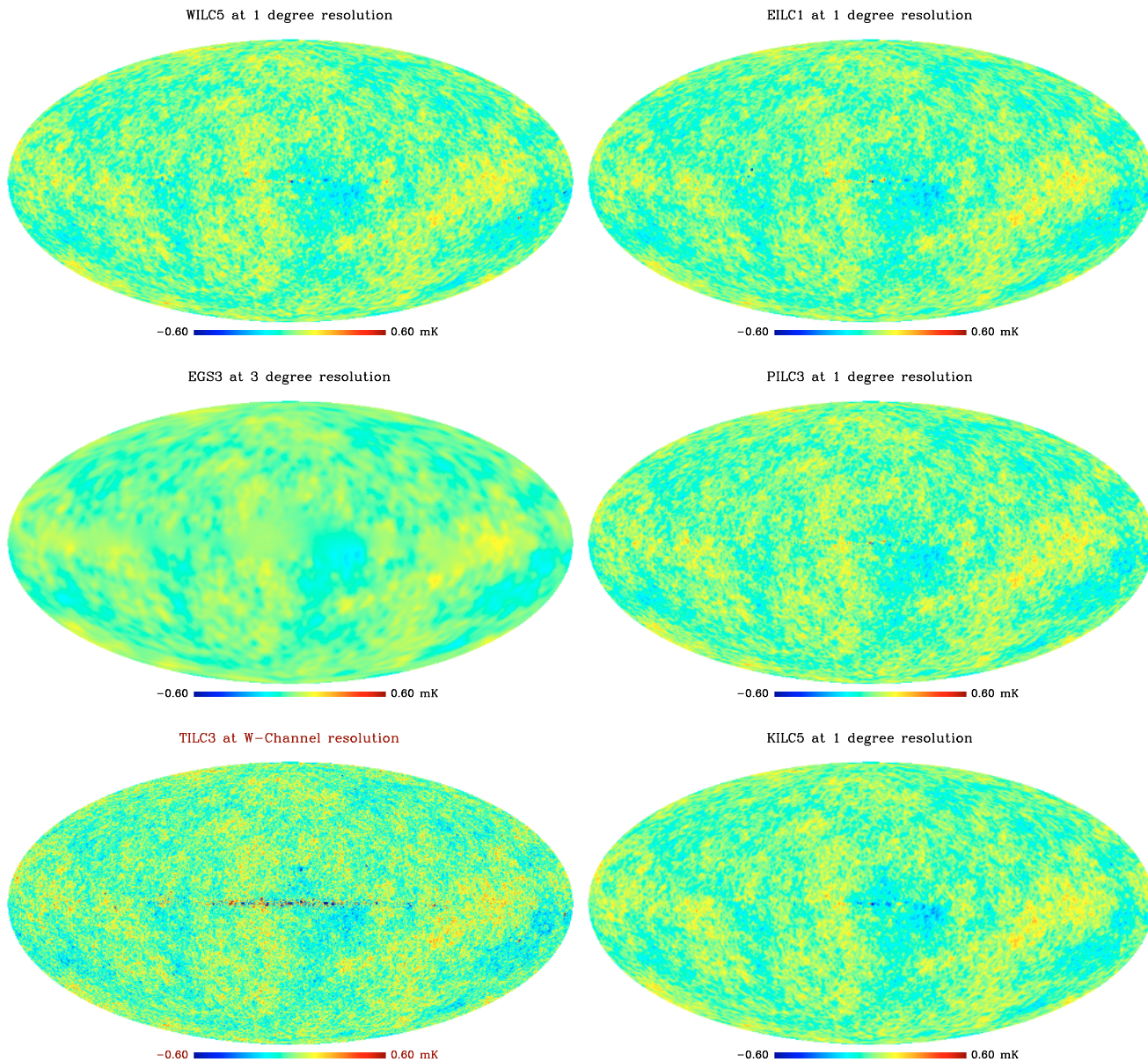


Fig. 1. A selection of presently available CMB maps from WMAP data. Contamination by galactic emission is visible in all of them at various levels, except for the 3-degree resolution EGS3 map (for which a galactic cut was applied and then filled by a plausible CMB extrapolated from higher galactic latitude data).

Bonaldi et al. (2006). Maino et al. (2007) obtain also several CMB maps, using the FastICA method (Maino et al. 2002). None of the CMB maps obtained is full sky, nor publicly available yet. They are not considered further in this analysis.

Finally, some foreground cleaning has also been performed by Saha et al. (2007). Their paper also includes an interesting analysis of the ILC bias. The primary goal of that work, however, is to compute the CMB power spectrum, rather than producing a CMB map.

3.1.10. Existing map summary

Figure 1 shows six available maps, all displayed on the same colour scale. It illustrates the resolution and foreground contamination of the various maps. Table 1 summarises the main properties of the maps. Only the TILC1 and TILC3 maps are high

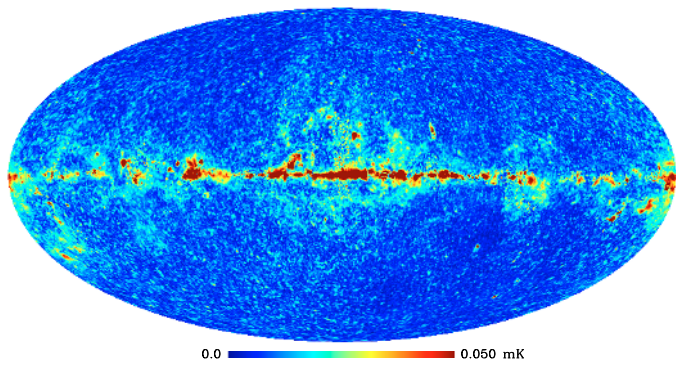
resolution attempts at component separation everywhere, including the galactic plane, combining all WMAP observations. All other maps (WILC maps, EILC1, EGS3, PILC3, KILC5) are at reduced resolution.

3.2. Map comparison

CMB maps produced from WMAP one year data have been compared by Eriksen et al. (2004), showing quite large differences, ranging from -100 to $100 \mu\text{K}$. Similarly, Park et al. (2007) compare their PILC3 map with the WILC3 and TILC3 maps at 1.4 degree resolution, showing differences in excess of $40 \mu\text{K}$. In the following, discrepancies between these various solutions are further investigated.

As a first step, we evaluate by how much the various maps at one degree resolution disagree over the full sky. The top panel

Standard Deviation of EILC1, TILC3, PILC3, WILC5 and KILC5 (1 deg. resolution)



Standard Deviation of EILC1, TILC3, PILC3, WILC5 and KILC5 (3 deg. resolution)

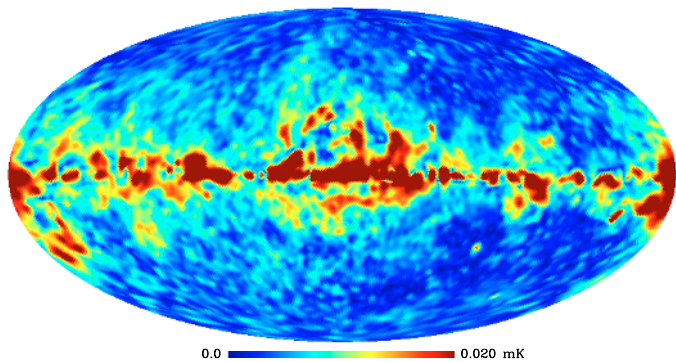


Fig. 2. *Top:* the standard deviation, per pixel for $n_{\text{side}} = 512$, of the EILC1, TILC3, PILC3, WILC5 and KILC5 maps, at 1° resolution. *Bottom:* same at 3° resolution. Note the different color scales for the two panels.

of Fig. 2 shows the pixel based standard deviation of five such maps. In this evaluation, we exclude the EGS3 map (which is at 3 degree resolution, and is the source of additional variance in the Kp2 sky mask, where the CMB is estimated at even lower resolution).

All maps are obtained from observation of the same sky (sometimes starting from the same data set) and are smoothed to the same resolution (one degree) where noise is small. Discrepancies originate essentially from systematic differences in the methods. In the central regions of the galactic plane, discrepancies are significantly higher than $50 \mu\text{K}$ (with bright spots above $90 \mu\text{K}$). In other sky regions, they are typically in the $10\text{--}20 \mu\text{K}$ range, except on compact spots again, where they are above $50 \mu\text{K}$. The latter are probably due to residuals of the emission from strong compact sources. The discrepancy between the maps is greater in the ecliptic plane, which is a signature of the impact of the instrumental noise, uneven because of the WMAP scanning strategy. This is due to the different attention given to minimising the contribution of instrumental noise on small scales (rather than foreground emissions) in the final map (and to a much lesser extent to the difference of noise level between the various WMAP releases).

The bottom panel of Fig. 2 focuses the comparison to larger scales (3 degree beam). At high galactic latitudes, the discrepancies are below $10 \mu\text{K}$ except for a few localised regions (LMC, Ophiuchus complex) where they reach about $30 \mu\text{K}$. Differences in the North Polar Spur, at the level of $10\text{--}15 \mu\text{K}$, are also clearly visible. Close to and inside the galactic plane, systematic discrepancies significantly exceed $30 \mu\text{K}$.

Including the EGS3 map in the comparison at high galactic latitudes does not change these conclusions. Pairwise comparisons of the available maps typically show the same level of discrepancies, which indicates that the variance of the solutions is not due to one single map being in strong disagreement with the others.

The conclusion of this comparison is that foreground residuals exist in the published CMB maps at the level of about 50 to $100 \mu\text{K}$ in the galactic ridge, 20 to $50 \mu\text{K}$ at low galactic latitudes, and 10 to $20 \mu\text{K}$ at higher galactic latitudes (above 30 degrees).

This observation calls for localised weightings, adapted to local properties of the foreground and the noise. This, however, is not easily compatible with the recovery of the largest modes of the CMB map, as pointed out before by Eriksen et al. (2004), and as demonstrated by the discontinuities observed between the CMB solutions in the different regions when the sky is cut, as in the TILC, EILC1, and WILC maps.

Our approach, then, will be to vary the relative weightings on small scales for small scale CMB reconstruction, and keep the weighting uniform over large regions of the sky for the recovery of the largest scales. This can be achieved quite straightforwardly by using the spherical needlets discussed in Sect. 4.

4. The ILC needlet method

4.1. The choice of the ILC

It is striking that all the presently available full sky CMB maps derived from an analysis of the WMAP data have been obtained by an implementation of the ILC method.

The ILC, indeed, has many advantages:

- the method relies only on two very safe assumptions: the CMB emission law, and the fact that the CMB template is not correlated to foreground emission³;
- under these assumptions, the method minimises the empirical variance of the reconstruction error;
- the ILC is very simple in implementation.

The ILC has two major drawbacks.

- As noted for instance by Hinshaw et al. (2007), Delabrouille & Cardoso (2007), Saha et al. (2007), empirical correlations between the CMB and the source of contamination results in a bias; this bias is discussed in more detail below;
- In the absence of a model of the contaminants (foregrounds and noise), it is not possible to predict the reconstruction errors, which somewhat annihilates the benefit of making very safe assumptions about the properties of the data set.

4.2. The ILC bias

The existence of a “bias” in maps obtained by an ILC method is a well established fact. The derivation of this bias (which is, in fact, the systematic cancelling of a fraction of the projection of the CMB map onto the vector space spanned by the noise realisations for all the considered input maps), is given in the Appendix.

The order of magnitude to keep in mind is that about $(m - 1)$ “modes” of the original CMB, out of N_p , are cancelled by

³ In reality, it is likely that the CMB map actually *is* somewhat correlated to the foreground emissions (extragalactic point sources and SZ effect), because of the ISW effect. The implication of this is not studied further in the present paper.

the ILC, where m is the number of channels used, and N_p the number of independent pixels or modes in the regions for which the ILC is implemented independently. Note however that when the signal is strongly correlated between pixels, the bias can be significantly larger – see the appendix for details.

The practical consequences are:

- a loss of CMB power, which has to be taken into account for power spectrum estimation;
- an anti-correlation of the map reconstruction error with the real CMB sky.

The level of the bias induced by our method is investigated both theoretically (Appendix A), and through Monte-Carlo simulations.

4.3. Needlets

A *frame* is a collection of functions with properties close to those of a basis. *Tight frames* share many properties with orthonormal bases, but are redundant (see Daubechies 1992, for details).

Needlets were introduced by Narcowich et al. (2006) as a particular construction of a wavelet frame on the sphere. They have been studied in a statistical context (e.g. Baldi et al. 2008; Baldi et al. 2007) and have also been used recently for cosmological data analysis problems (e.g. Pietrobon et al. 2006). The most distinctive property of the needlets is their simultaneous perfect localisation in the spherical harmonic domain (actually they are spherical polynomials) and potentially excellent localisation in the spatial domain.

We recall here the definition and practical implementation of the needlet coefficients, following the generalised formulation by Guilloux et al. (2008). Let $h_\ell^{(j)}$, $j \in \mathcal{J}$ be a collection of window functions in the multipole domain, indexed by j . Suppose that for each scale j , $\xi_k^{(j)}$ is a grid of points (indexed by $k \in K^{(j)}$) which satisfies an exact⁴ quadrature condition with weights $\lambda_k^{(j)}$. The needlets are axisymmetric functions defined by

$$\psi_k^{(j)}(\xi) = \sqrt{\lambda_k^{(j)}} \sum_{\ell=0}^{\ell_{\max}} h_\ell^{(j)} L_\ell(\xi \cdot \xi_k^{(j)}), \quad (6)$$

where the functions L_ℓ denote the Legendre polynomials.

