Mon. Not. R. Astron. Soc. **344,** 544–552 (2003)

Astrophysical component separation of *COBE*-DMR 4-yr data with FastICA

D. Maino,[1⋆] A. J. Banday,[2] C. Baccigalupi,[3] F. Perrotta[3] and K. M. Górski[4,5†]

[1]*Dipartimento di Fisica, Università di Milano, Via Celoria 16, 20133, Milano, Italy*
[2]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschildstrasse 1, 85741, Garching bei München, Germany*
[3]*SISSA/ISAS, Astrophysics Sector, Via Beirut 4, 34014, Trieste, Italy*
[4]*ESO, Karl-Schwarzschildstrasse 2, 85740, Garching bei München, Germany*
[5]*Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland*

Accepted 2003 May 22. Received 2003 May 22; in original form 2003 March 28

ABSTRACT

We present an application of the Fast Independent Component Analysis (FastICA) method to the *Cosmic Background Explorer* Differential Microwave Radiometer (*COBE*-DMR) 4-yr data. Although the signal-to-noise (S/N) ratio in the *COBE*-DMR data is typically ∼1, the approach is able to extract the cosmic microwave background (CMB) signal with high confidence from high galactic latitude regions. However, the foreground emission components have too low a S/N ratio to be reconstructed by this method (moreover, the number of components which can be reconstructed is directly limited by the number of input channels).

The reconstructed CMB map shows the expected frequency scaling of the CMB and, fitting for the rms quadrupole normalization $Q_{\rm rms-PS}$ and primordial spectral index n we find results in excellent agreement with those derived from the minimum-noise combination of the 90- and 53-GHz DMR channels without galactic emission correction.

We extend the analysis by including additional channels (priors): the Haslam map of radio emission at 408 MHz and the Diffuse Infrared Background Experiment (DIRBE) 140-μm map of galactic infrared emission. The FastICA algorithm is now able to both detect galactic foreground emission and separate it from the dominant CMB signal. Fitting again for $Q_{\rm rms-PS}$ and n we find good agreement with the results of Górski et al. where galactic emission has been taken into account by means of correlation analysis of the DMR signal.

We investigate the ability of FastICA to evaluate the extent of foreground contamination in the *COBE*-DMR data further including an all-sky Hα survey to determine a reliable free–free. The derived frequency scalings of the recovered foregrounds are consistent with previous correlation studies. After subtraction of the thermal dust emission as in model 7 of Finkbeiner, Davis & Schlegel, we find a clear indication of an anomalous dust-correlated component which is the dominant foreground emission at 31.5 GHz and which is well fitted by a power-law spectral shape $\nu^{-\beta}$ with $\beta \sim 2.5$ in agreement with Banday et al.

Key words: methods: data analysis – techniques: image processing – cosmic microwave background.

1 INTRODUCTION

Current and future cosmic microwave background (CMB) space missions such as the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) (launched in 2001 June; Bennett 1996) and *Planck* (scheduled for launch in 2007 February; Tauber 2000) will map the microwave sky emission over the entire sky with an unprecedented combination of sensitivity and angular resolution. In the meantime, a large variety of ground-based and balloon-borne experiments will provide accurate data over various regions of the sky.

The main goal of such experiments is to produce a map cleaned of any contributions from other sources between us and the Last Scattering Surface and thus containing only genuine CMB anisotropy. From such maps the angular power spectrum is evaluated and used to determine cosmological parameters with high accuracy and thereby to probe various scenarios for structure formation. Potential

⋆E-mail: davide.maino@mi.infn.it
†Current address: Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena CA 91109, USA.

foreground contaminants include emission from our Galaxy (mainly synchrotron, free–free and dust emission), compact galactic and extragalactic sources, and the thermal and kinematic Sunyaev–Zel'dovich effect from clusters of galaxies. In order to exploit the cosmological information encoded into the CMB angular power spectrum, it is crucial to be able to identify and remove such signal components with high accuracy and reliability (Bouchet & Gispert 1999).

Much recent work has been performed in this area by several authors (de Oliveira-Costa & Tegmark 1999; Tenorio et al. 1999; Hobson et al. 1998, 1999; Stolyarov et al. 2001; Prunet et al. 2001) employing a range of techniques from the classical Wiener filter to maximum entropy methods. These algorithms are referred to as 'non-blind' in the sense that they require some a priori information on the signal to be separated – e.g. spatial templates and frequency dependences of the underlying components, although see Barreiro et al. (2003) for a variation on this approach using a maximum entropy based technique. Recently, a blind approach has been proposed (Baccigalupi et al. 2002; Maino et al. 2002) and applied to simulated sky maps similar to those that *Planck* will produce. The approach appears to be very promising in that it is fast and does not need priors about the underlying signals (at least for high sensitivity missions such as *WMAP* or *Planck*). It is therefore possible to use such blind algorithms to obtain priors that can be fed into classical Bayesian separation algorithms.

However, to date, these blind algorithms have not been applied to real CMB data to assess and validate their reliability in a real world situation. We present here the first application of the Fast Independent Component Analysis (FastICA) to real CMB measurements from the *Cosmic Background Explorer* Differential Microwave Radiometer (*COBE*-DMR) 4-yr data.