Any square integrable function f on the sphere can be analysed by the scalar products $\beta_k^{(j)} := \langle f, \psi_k^{(j)} \rangle$ of the function f with *analysis needlets*. All the needlet coefficients of scale j are advantageously computed in the spherical harmonic domain, as the evaluation at points $\xi_k^{(j)}$ of a function whose multipole moment are simply $h_\ell^{(j)} a_{\ell m}$. These needlet coefficients, denoted $\gamma_k^{(j)}$, are given by:

$$\gamma_k^{(j)} = \sqrt{\lambda_k^{(j)}} \beta_k^{(j)}.$$

Each field of needlet coefficients can in turn be convolved with some *synthesis needlets*

$$\tilde{\psi}_k^{(j)}(\xi) = \sqrt{\lambda_k^{(j)}} \sum_{\ell=0}^{\ell_{\max}} \tilde{h}_\ell^{(j)} L_\ell(\xi \cdot \xi_k^{(j)}), \quad (7)$$

using the same procedure and leading to the map $X^{(j)}$ whose multipole moments are $h_\ell^{(j)} \tilde{h}_\ell^{(j)} a_{\ell m}$. The analysis and synthesis operations are summed up as: *Analysis*:

$$X \xrightarrow{\text{SHT}} a_{\ell m} \xrightarrow{\times} h_\ell^{(j)} a_{\ell m} \xrightarrow{\text{SHT}^{-1}} \gamma_k^{(j)}.$$

⁴ Or almost exact, for all practical purposes.

Synthesis:

$$\gamma_k^{(j)} \xrightarrow{\text{SHT}^{-1}} h_\ell^{(j)} a_{\ell m} \xrightarrow{\times} \tilde{h}_\ell^{(j)} h_\ell^{(j)} a_{\ell m} \xrightarrow{\text{SHT}} X^{(j)}.$$

Double arrows denote as many transforms as scales in \mathcal{J} . If X is band-limited to $\ell \leq \ell_{\max}$ and if the reconstruction condition $\sum_j h_\ell^{(j)} \tilde{h}_\ell^{(j)} = 1$ holds for all $\ell \leq \ell_{\max}$, then the complete process yields a decomposition of X in smooth maps, namely

$$\forall \xi, \quad X(\xi) = \sum_j X^{(j)}(\xi). \quad (8)$$

Note that the existence of a fast inverse spherical harmonic transform using the quadrature points $\xi_k^{(j)}$ is required in practice, and that HEALPix pixels and weights fulfil the quadrature condition only approximately. Further details can be found in Guilloux et al. (2008), with an extensive discussion on the choice of the spectral window functions.

A key feature of the needlet decomposition follows from the localisation of the analysis functions which allows for localised processing (such as denoising, signal enhancement, masking etc.) in the needlet coefficient domain, i.e. applying some non-uniform transforms to the coefficients $\gamma_k^{(j)}$. Other types of wavelets (for instance, the steerable wavelets of Wiaux et al. 2008, or the wavelets, ridgelets and curvelets of Starck et al. 2006) could also be used for localised processing on the sphere, although the quality of the localisation depends much on the details of the wavelet design. Needlets are compactly supported in the multipole domain and can be further designed to be well localised in the direct domain according to various criteria. This permits one to work on full sky data without real need for masking the galactic ridge.

4.4. The method

The method implemented in this work, and applied both to simulations and to the real WMAP data sets (for all releases), consists of the following steps:

- we start with the data set consisting of band-averaged temperature maps from WMAP (simulated or real data), to which we add the IRIS 100 micron map;
- WMAP-detected point sources are subtracted from the WMAP maps;
- we apply a preprocessing mask, in which a very small number of very bright compact regions are blanked (see Table 2); blanked regions are filled-in by interpolation; this is done only on the real WMAP data (not on simulations);
- all maps are deconvolved to the same resolution (that of the W channel⁵); this operation is performed in harmonic space;
- maps are analysed into a set of needlet coefficients $\gamma_k^{(j)}$ following the method described in 4.3;
- for each scale, the covariance matrix \widehat{R} of the observations is computed locally (using an average of 32×32 needlet coefficients);
- the ILC solution is implemented for each scale in local patches;
- an output CMB map is reconstructed from the ILC filtered needlet coefficients; this map constitutes our main CMB product;

⁵ A noise weighted average beam is obtained from the W1 W2 W3 W4 beam coefficients provided by the WMAP team.

- that map is Wiener-filtered in harmonic space, to make an alternate CMB map with lower integrated error (our best guess CMB map);
- in parallel, the actual ILC filter used on the analysed data set is applied to 100 different simulations of the WMAP noise, to estimate the noise contribution to the final map;
- the level of the biasing, which depends on the geometry and not much on the actual templates of CMB and foreground and noise, is estimated on a set of fully simulated data.

Each of these steps is described in more detail below.

4.5. Point source subtraction

Strong point sources in the input data set typically leave detectable residuals in the output ILC map, and modify at the same time the estimation of the background, leading to less rejection of other contaminants. On the other hand, their specific shape usually allows effective estimation and removal by other methods. For the purpose of this study we used information from the WMAP source catalogue (Hinshaw et al. 2007) which provides characterisation for all point sources detected above a 5σ threshold away from the galactic plane.

For all sources identified, we subtract from the input maps a Gaussian profile at the given position and with the given flux. Conversion factors between flux density and Gaussian amplitude, as well as the FWHM of the Gaussian profile are taken from Table 5 of Page et al. (2003) for one year data, and corresponding updates for the more recent releases.

For the simulated data set, the subtraction of detected point sources is mimicked by removing from the simulation all sources above 1 Jy (independently in all channels).

4.6. Blanking of compact regions

In addition to the point sources subtracted above, some compact regions of strong emission (mostly in the galactic plane) exceed the rejection capabilities of the method used in this analysis, because they are too local and/or too specific. Their contribution in the wings of the needlets also contaminates the solution far from the centre of the sources. Those sources cannot be satisfactorily subtracted in the same way as the previous ones, either because they are not strictly speaking point-like, or because they are bright enough that small departures of actual beam shapes from the Gaussian model used in the subtraction step leave significant residuals. As they represent only a very tiny fraction of the sky (we single out eleven such sources), we blank out these regions in all WMAP channels, cutting out circular patches adapted to the size of the beam and of the source. Table 2 gives the list of those regions with their main characteristics.

To reduce local pollution of the needlet coefficients by the sharp cut, the small blanked regions are filled in by a smooth interpolation, so that fluctuations at a larger scale than the hole size are at least coarsely reconstructed. More precisely, interpolation is made by diffusion of the boundary values inside the hole.

Although this masking and interpolation has no reason to be optimal, it is an efficient way of reducing the impact of very strong sources on their environment. The CMB inside the masked patches is recovered (to some extent) both by the interpolation of original maps (which avoids sharp discontinuities) and by the needlet decomposition and ILC reconstruction. The masked region is tiny: 0.058% of the sky in the K channel (the most affected).

Table 2. List of compact regions blanked in the pre-processing step.

Name	Galactic coordinates	Type
Crab neb	184.5575 –05.7843	SNR
sgr A	000.064 +00.147	Radio-Source
sgr B	000.599 +00.002	Radio-Source
sgr C	359.4288 –00.0898	HII region
sgr D	001.131 –00.106	Molecular cloud
Orion A	209.0137 –19.3816	HII region
Orion B	206.5345 –16.3539	Molecular cloud
Omega neb	015.051 –00.674	HII region
Cen A	309.5159 +19.4173	QSO
Cas A	111.735 –02.130	SNR
Carina neb	287.6099 –00.8542	HII region

A circular patch centered on the source, of radius 75, 55, 45, 45, 34 and 35 arcmin for the K, Ka, Q, V, W and IRIS 100 μ m bands respectively, is masked. The masked regions are then filled with an extrapolation of edge values.

Table 3. Spectral bands used for the needlet decomposition in this analysis.

Band (j)	ℓ_{\min}	ℓ_{\max}	nside(j)
1	0	15	8
2	9	31	16
3	17	63	32
4	33	127	64
5	65	255	128
6	129	511	256
7	257	767	512
8	513	1023	512
9	769	1199	512

Needlet coefficient maps are made at different values of nside, given in the last column.

4.7. Needlet decomposition

The original observations (WMAP and IRIS) are decomposed into a set of filtered maps represented by their spherical harmonic coefficients:

$$a_{\ell m}^{(j)} = h_{\ell}^{(j)} a_{\ell m} \quad (9)$$

where $a_{\ell m}$ are the spherical harmonic coefficients of the original map, and $a_{\ell m}^{(j)}$ those of the same map filtered by the window function j . Needlet coefficients $\gamma_k^{(j)}$ are obtained as the value of the filtered map at points ξ_k .

For each scale j , the coefficients $\gamma_k^{(j)}$ are computed on a HEALPix grid at some value of nside, compatible with the maximum value of ℓ of band j . We use for nside(j) the smallest power of 2 larger than $\ell_{\max}/2$, with a maximum of 512. Details about the bands used are given in Table 3 and Fig. 3.

4.8. ILC implementation on needlet coefficients

The general idea is to independently implement the ILC on subsets of the needlet coefficients $\gamma_k^{(j)}$. For a given scale, these coefficients come in the format of a set of HEALPix maps (one per frequency channel). The ILC is implemented locally in space and locally in ℓ as follows.

Covariance matrices $\mathbf{R}_k^{(j)} = \langle \gamma_k^{(j)} \gamma_k^{(j)T} \rangle$ for scale j at pixel k are estimated as the average of the product of the computed needlet coefficients over some space domain \mathcal{D}_k . Because of this, there is a trade-off between localisation and accuracy of estimation. A better estimate of the true covariance is obtained by

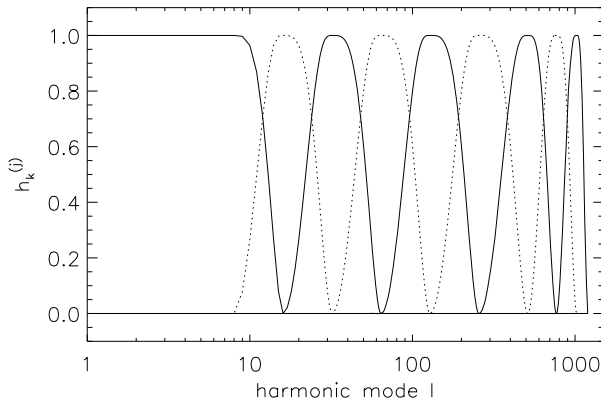


Fig. 3. The spectral bands used in this work for the definition of the needlets.

averaging the products of needlet coefficients between all pairs of maps over a large area, but this provides an estimate of the average covariance over that area. In practice, for the present analysis, we make use of the hierarchical property of the HEALPix pixelisation, and compute covariance matrices as the average in larger pixels, corresponding to a HEALPix pixelisation with $n_{\text{side}} = n_{\text{side}}(j)/32$. This provides a computation of the statistics by averaging $32^2 = 1024$ samples, which results in a precision of the order of 3 per cent for all entries of the $\mathbf{R}_k^{(j)}$ matrix. It implies an ILC bias of order $5/1024$ (for $m = 6$ channels and $N_p = 1024$ coefficients per domain on which the ILC filter is estimated independently (see Appendix for details). Choosing a smaller area results in excessive error in the covariance estimates, and hence excessive bias. Choosing a larger area results in less localisation, and hence some loss of efficiency of the needlet approach.

We denote as $\widehat{\mathbf{R}}_{k' \in \mathcal{D}_k}^{(j)}$ the estimate of $\mathbf{R}_k^{(j)}$ obtained by averaging the value of $\boldsymbol{\gamma}_{k'}^{(j)} \boldsymbol{\gamma}_{k'}^{(j)T}$ in domain \mathcal{D}_k .

On the largest scales ($\ell \leq 50$), the typical angular extent of a needlet is larger than 5 degrees, and the value of n_{side} for the map of needlet coefficients is less than 32. Covariance matrices are then computed on the full sky rather than on the largest possible HEALPix grid, i.e. \mathcal{D}_k is the complete sky, rather than one of the 12 basis Healpix pixels.

Using these covariance matrices, the ILC is implemented using Eq. (4) for each domain. The estimated CMB needlet coefficients are:

$$\left[\widehat{\boldsymbol{\gamma}}_k^{(j)} \right]_{\text{CMB}} = \frac{\mathbf{a}^T \left[\widehat{\mathbf{R}}_{k' \in \mathcal{D}_k}^{(j)} \right]^{-1}}{\mathbf{a}^T \left[\widehat{\mathbf{R}}_{k' \in \mathcal{D}_k}^{(j)} \right]^{-1} \mathbf{a}} \boldsymbol{\gamma}_k^{(j)}. \quad (10)$$

4.9. Full map reconstruction

The full CMB map reconstructed from this set of needlet coefficients is our basic needlet ILC (NILC) CMB map.

4.10. Final Wiener filtering

For a number of purposes, in particular subtraction of an estimate of the CMB to study other emissions, it is interesting to use, instead of our ILC map at the resolution of the WMAP W channel, a map with minimal error. Such a map is obtained from the ILC map by one-dimensional Wiener filtering.

As a last processing step towards a minimum variance CMB map, we thus Wiener-filter our CMB map, to get rid of the large noise contamination at high ℓ . The Wiener filter is performed in harmonic space as described in 2.3.2.

The harmonic Wiener filter is given by formula 5, i.e. $w_\ell = b_\ell^2 C_\ell / (b_\ell^2 C_\ell + N_\ell)$. For its implementation, we need to know the relative power of CMB and noise. We assume that the best fit CMB power spectrum of the WMAP team is correct, hence C_ℓ is known. The beam factor b_ℓ is assumed perfectly known as well. The denominator $b_\ell^2 C_\ell + N_\ell$ can be estimated directly as the power spectrum of our output needlet ILC map.

In practice, we smooth the power spectra with $\delta\ell/\ell = 0.1$ to lower the variance of the power spectrum estimator on the output needlet ILC map. Even with this, the filter is poorly estimated for low modes, because of the large cosmic variance. As can be seen in Fig. 5, the signal to noise ratio of our reconstructed map is expected to be quite high at low ℓ . Therefore, the Wiener filter for low modes is expected to be very close to 1. For this reason, we set $w_\ell = 1$ for $\ell < 200$, and use a linear interpolation between $\ell = 200$ and $\ell = 250$.

4.11. Noise level estimate

The level of noise contamination (variance per pixel, and average power spectrum) in the output map is estimated by Monte-Carlo simulations, using the average of 100 realisations of the WMAP noise maps. For each initial set (i) of five noise maps (one noise map per WMAP channel), a single output noise map $n_p^{(i)}$ is obtained by performing the needlet decomposition of the initial noise maps, and filtering needlet coefficient maps with the same filter as that used on the single full simulated data set.

Denoting as $n_p^{(i)}$ and $n_{\ell m}^{(i)}$ respectively the pixel value and the harmonic space value of the noise map number i , we compute:

$$\sigma_p^2 = \frac{1}{N_i} \sum_i (n_p^{(i)})^2 \quad (11)$$

and

$$\sigma_\ell^2 = \frac{1}{N_i(2\ell + 1)} \sum_i \sum_m (n_{\ell m}^{(i)})^2. \quad (12)$$

These are respectively estimates of the noise pixel variance and of the noise power spectrum of our final map.

4.12. Bias estimates

The impact of the ILC bias is estimated by Monte-Carlo simulations on simulated data sets. The corresponding error is of the order of 2% of the CMB.

4.13. Practical implementation

The practical implementation of this processing pipeline is made essentially using the octave language (the free software version of Matlab). The analysis is done in the framework of the pipeline tool developed by the ADAMIS team at the APC laboratory. This tool provides a flexible and convenient web interface for running our data analysis on simulations or real data with easy handling and tracing of the various pipeline options⁶. Single runs of the full pipeline require less than half an hour on a single processor

⁶ See http://www.apc.univ-paris7.fr/APC_CS/Recherche/Adamis/ in the “outreach” section.

of a standard desktop computer (dominated by harmonic transforms), whereas numerous pipelines on simulated data sets for Monte-Carlo are run on the ADAMIS 88-processor cluster, optimised for efficient I/O.

5. Simulations

5.1. The simulated data

We start with a validation of our method on simulated data sets. For this experiment, synthetic observations of the sky emission are generated using the Planck Sky Model (PSM). The PSM is a flexible software library, designed for simulating the total sky emission in the 10–1000 GHz frequency range, and developed as part of the foreground modelling activities of the Planck working group on component separation (Planck WG2). Sky emission comprises galactic components of four origins (free-free, synchrotron, thermal dust, and spinning dust, with spectral emission laws for dust and synchrotron varying from pixel to pixel), CMB, kinetic and thermal SZ effects, and the emission from a population of galactic and extragalactic point sources which includes radio sources, infrared sources, and an infrared background. Although not perfect, the model sky is thought to be sufficiently representative of the complexity of the real sky emission for our simulations to be meaningful.

Sky maps are produced at WMAP central frequencies for the K, Ka, Q, V and W band, and convolved in harmonic space with approximate WMAP instrumental beams (Gaussian symmetric beams are used for these simulations). Uncorrelated, non-stationary Gaussian noise is added, with a pixel variance deduced from the WMAP sensitivity per channel and effective hit count. To mimic the subtraction of the brightest point sources detected by WMAP, we remove from the model sky, at each frequency, all sources with flux above 1 Jy (assuming they would have been detected, and can be subtracted from the data set). The 11 compact regions listed in Table 2 however, being specific to the real sky, are not blanked for the simulations.

Although these simulations provide only an approximation of the real WMAP data sets, they are representative enough that the simulated data offer a component separation challenge close to that of the real data set. The IRIS map is used as part of the full set of data for the ILC implementation on simulations.

5.2. Results

Figure 4 shows the input simulated CMB, the output CMB, and the difference of the two for one particular simulation. The reconstruction is visually good except in regions of local strong galactic emission (in the galactic ridge, for example). This is to be expected: not only can the method not remove foreground emission perfectly, but in addition the price to pay to remove the foreground (even imperfectly) is more noise (because of sub-optimal weighting of the observations as far as noise contamination is concerned).

A more quantitative estimate of the level of contamination of the CMB map by foreground emissions and noise is obtained by looking at power spectra. Figure 5 shows the input simulated CMB power spectrum (dotted line), the spectrum of the map of residuals (difference between output and input, dashed line), both full sky (top panel) and in the HGL region (bottom panel). The angular power spectrum of the residual map is seen to be small compared to the CMB power on large scales, the two being comparable at $\ell \simeq 500$. Noise dominates on smaller scales.

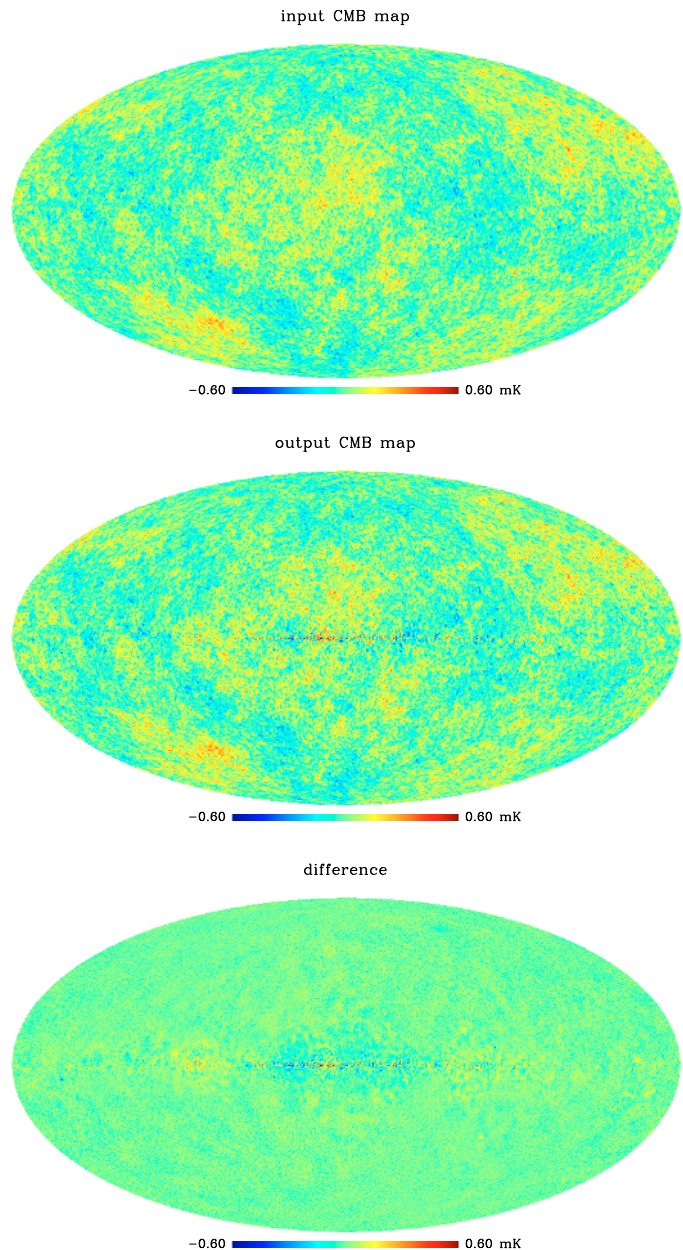


Fig. 4. *Top:* the simulated input CMB map. *Middle:* the reconstructed CMB. *Bottom:* the difference (output-input), displaying the residuals left by the method. All three maps share the same colour scale, and are at the resolution of the WMAP W channel.

The residuals due to the presence of galactic emission are seen to contribute power essentially below $\ell = 400$, where the power of the difference map is seen to be slightly higher in the full sky power spectrum than in the HGL power spectrum (this is visible, in particular, at the top of the first acoustic peak).

5.3. Bias

As discussed in 4.2, we expect a (small) bias in the ILC map, due to empirical correlations between the CMB emission and contaminants (including noise and foreground emission). This is not particular to our approach, and is expected for any ILC method. For better characterisation of our output map, we evaluate the effect both theoretically (in the Appendix) and numerically.

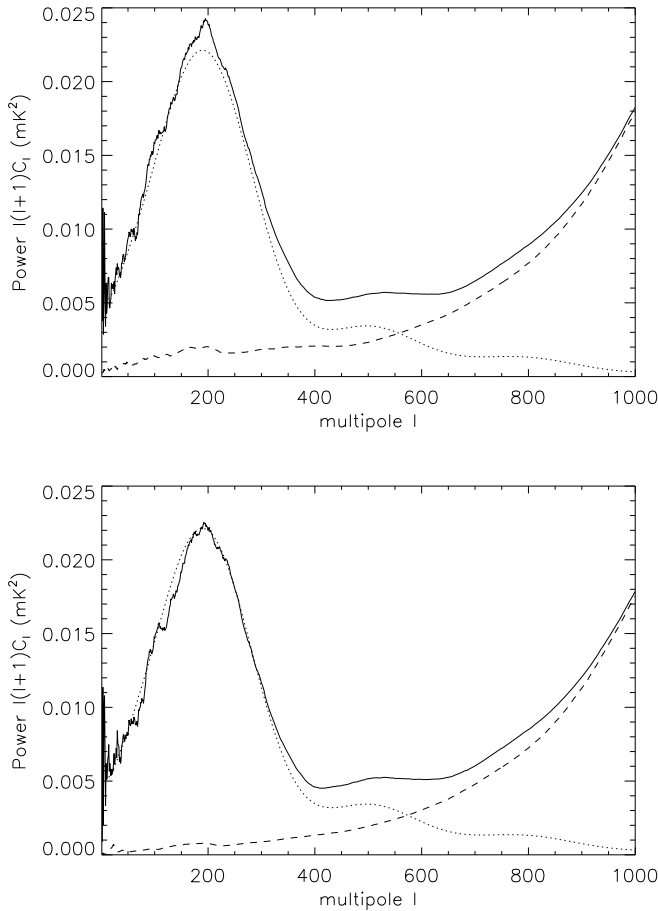


Fig. 5. *Top:* for simulated data sets, full sky power spectra of the output CMB map (plain line), and the difference map (dashed line). The CMB model used in the simulation is over-plotted as a dotted line. *Bottom:* same at high galactic latitude only (HGL region).

Although a general analytic estimate of the bias is complicated, Appendix A shows that (to first order at least and for reasonable assumptions about the CMB, foreground emissions, and noise) the amplitude of the effect does not depend much on what the actual foreground emissions are in detail, but is set essentially by the geometry of the domains considered (through a number of effective modes). It is then possible to estimate the amplitude of the effect by Monte-Carlo simulations on synthetic data sets resembling the actual WMAP observations.

Figure 6 illustrates an estimate of the bias $b(\ell)$ as a function of the harmonic mode, computed as a fractional error:

$$b(\ell) = \frac{\sum_m (\widehat{a}_{\ell m} - a_{\ell m}) a_{\ell m}^*}{\sum_m |a_{\ell m}|^2} \quad (13)$$

where $a_{\ell m}$ are the harmonic modes of the input CMB map, and $\widehat{a}_{\ell m}$ the harmonic modes of the output CMB map. The numerator in this equation computes the covariance of the residual map and the input map as a function of ℓ , and the denominator is a normalisation factor. For an error uncorrelated with the input, $b(\ell)$ should be close to 0 on average. Analytical estimates of the effect (see Appendix) suggest a bias of the order of 2% for our implementation (taking into account mode correlation). The numerical estimate of Fig. 6, obtained as the average bias for 500 simulated data sets, is in good agreement with this prediction, with slight variations due to varying numbers of effective modes for

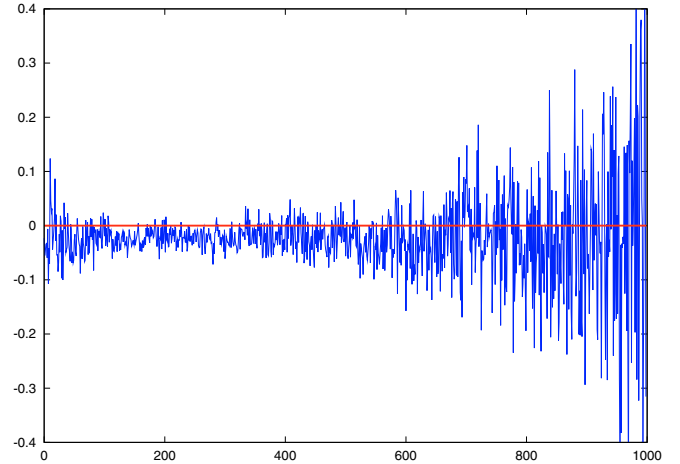


Fig. 6. The fractional bias in the ILC map as a function of ℓ , for one single simulation. This figure is illustrative both of the amplitude of the effect, and on its variance for one single realisation. Bias, and standard deviation of the bias, are of the same order of magnitude for most of the ℓ range.

different needlet scales. The average value of $b(\ell)$ between $\ell = 2$ and 1000 in that simulation is about -2.2% .