This paper is organized as follows. In Section 2 we briefly present the component separation problem and the main simplifications required for the application of FastICA. In Sections 2.1 and 2.2 we summarize the basics of the FastICA algorithm and its application to *COBE*-DMR data. In Section 3 we show the impact of the noise amplitude and distribution on FastICA outputs, which drives the choice of optimal separation technique. The main CMB-related results are presented in Section 4 while in Section 6 we consider the implications for foreground emission. A critical discussion follows in Section 7.

2 COMPONENT SEPARATION PROBLEM

Let us suppose that the observed sky radiation is the superposition of N different physical processes and its frequency and spatial dependences can be factorized into two separated terms:

$$\tilde{x}(\boldsymbol{r}, \nu) = \sum_{j=1}^{N} \bar{s}_j(\boldsymbol{r}) f_j(\nu). \tag{1}$$

This signal is, in general, observed by an experiment with M frequencies through an optical system, whose beam pattern is in general modelled at each frequency as a shift-invariant point spread function $B(\boldsymbol{r}, \nu)$. Let us further suppose that $B(\boldsymbol{r}, \nu)$ is frequency-independent at least within each frequency bandwidth $t(\nu)$. In addition, any real experiment adds some instrumental noise to the output $\epsilon_\nu(\boldsymbol{r})$. Following our assumptions the observed signal at a frequency ν is given by

$$x_\nu(\boldsymbol{r}) = \sum_{j=1}^{N} B_\nu(\boldsymbol{r}) * \bar{s}_j(\boldsymbol{r}) \int t_\nu(\nu') f_j(\nu')\, \mathrm{d}\nu' + \epsilon_\nu(\boldsymbol{r}) \tag{2}$$

$$= B_\nu(\boldsymbol{r}) * \sum_{j=1}^{N} a_{\nu j} \bar{s}_j(\boldsymbol{r}) + \epsilon_\nu(\boldsymbol{r}) \tag{3}$$

where $*$ denotes convolution, and

$$a_{\nu j} = \int t_\nu(\nu') f_j(\nu')\, \mathrm{d}\nu'. \tag{4}$$

Our data model can be further simplified assuming that the radiation pattern of the telescope is frequency-independent, i.e. $B_\nu(\boldsymbol{r}) = B(\boldsymbol{r})$. In this case, equation (2) can be written in vector form as

$$\boldsymbol{x}(\boldsymbol{r}) = \mathbf{A}\bar{\boldsymbol{s}}(\boldsymbol{r}) * B(\boldsymbol{r}) + \boldsymbol{\epsilon}(\boldsymbol{r}) = \mathbf{A}\boldsymbol{s}(\boldsymbol{r}) + \boldsymbol{\epsilon}(\boldsymbol{r}), \tag{5}$$

where each component, s_j, of the vector $\boldsymbol{s}$ is the corresponding source function convolved with the B beam pattern. The matrix $\mathbf{A}$ is the mixing matrix with elements given by the $a_{\nu j}$ coefficients.

We present this derivation in order to stress data model assumptions. It is worth noting that one or more of these might be not justified in real cases. For instance, both the *WMAP* and *Planck* experiments will observe the sky radiation through a telescope with multifrequency receivers and the resulting beam pattern is strongly frequency-dependent. This is up to now one of the main limitations of the FastICA approach to astrophysical component separation and this forced Maino et al. (2002) to further convolve simulated *Planck* sky maps in order to obtain similar beam functions. Furthermore, the noise term $\boldsymbol{\epsilon}(\boldsymbol{r})$ is usually assumed to be additive, signal-independent, white, Gaussian, stationary and uniformly distributed on the sky. These are additional strong assumptions, as the noise spectrum of a real experiment may contain a low-frequency tail due to the so-called $1/f$ noise and different scanning (observing) strategies which distribute the integration time on the sky in a non-uniform way.

2.1 FastICA algorithm

We briefly summarize here the FastICA algorithm. The problem of obtaining both the mixing matrix $\mathbf{A}$ and the signals $\boldsymbol{s}$ from observed data $\boldsymbol{x}$ is unsolvable if additional information is not provided. The ICA approach assumes that

(i) the signals $\boldsymbol{s}$ are independent random processes on the map domain; and

(ii) all the signals, but at most one, have non-Gaussian distribution.

The strategy we exploit here is described in detail in Hyvärinen & Oja (1997) and Hyvärinen (1999) and its application in an astrophysical context can be found in Maino et al. (2002); independent components are extracted maximizing a suitable measure of non-Gaussianity. Indeed, the central limit theorem states that a variable which is a mixture of independent variables is 'more Gaussian' than the original ones. Therefore, we have to find a transformation such that the Gaussianity of the variables is reduced; this is equivalent to finding a set of transformed variables that are 'more independent' than the original ones. Furthermore, as our data model is noisy, we have to define a measure of non-Gaussianity that is robust against noise; this is the the so-called neg-entropy.