6. Application to WMAP data

6.1. ILC result

We now turn to the description of the results obtained on the real WMAP data sets. In order to facilitate the comparison with existing maps, we process independently one year, three year and five year data, to obtain three CMB maps (hereafter NILC1, NILC3 and NILC5). For each year, we use the beam estimates, noise level, and point source catalogue provided with the corresponding release.

The improvement of CMB reconstruction with consecutive data releases is illustrated in Fig. 7, which shows the full sky power spectra of the NILC CMB maps obtained with one year, three year, and five year WMAP data. The power spectra displayed are the raw power spectra of the output map, computed for the full sky, and smoothed with a variable window in ℓ of 10% width. While the lower part of the spectrum, cosmic variance limited and CMB dominated, does not change much, the high ℓ spectrum of the map, dominated by noise, decreases substantially with increasing observation time – as expected. The excellent agreement at low ℓ (up to $\ell \simeq 300$) between the power spectra and the model is striking. The bumps in the spectrum due to the first and second acoustic peaks are clearly visible on the five year map spectrum.

Our full-sky reconstructed CMB map for the five year observations, at the resolution of the W-channel, is displayed in the top panel of Fig. 8.

7. Discussion

7.1. Comparison with other maps

A full comparison of our needlet ILC maps (for all data releases) with all the other available maps would be too long for the present discussion. Rather, we decide to compare our five year result only with the TILC3 map on small scales (choice is motivated by the fact that the TILC is the only other full sky high resolution map available), and with the EGS3 map on large scales.

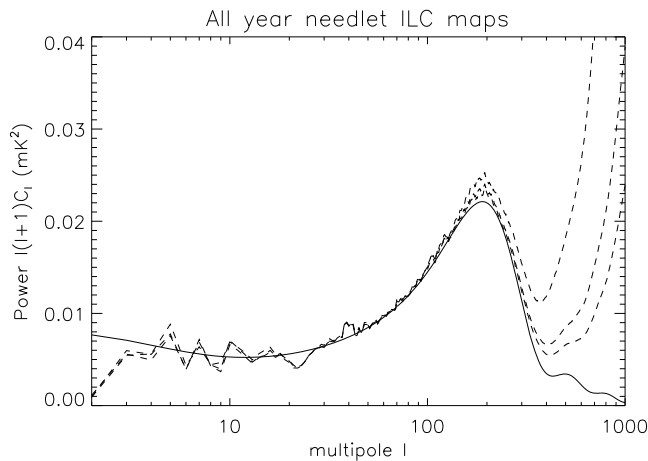


Fig. 7. Power spectra of the NILC map for one year, three year, and five year WMAP data.

This is of particular interest, as the EGS3 is the only map obtained with a method not based on the ILC, and also is a method specifically implemented for the recovery of the largest scales.

7.1.1. Comparison at the pixel level – small scales

On the smallest scales, we compare our needlet ILC map with the TILC and with the WMAP foreground-reduced *W* band. Figure 10 shows local regions of the foreground-reduced map, the NILC5 map, and the TILC3 map, in the galactic plane and at the galactic pole. Our needlet map is clearly less contaminated by galactic emission than the other two. At high galactic latitude, the NILC5 and TILC3 are visually comparable, while the foreground-reduced map appears to be more noisy, as expected.

The power spectrum of the output map for the five year data (NILC5 map), for three different sky coverages, is shown in the bottom panel of Fig. 9. On the same panel, we plot the angular power spectrum C_ℓ corresponding to the WMAP best fit model, corrected for the *W*-channel beam. On the top panel of the same figure, we show the same power spectrum estimates for the TILC3 map. This shows the improvement achieved by our method close to the galactic plane. This improvement is due both to the needlet approach and to the use of the IRIS map to help with dust subtraction. As seen in Fig. 7, the difference in quality between NILC5 and TILC3 cannot be explained solely by reduced noise (NILC5 and NILC3 being very close in quality for all scales except the smallest ones).

7.1.2. Comparison at the pixel level – large scales

Figure 11 gives a visual comparison of NILC5 (this work) and the EGS3 (Eriksen et al. 2007), as well as of NILC5 and KILC5 (Kim et al. 2008). In the top row, we display the EGS3 and KILC5 at a resolution of 3 degrees, and degraded to $n_{\text{side}} = 64$. The bottom row shows the difference between our needlet ILC solution (displayed on the bottom panel of Fig. 8) and these two maps.

The most striking difference between the five year needlet ILC map and the EGS3 is in the galactic plane, where the EGS3 does not recover the intermediate angular scales. The difference, however, shows no particular clear structure, as expected if it is the random realisation of a Gaussian random field. It is thus probably essentially due to the difference between our CMB

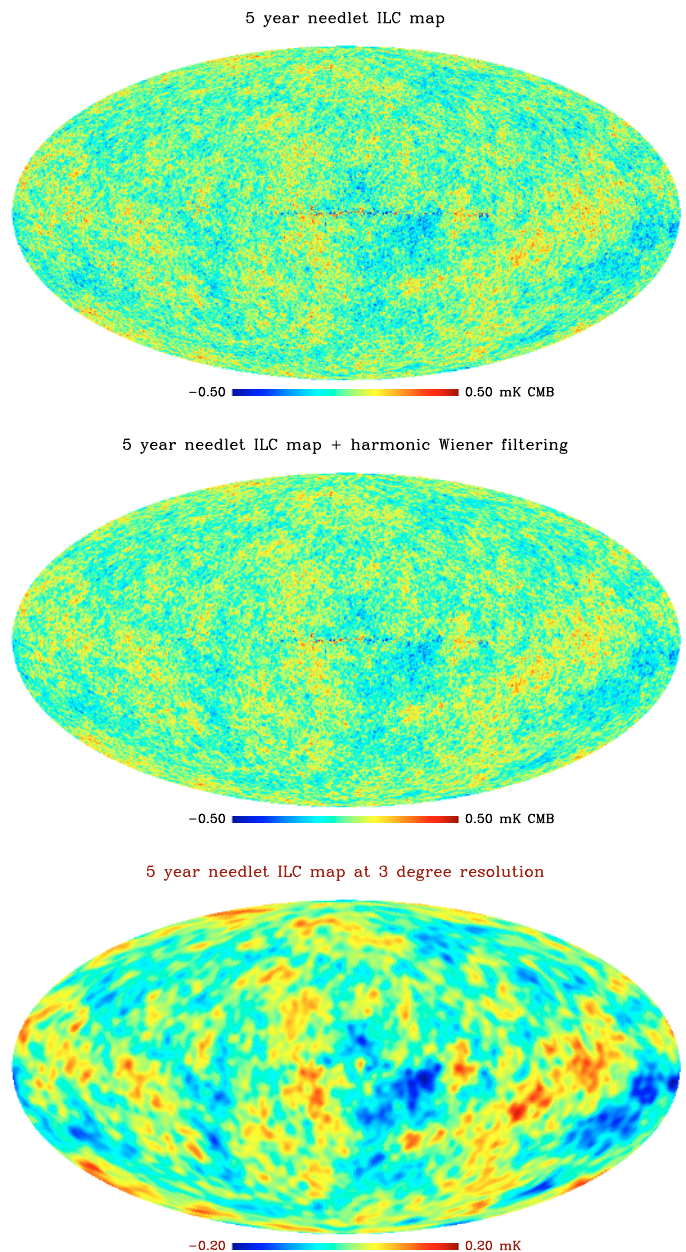


Fig. 8. *Top:* the NILC5 reconstructed WMAP CMB at the resolution of the *W* channel. *Middle:* the harmonic Wiener NILC5 CMB map. *Bottom:* the NILC5 CMB map at 3 degree resolution.

reconstruction on scales larger than 3 degrees, and the CMB on larger scales that can be inferred from the CMB reconstructed by Eriksen et al. outside of their galactic mask. At higher galactic latitude, the two maps are in good agreement, with no obvious feature which could be correlated to foreground emissions or to the CMB itself, with the exception of a hot spot in the large Magellanic cloud (which might be residual emission of the LMC in our map, as Eriksen et al. (2007) actually mask the centre of the LMC and obtain a solution in the direction of the LMC by extrapolation from nearby pixels). Above 30 degree absolute galactic latitude, the rms of the difference between our 3 degree map and the EGS3 map is $5.7 \mu\text{K}$. The two maps are in much better agreement than the EGS3 and the WMAP MEM model maps (see Fig. 3 in Eriksen et al.). Note however that theoretically, if there were no foreground in the WMAP data, the noise standard

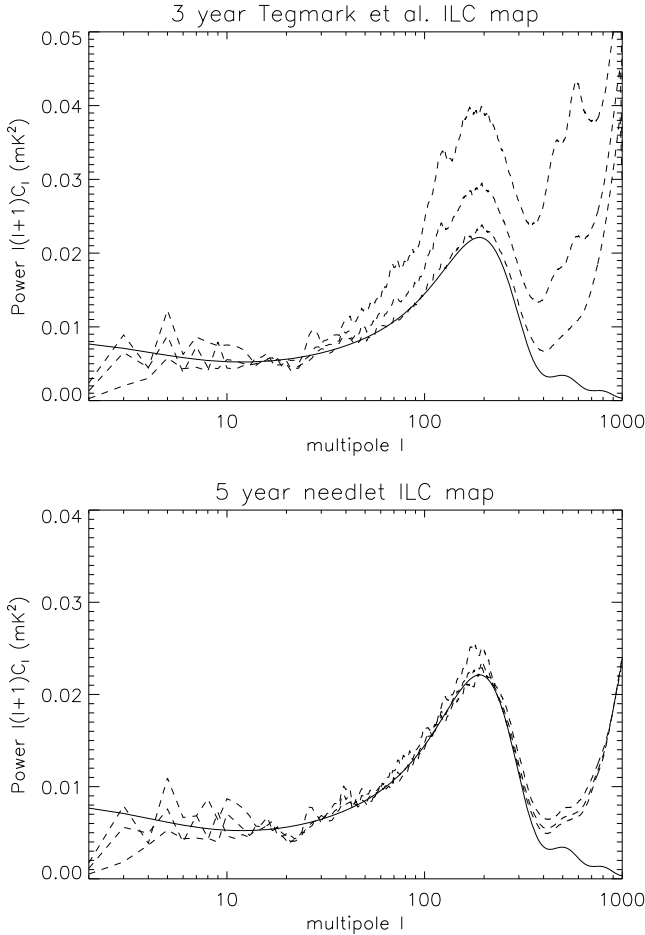


Fig. 9. Power spectrum of the reconstructed WMAP CMB map. For each of the two panels, the CMB best fit model is shown as a solid black line, and power spectra computed at low galactic latitudes (using the LGL mask), on the full sky (no mask), and at high galactic latitudes only (HGL mask) are displayed as dashed lines. Note that the map spectra plotted here are directly those of the maps, without any correction for the beam. *Top panel:* ILC map of (Tegmark et al. 2003, TILC3). *Bottom panel:* this work, with five year WMAP data. The power spectrum of the needlet ILC CMB is significantly more homogeneous than the power spectrum of the TILC3 map. We interpret this difference as an indication that the TILC3 map is significantly more contaminated by residuals of galactic emission. Note the different scales of the y -axis, and the improvement on small scales, with a noise power of about 0.024 mK^2 at $\ell = 1000$ for the NILC5 map at high galactic latitude, instead of about 0.040 mK^2 for TILC3. As indicated by Fig. 7, this is due essentially to the better quality of the five year release, since the NILC3 map also has a noise power spectrum of about 0.040 mK^2 at $\ell = 1000$.

deviation σ_n on a 3 degree map obtained by noise-weighted averaging using all WMAP channels would be about $3.2 \mu\text{K}$ for three year data. If instead we assumed that only the three highest frequency channels are free of foreground contamination, σ_n would be $4.4 \mu\text{K}$.

The difference between our map and the KILC5 map is more systematic with, in particular, stronger differences in the galactic plane, in spite of the fact that the two methods work on the same input data set. A careful visual inspection of the CMB maps themselves gives the impression that the KILC5 map is probably systematically negative towards the galactic central regions. There is, however, no secure way to be certain which map is best.

7.2. Map characterisation

7.2.1. Beam

The effective beams of the reconstructed maps are plotted in Fig. 12, for both the full resolution five year needlet ILC map, and for the Wiener-filtered version. The map has been reconstructed for the range $0 \leq \ell \leq 1200$, with a smooth transition of the response, between ℓ of 1024 and 1200, from the nominal W band beam value to 0. This smooth transition allows us to avoid ringing effects which happen in the case of a sharp cut-off in ℓ . The ratio of the Wiener beam (dashed line) and ILC beam (solid line) gives a measure of the signal to noise ratio in each mode.

7.2.2. Instrumental noise

Given the ILC filter computed on the real data set, the level and properties of the instrumental noise can be straightforwardly computed by applying the *same* filter to simulated WMAP noise maps. From 100 noise realisations, we compute the average full-sky noise power spectrum (Fig. 13), as well as the noise standard deviation per pixel (Fig. 14). Noise properties are not as simple as one may wish: the noise is non stationary, because of both the uneven sky coverage and of the localised processing. It is also somewhat correlated pixel-to-pixel, in particular close to the galactic plane. This is unavoidable, but fortunately our pipeline allows us to make as many Monte-Carlo realisations of the instrumental noise as needed for any scientific study made using our needlet ILC map. One hundred such simulations are made available as part of our main data products.

7.2.3. Foreground residuals

More problematic is the evaluation of the contamination of the CMB map by foreground residuals. It requires prior information about the foregrounds, which the ILC method avoids using. An indication of the level of systematic uncertainty is obtained from the comparison of the various solutions (Fig. 2). An other option consists of checking the contamination on simulations. Figures 4 and 5 give an idea of the expected contamination from such simulations. This, however, is good only as long as the simulations are representative enough.

The comparison of the power spectrum of the output CMB map computed at high and low galactic latitude (Fig. 9), and a visualisation of the output CMP map at high and low galactic latitude (Fig. 10) also give an idea of the amount of galactic residuals, but none of these estimates is fully satisfactory for careful CMB science. This is, however, not particular to our map. No published CMB map is available with a good estimate of foreground contamination. Although this, clearly, is not fully satisfactory, we leave further investigations on this question for future work.

7.3. Final comments

7.3.1. Is our map optimal?

In the present work, we have obtained a CMB map which has been shown to be significantly less contaminated by foreground and noise than the other existing maps obtained from WMAP data. A natural question is whether we can do even better.

In the following, we outline where there is margin for improvement, and explain why we have stopped the analysis at its present state.

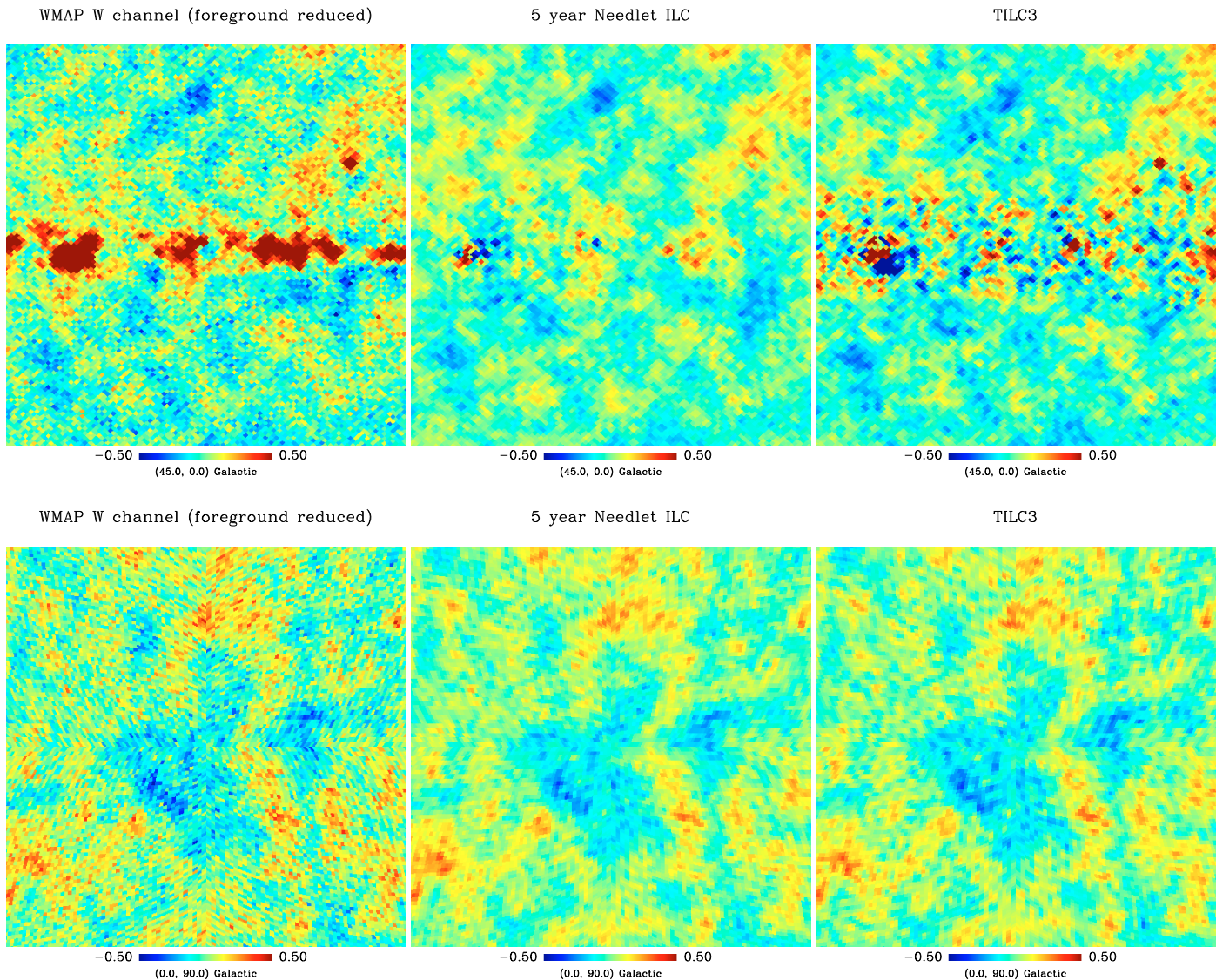


Fig. 10. CMB maps from WMAP three year data obtained with three techniques: *left column*: WMAP foreground reduced, W channel; *middle column*: our needlet ILC map; *right column*: the TILC map. The top line corresponds to a patch located in the galactic plane, centred around coordinates $(l, b) = (45^\circ, 0^\circ)$. The bottom line shows the recovered CMB around the North Galactic Pole.

First, the present analysis uses only limited external information and data sets: WMAP point source detections, and the IRIS 100 micron map. It is likely that something could be gained by using additional observations to help constrain the galactic emission.

Second, there is a trade-off between localisation of the estimation of covariance matrices, and bias in the ILC. The estimation of covariance matrices \mathbf{R}_x over sets of $N_p = 32 \times 32$ needlet coefficients results from a compromise which has been made based on varying N_p in simulations, but has not been optimised in any way. In addition, the optimal solution is probably different at high galactic latitudes, where weights given to different channels probably do not vary much and thus require less localisation, and at low galactic latitudes, where the complexity of galactic emission calls for more localisation. We have tried to use $N_p = 16 \times 16$ (more localised ILC filters, but more bias) and $N_p = 64 \times 64$ (less localised ILC filters, but less bias). Our choice of $N_p = 32 \times 32$ seems, in simulations, not worse than anything else (nor much better either). The bias error has been

verified to remain below the reconstruction error due to the contribution of the noise for small scales, and below cosmic variance uncertainty for large scales (see Fig. 13), and remains below an acceptable level of a few percent.

Similarly, the choice of the spectral window functions used on this data set has not been the object of specific optimisation. At low ℓ , we follow a “dyadic” scheme, where each window reaches an ℓ_{\max} of about twice the previous one. Wide spectral window functions allow for more localisation in pixel space, but narrow window functions allow for more accuracy in the harmonic domain. At high ℓ , because of the variation of the beams with ℓ , the relative noise levels of the different channels change quite fast with ℓ , which calls for more localisation in ℓ space. Here again, the optimal window functions are probably not the same at high and at low galactic latitudes. In practice, we chose a small number of bands to limit the number of harmonic transforms in the pipeline and allow reasonable localisation of the analysis.

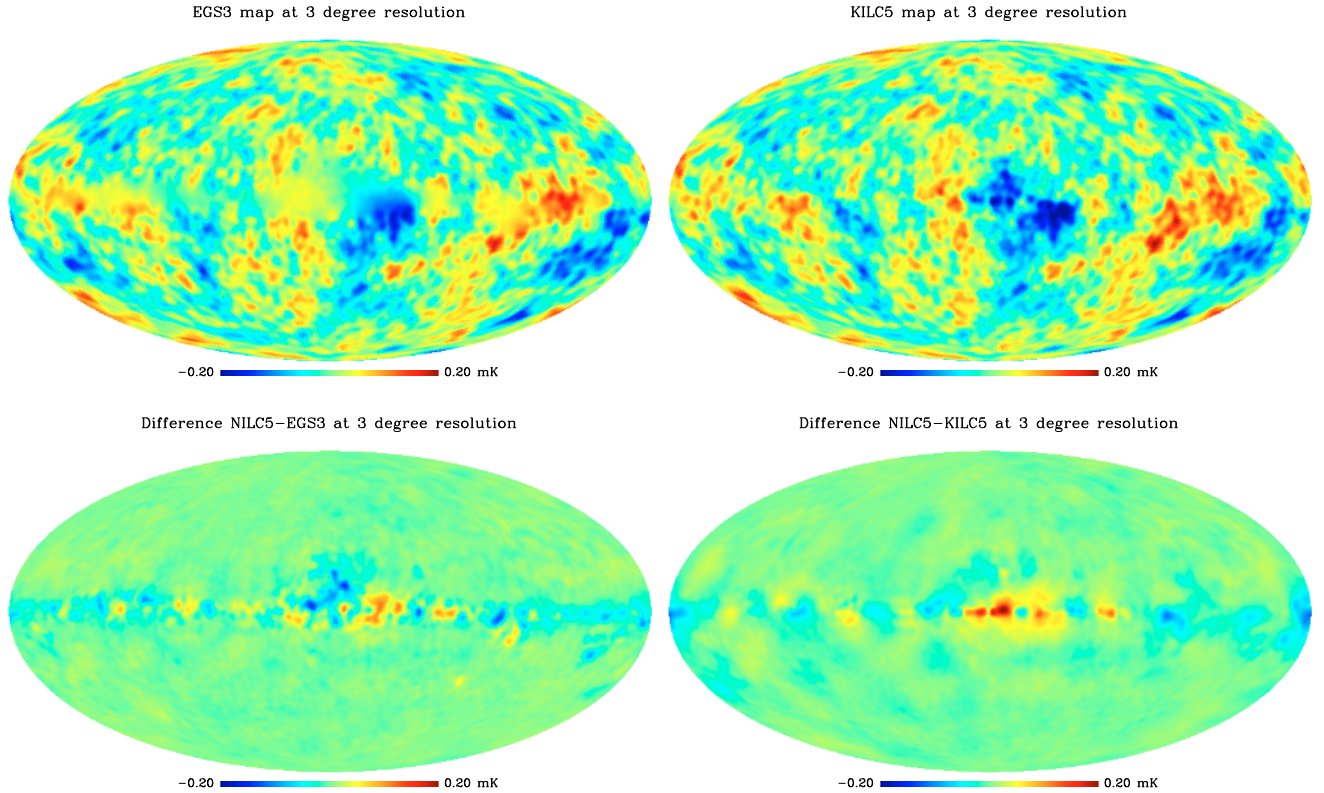


Fig. 11. *Top left:* the EGS3 map. *Top right:* The KILC5 map, smoothed at 3 degree resolution. *Bottom left:* the difference NILC5-EGS3; we see a clear difference in the galactic plane with no particular structure, compatible with a smooth Gaussian field, where the EGS3 solution poorly estimates intermediate scales; a patch in the difference map at the location of the Large Magellanic Cloud is clearly visible also. *Bottom right:* the difference NILC5-KILC5; a clear structure aligned with the galactic plane is clearly visible, with a large difference towards the galactic center.

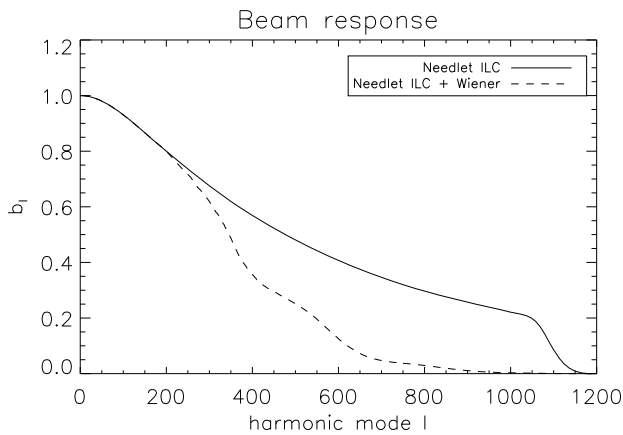


Fig. 12. Harmonic response of the beam of the NILC5 CMB map and of the Wiener-filtered version.

The choice we have made results from the principle of simplicity. We have tried to devise a pipeline which depends as little as possible on external data, on priors, or on fine tuning. A very simple scheme has permitted us to obtain a CMB map convincingly better than other maps available. This does not preclude any attempt at more optimisation for future work if needed.

7.3.2. Why ILC and not ICA?

It is certainly possible to tune our pipeline, changing some of its parameters. It would be possible also to use methods other than

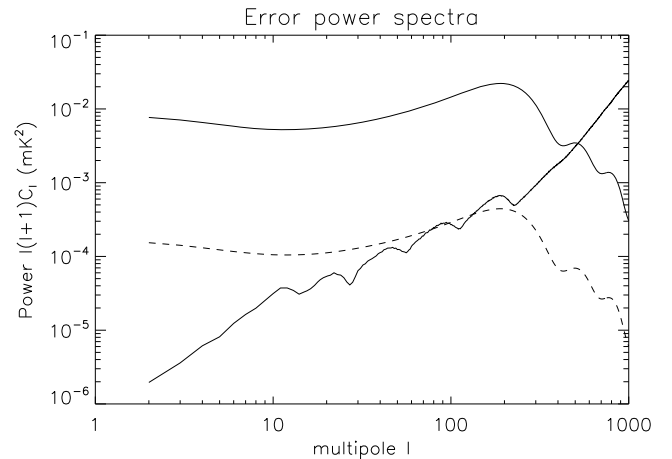


Fig. 13. Plot of the power spectrum of the noise (solid line). The spectrum for the WMAP best fit model is shown as a solid line, for comparison. The dashed line is 2% of the WMAP best fit C_ℓ , indicative of the level of the expected ILC bias. The bias is seen to dominate on large scales. There is, however, little margin for improvement, as few independent modes (or needlet coefficients) are available on the largest scales.

an ILC, for instance Independent Component Analysis (ICA) methods such as SMICA (Delabrouille et al. 2003; Cardoso et al. 2008), or more generally maximum likelihood methods fitting parametric models of the foreground emissions in the data sets. Such methods extract information about the foreground from the

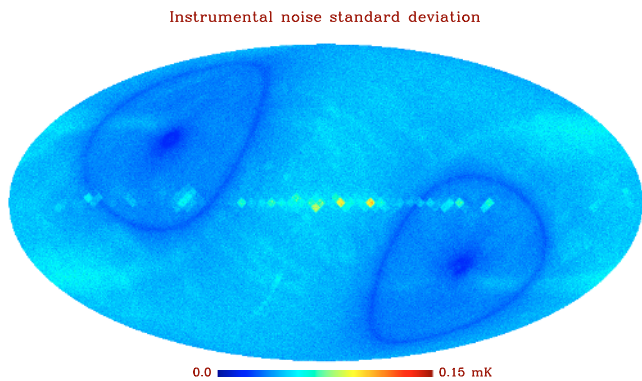


Fig. 14. Map of the standard deviation of the noise, per pixel at $n_{\text{side}} = 512$, for the five year needlet ILC map at the resolution of the W channel. The adaptation of the filter to local contamination is obvious from the uniform noise level in large healpix pixels, in particular in the galactic plane. These large healpix pixels correspond to $n_{\text{side}} = 16$, which is the size used to compute the average filter for the smallest scales. This illustrates the impact of the compromise between subtracting foreground emissions and increasing the noise with the use of the lower frequency channels, which have poor sensitivity on small scales.

data directly, possibly with the help of ancillary data sets, and use this information to clean the observations in some way.

In the present case however, it is not very likely that such attempts would give much better results than what is obtained here, unless one uses a very significant number of additional data sets and safe prior information. Indeed, the WMAP data consists of five channels only, from which a component separation method based on a meaningful model of foreground emission needs to extract CMB, synchrotron, free-free, spinning dust, thermal dust (i.e. five templates), and possibly also point sources, and variations of the spectral indices of some of the components – not to mention special regions of galactic emission such as cold cores and H-II regions, nor particular objects such as nearby resolved galaxies or galaxy clusters. Any component separation method based on the estimation of parameters for such a rich model would almost certainly be confronted with indeterminacy issues. Methods based on a precise model are expected to be effective when the data are very redundant as compared with the number of “parameters” of the emission model, which would not be the case here. Hence, the ILC is probably one of the best approaches for performing component separation on WMAP data. It is not surprising, then, that all methods producing full sky CMB maps from WMAP, or nearly so, implement some variant of the ILC.

Incidentally, ICA methods could benefit from the needlet framework.

8. Conclusions

In this paper, we have described a new approach to implement CMB extraction in WMAP data, using the ILC method on a needlet frame. Tests on simulations show excellent performance of the method, thanks to localisation both in pixel space and in harmonic space. Localisation in pixel space allows the ILC weights to adapt themselves to local conditions of foreground contamination and instrumental noise (this is essentially the reason why the NILC performs better than the TILC, in particular in the vicinity of the galactic plane). Localisation in harmonic space allows us to favour foreground rejection on large

scales (where foreground emissions dominate the total error) and instrumental noise rejection on small scales (where foreground emissions are negligible but where, after beam deconvolution, the relative noise level between the various WMAP channels varies a lot as a function of scale). Needlets permit us to vary the weights smoothly on large scales, and rapidly on small scales, which is not possible by dividing the sky in zones prior to any processing.

As a further improvement on previous work on WMAP data, we include a dust template in the set of analysed observations. This is motivated by the fact that on the smallest scales, observed with reasonable signal to noise ratio by the W channel only, dust emission contributes a significant fraction of the total reconstruction of the map. Using the IRIS 100 micron map as an additional observation enables the ILC to reduce the final contamination by dust – thanks to correlations of dust emission between the W channel and the 100 micron map. Special care was also taken to subtract a number of strong point sources from the data prior to the ILC.

As discussed at length in the main text and in the Appendix, the implementation of a filter (the ILC) based on empirical estimates of covariance matrices leads to a bias. This is not particular to our map, but is the case for any ILC map. We have estimated the level of this systematic effect, both analytically and numerically, to be at the level of a few per cent on all scales. Our simulation tool allows us to make accurate estimates of the amplitude of the effect, if needed for any scientific exploitation of the NILC5 map.

The application of the method to WMAP one year and three year data (in addition to five year data) allows us to compare the needlet ILC solution to previous work. Our map is seen to be at least as good as others on large scales, while being significantly less contaminated by residual foreground and noise than others on small scales, in particular in the vicinity of the galactic plane. The application of the method to WMAP five year data yields a CMB map which we believe to be the cleanest full sky map of the CMB to date. Contamination by noise, and the power loss due to the use of the ILC method, are characterised by means of Monte Carlo simulations. The map is available for download on the ADAMIS web site⁷, and can be used for a variety of science projects relying on accurate maps of the CMB.

Acknowledgements. The ADAMIS team at APC has been partly supported by the Astro-Map and Cosmostat ACI grants of the French ministry of research, for the development of innovative CMB data analysis methods. We acknowledge the use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science. The results in this paper have been derived using the HEALPix package (Górski et al. 2005). The authors acknowledge the use of the Planck Sky Model⁸, developed by the Planck working group on component separation, for making the simulations used in this work. Our pipeline is mostly implemented in octave (<http://www.octave.org>).

Appendix A: Derivation of the ILC bias

In this appendix, we compute the error made after CMB reconstruction with the ILC and some of its statistical properties. In particular, we derive the correlation of the error with the true CMB signal, which yields a non unit effective “response” of the ILC filter – and hence a *bias* in the reconstructed map and in the CMB power spectrum computed from it. This bias has to

⁷ http://www.apc.univ-paris7.fr/APC_CS/Recherche/Adamis/cmb_wmap-en.php

⁸ http://www.apc.univ-paris7.fr/APC_CS/Recherche/Adamis/PSM/psky-en.php

be accounted for properly for further use of the reconstructed CMB map.

We model the data as:

$$\mathbf{x}_p = \mathbf{a}s_p + \mathbf{n}_p \quad (\text{A.1})$$

where s_p are the coefficients of the map of interest over some domain \mathcal{D} (e.g. needlet coefficients of the CMB map for a given scale and a given patch of the sky, or pixel values in a certain region of the sky, or values of the harmonic coefficients in some band of ℓ). p indexes the coefficient (i.e. pixel, or harmonic mode, or needlet coefficient). \mathbf{x}_p are the observations of coefficient p for the set of available observed maps, and \mathbf{n}_p the corresponding “noise” terms (including foreground contaminants).

The ILC is best applied over domains of p where all coefficients have (near) uniform expected signal and noise properties, so that the ILC weights are optimal simultaneously for all p . In particular, the rms values of all maps do not depend (much) on p in a given domain. Hence, we concentrate on one given domain of p for which we assume that the sequences s_p and \mathbf{n}_p are independent, Gaussian random variables with distribution $\mathcal{N}(0, \sigma_s^2)$ and $\mathcal{N}(0, \mathbf{R}_n)$, with σ_s^2 the variance of s_p (the CMB) and \mathbf{R}_n the covariance matrix of the noise (including foreground emissions).

The ILC estimate of s_p in domain \mathcal{D} is given, for all p , by

$$\hat{s}_p = \frac{\mathbf{a}^t \widehat{\mathbf{R}}_x^{-1} \mathbf{x}_p}{\mathbf{a}^t \widehat{\mathbf{R}}_x^{-1} \mathbf{a}} \quad (\text{A.2})$$

where $\widehat{\mathbf{R}}_x$ is an estimate of the covariance matrix of the observations, obtained as:

$$\begin{aligned} \widehat{\mathbf{R}}_x &= \frac{1}{N_p} \sum_p \mathbf{x}_p \mathbf{x}_p^t \\ &= \frac{1}{N_p} \sum_p (\mathbf{a}s_p + \mathbf{n}_p)(\mathbf{a}s_p + \mathbf{n}_p)^t \end{aligned} \quad (\text{A.3})$$

with N_p the number of coefficients in domain \mathcal{D} . In the limit of large N_p , $\widehat{\mathbf{R}}_x$ approaches its expectation (ensemble average) value $E(\widehat{\mathbf{R}}_x) = \mathbf{R}_x$. For finite N_p , we have instead

$$\widehat{\mathbf{R}}_x = \mathbf{R}_x + \Delta_x \quad (\text{A.4})$$

where Δ_x is a correction term corresponding to the departure of the empirical correlation from its ensemble average due to the finite sample size N_p . From now on, we assume that N_p is large enough that this correction is small: we investigate effects at first order in $1/N_p$.

A.1. First order expansion of the reconstruction error

We are interested in the reconstruction error:

$$d_p = \hat{s}_p - s_p. \quad (\text{A.5})$$

The first and second moments of d_p , i.e. the mean value $E(d_p)$ of the reconstruction error, as well as its variance $E(d_p^2)$, are of particular interest for the interpretation of the reconstructed map. In particular, we have

$$E(\hat{s}_p^2) = E(s_p^2) + E(d_p^2) + 2E(s_p d_p). \quad (\text{A.6})$$

In our case, $E(\hat{s}_p^2)$ can be used to estimate $E(s_p^2)$. In the case where domain \mathcal{D} is a harmonic domain, p indexes harmonic modes (ℓ, m) , and $E(s_p^2)$ is a term of the CMB power spectrum.

In our needlet approach, $E(s_p^2)$ is also directly connected to the CMB power spectrum. For this reason, it is important to characterise in the best way we can the “noise bias” $E(d_p^2)$ and the covariance of the error with the CMB, $E(s_p d_p)$.

The ILC being constructed so that the response to the signal of interest is unity, only the filtered noise term contributes to the error d_p , which can then be written as:

$$d_p = \frac{\mathbf{a}^t \widehat{\mathbf{R}}_x^{-1} \mathbf{n}_p}{\mathbf{a}^t \widehat{\mathbf{R}}_x^{-1} \mathbf{a}} \quad (\text{A.7})$$

$$= \frac{\mathbf{a}^t [\mathbf{R}_x + \Delta_x]^{-1} \mathbf{n}_p}{\mathbf{a}^t [\mathbf{R}_x + \Delta_x]^{-1} \mathbf{a}} \quad (\text{A.8})$$

where Δ_x is a small perturbation to \mathbf{R}_x . We use the first order expansion:

$$[\mathbf{R} + \Delta_x]^{-1} \simeq \mathbf{R}^{-1} - \mathbf{R}^{-1} \Delta_x \mathbf{R}^{-1} \quad (\text{A.9})$$

which yields

$$d_p = \frac{\mathbf{a}^t [\mathbf{R}_x^{-1} - \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1}] \mathbf{n}_p}{\mathbf{a}^t [\mathbf{R}_x^{-1} - \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1}] \mathbf{a}}. \quad (\text{A.10})$$

Writing:

$$\begin{aligned} \frac{1}{\mathbf{a}^t [\mathbf{R}_x^{-1} - \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1}] \mathbf{a}} &= \frac{1}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} \frac{1}{(1 - \epsilon)} \\ &\simeq \frac{1}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} (1 + \epsilon) \end{aligned}$$

where ϵ is

$$\epsilon = \frac{\mathbf{a}^t \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1} \mathbf{a}}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} \quad (\text{A.11})$$

we get

$$d_p \simeq \frac{\mathbf{a}^t [\mathbf{R}_x^{-1} - \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1}] \mathbf{n}_p}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} (1 + \epsilon). \quad (\text{A.12})$$

Keeping only first order terms in (Δ_x) yields:

$$\begin{aligned} d_p &= \frac{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} - \frac{\mathbf{a}^t \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1} \mathbf{n}_p}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} \\ &\quad + \frac{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p] [\mathbf{a}^t \mathbf{R}_x^{-1} \Delta_x \mathbf{R}_x^{-1} \mathbf{a}]}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]^2}. \end{aligned} \quad (\text{A.13})$$

The first term on the right hand side, proportional to \mathbf{n} , is the “ideal” ILC error, i.e. the error we would get if we knew perfectly the “true” covariance matrix \mathbf{R}_x of the observations. The second and third terms, proportional to Δ_x , are corrections due to the fact that this covariance matrix is actually obtained empirically from the observations themselves.

From Eqs. (A.3) and (A.4), we can write Δ_x in the form:

$$\Delta_x = \delta_s \mathbf{a} \mathbf{a}^t + \Delta_n + \widehat{\mathbf{C}} \quad (\text{A.14})$$

where

$$\delta_s = \widehat{\sigma}_s^2 - \sigma_s^2 \quad (\text{A.15})$$

$$\Delta_n = \widehat{\mathbf{R}}_n - \mathbf{R}_n \quad (\text{A.16})$$

$$\widehat{\mathbf{C}} = \frac{1}{N_p} \sum_q (\mathbf{n}_q s_q \mathbf{a}^t + \mathbf{a} s_q \mathbf{n}_q^t). \quad (\text{A.17})$$

These three quantities correspond respectively, in pixel (or mode, or needlet coefficient) p , to the uncertainty in CMB variance estimates due to “cosmic” (or sample) variance, to the error in the estimation of the “noise” covariance matrix alone (if maps of noise + foreground alone were available), and to a cross term, originating from the empirical covariance between CMB and noise due the finite sample size N_p .

The two last terms (small correction terms) in Eq. (A.13), being proportional to Δ_x , can be decomposed each into the sum of three terms, proportional to $(\delta_s \mathbf{a} \mathbf{a}^t)$, Δ_n , and $\widehat{\mathbf{C}}$ respectively. The signal and noise realisations enter the $(\delta_s \mathbf{a} \mathbf{a}^t)$ term as products of terms of the form $(\mathbf{a} \mathbf{a}^t s_q s_q \mathbf{n}_p)$ only, the Δ_n term as products of terms of the form $(\mathbf{n}_q \mathbf{n}_q^t \mathbf{n}_p)$. On the contrary, signal and noise realisations enter the $\widehat{\mathbf{C}}$ term as the product of terms in the form $(\mathbf{a} s_q \mathbf{n}_q^t \mathbf{n}_p)$, i.e. products of s and the second power of \mathbf{n} . Index q runs over domain \mathcal{D} .

Assuming that s and \mathbf{n} are centred variables, the mean value of the error is immediately seen to vanish:

$$E(d_p) = 0. \quad (\text{A.18})$$

The main contribution to the variance comes from the first term on the right hand side of Eq. (A.13). The second and third terms are small corrections to this variance estimate, so that to first order, we get:

$$E(d_p^2) \simeq \frac{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{R}_n \mathbf{R}_x^{-1} \mathbf{a}}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]^2}. \quad (\text{A.19})$$

Recalling that $\mathbf{R}_x = \mathbf{R}_n + \sigma^2 \mathbf{a} \mathbf{a}^t$, where σ^2 is the variance of the CMB, and making use of the inversion formula:

$$\left[\mathbf{R}_n + \sigma^2 \mathbf{a} \mathbf{a}^t \right]^{-1} = \mathbf{R}_n^{-1} - \sigma^2 \frac{\mathbf{R}_n^{-1} \mathbf{a} \mathbf{a}^t \mathbf{R}_n^{-1}}{1 + \sigma^2 \mathbf{a}^t \mathbf{R}_n^{-1} \mathbf{a}} \quad (\text{A.20})$$

we finally obtain:

$$E(d_p^2) \simeq \frac{1}{[\mathbf{a}^t \mathbf{R}_n^{-1} \mathbf{a}]}. \quad (\text{A.21})$$

The most interesting terms are those connecting the error to the signal of interest, $E(s_p d_p)$, which is necessary to compute the power spectrum of the output map according to A.6⁹.

As mentioned previously, under the assumption that the signal of interest is not correlated to the noise and the foregrounds, the first term (main term) of the r.h.s. of Eq. (A.13) does not give rise to multiplicative errors (or correlation of d_p with s_p). Similarly, the corrective term proportional to $\delta_s \mathbf{a} \mathbf{a}^t + \Delta_n$, multiplied by s , gives terms which contain an odd power of s and an odd power of \mathbf{n} , and does not give rise to correlations. This assumption is correct, to excellent accuracy, when the signal of interest is CMB anisotropies¹⁰. We are left with:

$$E(s_p d_p) = E \left(\sum_q \frac{s_p s_q}{N_p} \frac{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p] [\mathbf{a}^t \mathbf{R}_x^{-1} (\mathbf{n}_q \mathbf{a}^t + \mathbf{a} \mathbf{n}_q^t) \mathbf{R}_x^{-1} \mathbf{a}]}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]^2} \right) - E \left(\sum_q \frac{s_p s_q}{N_p} \frac{\mathbf{a}^t \mathbf{R}_x^{-1} (\mathbf{n}_q \mathbf{a}^t + \mathbf{a} \mathbf{n}_q^t) \mathbf{R}_x^{-1} \mathbf{n}_p}{\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}} \right). \quad (\text{A.22})$$

⁹ We warn the reader that some authors fail to make a clear distinction between the statistical (ensemble average) correlation, which is a deterministic quantity, and the “empirical correlations” computed, assuming some kind of ergodicity, as averages over finite sets of samples as in Eq. (A.3).

¹⁰ Certainly the CMB is not correlated to galactic components. Small correlations with large scale structure, and hence with SZ effect and emission from outer galaxies, may exist because of the integrated Sachs-Wolfe effect. We neglect this effect in the present discussion.

Multiplying the numerator and denominator of the second term by $\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}$ and expanding numerators, two terms cancel and two remain. If we assume in addition that signal and/or noise coefficients are independent, i.e. $E(s_p s_q) = \sigma_s^2 \delta_{pq}$ and/or $E(\mathbf{n}_p \mathbf{n}_q^t) = \mathbf{R}_n \delta_{pq}$, only the pp term is non vanishing, and we get

$$E(s_p d_p) = \frac{\sigma_s^2}{N_p} \left(\frac{E \left((\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)^2 \right) - (\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}) E \left(\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_p \right)}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]} \right). \quad (\text{A.23})$$

We compute

$$\begin{aligned} E \left((\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)^2 \right) &= \mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{R}_n \mathbf{R}_x^{-1} \mathbf{a} \\ &= \mathbf{a}^t \mathbf{R}_x^{-1} \left[\mathbf{R}_x - \sigma_s^2 \mathbf{a} \mathbf{a}^t \right] \mathbf{R}_x^{-1} \mathbf{a} \\ &= \left[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a} \right] \left[1 - \sigma_s^2 \mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a} \right] \end{aligned} \quad (\text{A.24})$$

and

$$\begin{aligned} E \left(\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_p \right) &= \text{Tr} \left(\mathbf{R}_x^{-1} \mathbf{R}_n \right) \\ &= \text{Tr} \left(\mathbf{R}_x^{-1} \left[\mathbf{R}_x - \sigma_s^2 \mathbf{a} \mathbf{a}^t \right] \right) \\ &= \text{Tr}(\text{Id}) - \sigma_s^2 \text{Tr} \left(\mathbf{R}_x^{-1} \mathbf{a} \mathbf{a}^t \right) \\ &= m - \sigma_s^2 \left(\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a} \right) \end{aligned} \quad (\text{A.25})$$

where m is the number of channels used for the ILC (here, 5 WMAP channels + 1 IRIS map, for a total of 6). Substituting the results of Eqs. (A.24) and (A.25) into Eq. (A.23), we get the simple final result:

$$E(s_p d_p) = \frac{\sigma_s^2 (1 - m)}{N_p}. \quad (\text{A.26})$$

The error in the reconstructed CMB map comprises a term proportional (on average) to the CMB. In our application, $m = 6$ and $N_p = 1024$, so that if indeed all needlet coefficients were independent, the amplitude of the effect should be $E(s_p d_p) \simeq 5 \times 10^{-3} \sigma_s^2$, i.e. a bias of about half a percent in the CMB reconstruction.

A.2. A geometric interpretation

Although allowing the statistical derivation of the (anti-) correlation of the reconstruction error with the original CMB, the above derivation is not very illuminating about the mechanism giving rise to this CMB power loss. A geometrical reasoning gives better insight into what is actually going on.

For a given data set, the ILC works on one single realisation of all random fields. For an independent implementation of the ILC on N_p pixels (or modes, or needlet coefficients) of the observations, each data set is represented by a vector in an N_p -dimensional vector space W . The CMB s , the observation \mathbf{x}_i for each channel i , and each of the noise realisations \mathbf{n}_i (including foreground emissions) are elements of W .

The collection of vectors \mathbf{n}_i defines an m -dimensional subspace V of W . This is true irrespective of the nature of the foreground contaminants: indeed, although in principle vectors \mathbf{n}_i could be linearly dependent, this happens in practice with vanishing probability (in particular if the observations are noisy).

Vector space W can thus be decomposed in two orthogonal subspaces, U and V , where V is the m -dimensional sub-space spanned by all vectors \mathbf{n}_i , and $U = V^\perp$ is a $(N_p - m)$ dimensional

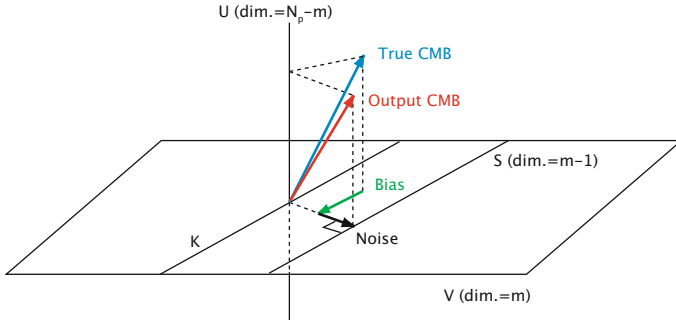


Fig. A.1. Geometric illustration of the ILC bias. The true CMB (blue) and the output CMB (red) are elements of $W = U \oplus V$. The difference between the two can be decomposed into the sum of two elements of V : a bias (green), which is an element of K , and a noise contribution (black), which is an element of H , the orthogonal of K in V .

vector space. The CMB itself can be decomposed into two components, one in U , and one in V :

$$\mathbf{s} = \mathbf{s}_U + \mathbf{s}_V \quad (\text{A.27})$$

where \mathbf{s}_U is the orthogonal projection of \mathbf{s} onto U , and \mathbf{s}_V its orthogonal projection onto V .

What happens when the ILC is made is the following: we look for weights w_i for all channels, that minimise the variance of the reconstructed map, i.e. minimise the norm of vector $\widehat{\mathbf{s}} = \sum_i w_i \mathbf{x}_i$, under the constraint that $\sum_i w_i = 1$. We have:

$$\begin{aligned} \widehat{\mathbf{s}} &= \mathbf{s}_U + \left(\mathbf{s}_V + \sum_i w_i \mathbf{n}_i \right) \\ &= \widehat{\mathbf{s}}_U + \widehat{\mathbf{s}}_V \end{aligned} \quad (\text{A.28})$$

where the first term is a vector of U and the second term a vector of V , and the second line of the equation defines $\widehat{\mathbf{s}}_U$ and $\widehat{\mathbf{s}}_V$. Since these two subspaces of W are orthogonal, the norm of $\widehat{\mathbf{s}}$ is the sum of the norms of the two vectors $\widehat{\mathbf{s}}_U$ and $\widehat{\mathbf{s}}_V$. The norm of $\widehat{\mathbf{s}}$ thus depends on w_i only through the norm of the projection of $\widehat{\mathbf{s}}$ on subspace V .

The noise contribution to $\widehat{\mathbf{s}}$ appears as a linear combination of vectors \mathbf{n}_i . For varying values of w_i such that $\sum_i w_i = 1$, this linear combination spans an *affine* subspace S of V . S is of dimension $m - 1$ (an hyperplane). Defining K as the vector subspace of V spanned by linear combinations $\sum_i w_i \mathbf{x}_i$ such that $\sum_i w_i = 0$, we obtain S as:

$$S = \mathbf{p} + K \quad (\text{A.29})$$

where \mathbf{p} is any element of S .

We note that the vector subspace K depends only on noise realisations, and not on \mathbf{s} nor on the final ILC weights (the latter only defining a single point on S – and on K by orthogonal projection). Hence, the direction of the one-dimensional vector subspace H of V orthogonal to K is also independent of \mathbf{s} and of the final ILC weights.

For any element $\widehat{\mathbf{s}}_V = \mathbf{s}_V + \sum_i w_i \mathbf{n}_i$ of affine space S the norm of $\widehat{\mathbf{s}}_V$ is the sum of the norms of its projections onto K and H . Allowing weights w_i to vary, vector $\widehat{\mathbf{s}}_V$ spans S , and hence only the norm of the projection of $\widehat{\mathbf{s}}_V$ onto K varies (and not its projection on H). The minimum is reached when the projection of $\widehat{\mathbf{s}}_V$ onto K vanishes. When this happens, the ILC has cancelled completely the linear combination of projections of \mathbf{s} and \mathbf{n}_i onto the $m - 1$ -dimensional space K , and left untouched the projections of \mathbf{s} and \mathbf{n}_i onto the $N_p - m + 1$ dimensional vector space $U \oplus H$.

Assuming the CMB to be Gaussian and statistically isotropic, its coefficients in any orthogonal basis are Gaussian distributed random variables with variance σ^2/N_p (since the sum must have total variance σ^2). It follows straightforwardly that the correlation of the recovered CMB map with the input CMB map is $(N_p - m + 1)/N_p$, and that the “bias” is due to the loss of $m - 1$ modes of the original CMB, which have been unlucky enough to “live” in the $(m - 1)$ dimensional space K .

A.3. Comment on coefficient independence

The above derivation in Sect. A.1 assumes the independence of coefficients \mathbf{n}_p and/or of coefficients s_p , i.e. $E(\mathbf{n}_p \mathbf{n}_q^t) = \mathbf{R}_n \delta_{pq}$ and/or $E(s_p s_q) = \sigma_s^2 \delta_{pq}$. When this assumption does not hold, we have:

$$\begin{aligned} E(s_p d_p) &= \frac{1}{N_p} \sum_q E \left(s_p s_q \frac{(\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)(\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_q)}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]} \right) \\ &\quad - \frac{1}{N_p} \sum_q E(s_p s_q (\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_q)). \end{aligned} \quad (\text{A.30})$$

Assuming that the noise and the CMB are independent, we have:

$$\begin{aligned} E(s_p d_p) &= \frac{1}{N_p} \sum_q E(s_p s_q) \frac{E((\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)(\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_q))}{[\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{a}]} \\ &\quad - \frac{1}{N_p} \sum_q E(s_p s_q) E(\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_q) \end{aligned} \quad (\text{A.31})$$

where

$$E((\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_p)(\mathbf{a}^t \mathbf{R}_x^{-1} \mathbf{n}_q)) = \mathbf{a}^t \mathbf{R}_x^{-1} E(\mathbf{n}_p \mathbf{n}_q^t) \mathbf{R}_x^{-1} \mathbf{a} \quad (\text{A.32})$$

and

$$E(\mathbf{n}_p^t \mathbf{R}_x^{-1} \mathbf{n}_q) = \text{Tr}(\mathbf{R}_x^{-1} E(\mathbf{n}_p \mathbf{n}_q^t)). \quad (\text{A.33})$$

When p and q index needlet coefficients as in the present work, we have:

$$E(s_p s_q) = \sum_{\ell} \frac{2\ell + 1}{N_{\text{tot}}} h_{\ell}^2 C_{\ell} P_{\ell}(\cos \theta_{qp}) \quad (\text{A.34})$$

where s_p and s_q are needlet coefficients of the CMB map, evaluated at two different points p and q , C_{ℓ} is the angular power spectrum of the CMB, θ_{qp} is the angle between q and p , and N_{tot} is the total number of pixels of the needlet coefficient map. For noise maps (including foreground emissions), which are not stationary Gaussian random fields on the sphere, the analogous formula is just an approximation, which can be written as:

$$E(\mathbf{n}_p \mathbf{n}_q^t) \simeq \sum_{\ell} \frac{2\ell + 1}{N_{\text{tot}}} h_{\ell}^2 \mathbf{R}_n(\ell) P_{\ell}(\cos \theta_{qp}) \quad (\text{A.35})$$

where \mathbf{n}_p and \mathbf{n}_q are needlet coefficients of all noise maps, evaluated at two different points p and q , $\mathbf{R}_n(\ell)$ is the covariance of the noise needlet coefficients (an $m \times m$ matrix for each ℓ), and θ_{qp} and N_{tot} are defined as above.

Assuming that neither the noise level nor the CMB power vary much over the spectral window h_{ℓ} , $\mathbf{R}_n(\ell)$ and C_{ℓ} are approximately independent of ℓ , and can be taken out of the integral. We get:

$$E(s_p s_q) = C_{\ell} k(\theta_{qp}) \quad (\text{A.36})$$

and

$$E(\mathbf{n}_p \mathbf{n}_q^t) = \mathbf{R}_n k(\theta_{qp}) \quad (\text{A.37})$$

with

$$k(\theta_{qp}) = \sum_{\ell} \frac{2\ell + 1}{N_{\text{tot}}} h_{\ell}^2 P_{\ell}(\cos \theta_{qp}). \quad (\text{A.38})$$

Hence, plugging this result, together with Eqs. (A.32) and (A.33), into Eq. (A.31) we get:

$$E(s_p d_p) = \frac{(1-m)}{N_p} C_{\ell} \sum_q k^2(\theta_{qp}). \quad (\text{A.39})$$

Finally, noting that $\sigma_s^2 = C_{\ell} k(0)$, we get

$$E(s_p d_p) = \frac{(1-m)}{N_p} \frac{\sigma_s^2}{k(0)} \sum_q k^2(\theta_{qp}). \quad (\text{A.40})$$

Equation (A.40) is the same as Eq. (A.26), except for a coefficient, which measures the correlation between signal and noise coefficients p and coefficients q in domain \mathcal{D} . In particular, the result is, again, independent of \mathbf{R}_n .

Hence, we define an *effective number of modes*, $N_p^{\text{eff}} = N_p/\alpha$, where

$$\alpha = \frac{\sum_q k^2(\theta_{qp})}{k(0)} \quad (\text{A.41})$$

and we get

$$E(d_p s_p) = \frac{\sigma_p^2(1-m)}{N_p^{\text{eff}}}. \quad (\text{A.42})$$

An approximation (and upper bound) for α is easily obtained in the special case where h_{ℓ} is a square spectral window, and when domains \mathcal{D} over which the ILC is implemented are small regions of the sky, so that $h_{\ell}^2 = 1$, and $P_{\ell}(\cos \theta_{qp}) \simeq 1$. In this case, we have:

$$k(\theta_{qp}) \simeq \frac{1}{N_{\text{tot}}} \left((\ell_{\text{max}} + 1)^2 - \ell_{\text{min}}^2 \right) \quad (\text{A.43})$$

and we have

$$N_p^{\text{eff}} \simeq \frac{N_p}{N_{\text{tot}}} \left((\ell_{\text{max}} + 1)^2 - \ell_{\text{min}}^2 \right). \quad (\text{A.44})$$

We note that $\left((\ell_{\text{max}} + 1)^2 - \ell_{\text{min}}^2 \right)$ simply is the number of harmonic coefficients selected by the spectral window h_{ℓ} , and that $N_p/N_{\text{tot}} = f_{\text{sky}}$ is a coefficient which takes into account the effect of partial sky coverage for the (local) calculation of the statistics of the data set, well in line with N_p^{eff} being understood as an effective number of modes.

References

Abrial, P., Moudén, Y., Starck, J.-L., et al. 2008, *Statistical Methodology*, 5, 289

- Aurich, R., Lustig, S., & Steiner, F. 2006, *MNRAS*, 369, 240
- Baldi, P., Kerkyacharian, G., Marinucci, D., & Picard, D. 2007, *ArXiv e-prints*, 706
- Baldi, P., Kerkyacharian, G., Marinucci, D., & Picard, D. 2008, *Annals of Statistics*, in press [arXiv.org: math.ST/0606599]
- Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2003, *ApJS*, 148, 97
- Bonaldi, A., Bedini, L., Salerno, E., Baccigalupi, C., & de Zotti, G. 2006, *MNRAS*, 373, 271
- Bonaldi, A., Ricciardi, S., Leach, S., et al. 2007, *ArXiv e-prints*, 707
- Cabella, P., Hansen, F. K., Liguori, M., et al. 2006, *MNRAS*, 369, 819
- Cabré, A., Gaztañaga, E., Manera, M., Fosalba, P., & Castander, F. 2006, *MNRAS*, 372, L23
- Caillerie, S., Lachièze-Rey, M., Luminet, J.-P., et al. 2007, *A&A*, 476, 691
- Cao, L., Chu, Y.-Q., & Fang, L.-Z. 2006, *MNRAS*, 369, 645
- Cardoso, J.-F., Martin, M., Delabrouille, J., Betoule, M., & Patanchon, G. 2008, *ArXiv e-prints*, 803
- Cayón, L., Jin, J., & Treaster, A. 2005, *MNRAS*, 362, 826
- Daubechies, I. 1992, *CBMS-NSF Regional Conference Series in Applied Mathematics*, Ten lectures on wavelets (Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM)), Vol. 61
- Delabrouille, J., & Cardoso, J. F. 2007, *ArXiv Astrophysics e-prints*
- Delabrouille, J., Cardoso, J.-F., & Patanchon, G. 2003, *MNRAS*, 346, 1089
- Elad, M., Starck, J. L., Donoho, D., & Querre, P. 2005, *Applied and Computational Harmonic Analysis*, 19, 340
- Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2004, *ApJ*, 612, 633
- Eriksen, H. K., Dickinson, C., Jewell, J. B., et al. 2007, *ArXiv e-prints*, 709
- Finkbeiner, D. P. 2003, *ApJS*, 146, 407
- Finkbeiner, D. P., Davis, M., & Schlegel, D. J. 1999, *ApJ*, 524, 867
- Fosalba, P., & Gaztañaga, E. 2004, *MNRAS*, 350, L37
- Gold, B., Bennett, C. L., Hill, R. S., et al. 2008, *ArXiv e-prints*, 803
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759
- Guilloux, F., Fay, G., & Cardoso, J.-F. 2008, *Applied and Computational Harmonic Analysis*, in press [arXiv:0706.2598]
- Hinshaw, G., Nolta, M. R., Bennett, C. L., et al. 2007, *ApJS*, 170, 288
- Huffenberger, K. M., Eriksen, H. K., Hansen, F. K., Banday, A. J., & Gorski, K. M. 2007, *ArXiv e-prints*
- Kim, J., Naselsky, P., & Christensen, P. R. 2008, *ArXiv e-prints*, 803
- Maino, D., Farusi, A., Baccigalupi, C., et al. 2002, *MNRAS*, 334, 53
- Maino, D., Donzelli, S., Banday, A. J., Stivoli, F., & Baccigalupi, C. 2007, *MNRAS*, 374, 1207
- Marinucci, D., Pietrobon, D., Balbi, A., et al. 2008, *MNRAS*, 383, 539
- McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2006a, *MNRAS*, 371, L50
- McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2006b, *MNRAS*, 369, 1858
- McEwen, J. D., Wiaux, Y., Hobson, M. P., Vanderghenst, P., & Lasenby, A. N. 2007, *ArXiv e-prints*, 704
- Miville-Deschênes, M.-A., & Lagache, G. 2005, *ApJS*, 157, 302
- Mukherjee, P., & Wang, Y. 2004, *ApJ*, 613, 51
- Narcowich, F., Petrushev, P., & Ward, J. 2006, *SIAM J. Math. Anal.*, 38, 574
- Niarchou, A., & Jaffe, A. 2007, *Phys. Rev. Lett.*, 99, 081302
- Nolta, M. R., Wright, E. L., Page, L., et al. 2004, *ApJ*, 608, 10
- Page, L., Barnes, C., Hinshaw, G., et al. 2003, *ApJS*, 148, 39
- Park, C.-G., Park, C., & Gott, J. R. I. 2007, *ApJ*, 660, 959
- Pietrobon, D., Balbi, A., & Marinucci, D. 2006, *Phys. Rev. D*, 74, 043524
- Rassat, A., Land, K., Lahav, O., & Abdalla, F. B. 2007, *MNRAS*, 377, 1085
- Saha, R., Prunet, S., Jain, P., & Souradeep, T. 2007, *ArXiv e-prints*, 706
- Starck, J.-L., Moudén, Y., Abrial, P., & Nguyen, M. 2006, *A&A*, 446, 1191
- Tegmark, M., de Oliveira-Costa, A., & Hamilton, A. J. 2003, *Phys. Rev. D*, 68, 123523
- Tojeiro, R., Castro, P. G., Heavens, A. F., & Gupta, S. 2006, *MNRAS*, 365, 265
- Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., & Cayón, L. 2004, *ApJ*, 609, 22
- Wiaux, Y., Vielva, P., Barreiro, R. B., Martínez-González, E., & Vanderghenst, P. 2007, *ArXiv e-prints*, 706
- Wiaux, Y., McEwen, J. D., Vanderghenst, P., & Blanc, O. 2008, *MNRAS*, 388, 770
- Wright, E. L., Chen, X., Odegard, N., et al. 2008, *ArXiv e-prints*
- Yadav, A. P. S., & Wandelt, B. D. 2007, *ArXiv e-prints*, 712