Approximations to neg-entropy have been given by Hyvärinen & Oja (2000) and Hyvärinen (1999) and, if the noise has the properties assumed previously and its covariance matrix is known, the Gaussian moments of the transformed variables $\boldsymbol{y} = \mathbf{W}\boldsymbol{x}$ are shown to be robust estimates of the desired functions. Here the matrix $\mathbf{W}$ is the separation matrix such that the $\boldsymbol{y}$ components are in fact independent.

The algorithm needs a pre-processing step (Hyvärinen 1999) in which the input maps are 'quasi-whitened'. This reduces the number of unknowns in the problem. Let us assume that we know the covariance matrix $\mathbf{\Sigma}$ of the instrumental noise; at each frequency, the mean value is removed from the data (the offsets of each independent component can be recovered at the end of the separation process) and their covariance matrix $\mathbf{C}$ is evaluated by computing the following expectation value:

$$\mathbf{C} = \langle \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}} \rangle. \tag{6}$$

A modified covariance matrix and quasi-whitened data sets are respectively given by

$$\hat{\mathbf{\Sigma}} = (\mathbf{C} - \mathbf{\Sigma})^{-1/2}\mathbf{\Sigma}(\mathbf{C} - \mathbf{\Sigma})^{-1/2}, \tag{7}$$

$$\hat{\boldsymbol{x}} = (\mathbf{C} - \mathbf{\Sigma})^{-1/2}\boldsymbol{x}. \tag{8}$$

The separation matrix $\mathbf{W}$ is estimated row by row, i.e. one component at a time. Let $\boldsymbol{w}$ be an M vector such that $\boldsymbol{w}^{\mathrm{T}}\hat{\boldsymbol{x}}$ gives one component of the transformed vector $\boldsymbol{y}$ ($\boldsymbol{w}^{\mathrm{T}}$ is a row of the separation matrix $\mathbf{W}$). In order to find an estimation of the transformed vector $\boldsymbol{y}$ that is robust against noise, the following iterative algorithm is applied, together with a convergence criterion:

(i) choose an initial vector $\boldsymbol{w}$;
(ii) update it by means of

$$\boldsymbol{w}_{\mathrm{new}} = \langle \hat{\boldsymbol{x}} g(\boldsymbol{w}^{\mathrm{T}}\hat{\boldsymbol{x}}) \rangle - (I + \hat{\mathbf{\Sigma}})\langle g'(\boldsymbol{w}^{\mathrm{T}}\hat{\boldsymbol{x}}) \rangle$$

where g is a regular non-quadratic function, i.e. $g(u) = u^3$, $g(u) = \tanh(u)$ and $g(u) = u\exp(-u^2)$;
(iii) normalize $\boldsymbol{w}_{\mathrm{new}}$ to be a unit vector;
(iv) compare $\boldsymbol{w}_{\mathrm{new}}$ with the old computed value; if not converged return to (ii); if converged, begin another row.

This procedure maximizes the non-Gaussianity of the component $\boldsymbol{w}^{\mathrm{T}}\hat{\boldsymbol{x}}$. If k rows of the matrix $\mathbf{W}$ are found at a given time, the $k + 1$ row is searched for in a subspace orthogonal to the first k rows. To this purpose a orthogonalization procedure (e.g. by mean of the Gram–Schmidt rule) is inserted between points (ii) and (iii). Once the separation matrix $\mathbf{W}$ is obtained, the underlying components are derived by using

$$\boldsymbol{x} = \mathbf{W}^{-1}\boldsymbol{y}. \tag{9}$$

This equation allows us to derive the frequency scalings for each independent component; the scaling between ν and ν' of the jth component is given by the ratio of $W^{-1}_{\nu j}/W^{-1}_{\nu' j}$.

It is also possible to recover the signal-to-noise (S/N) ratio for the reconstructed components. Because the noise covariance $\mathbf{\Sigma}$ is supposed to be known, noise constrained realization $\boldsymbol{n}_x$ for each frequency channel can be built. Once we have the separation matrix $\mathbf{W}$, the noise realization in FastICA, outputs are given by $\mathbf{W}\boldsymbol{n}_x$. Noise is transformed like signals and an estimation of the noise in the reconstructed component is achieved. This is quite useful. Suppose that we have M frequency channels but only $N < M$ components have to be separated. Using all M channels we can reconstruct M components and the 'fictitious' ones will be characterized by a S/N ratio lower than unity.

2.2 Application to *COBE*-DMR 4-yr data

The DMR experiment has observed full-sky microwave emission at three frequencies (31.5, 53 and 90 GHz). At each frequency, two separate channels (denoted A and B) measured the temperature difference between two horn antennas of angular resolution $\sim 7^\circ$ pointing in directions separated by 60° on the sky. This differencing information was then used to reconstruct full-sky maps of the microwave sky.

The instrumental noise is almost white and Gaussian but, due to the scanning strategy, it is not uniformly distributed on the sky; variations by a factor of $\sim$4 are typically observed at 90 and 53 GHz while a factor of $\sim$6 is found at 31.5 GHz. This larger factor reflects the fact that some of the original 31-GHz data have been discarded due to a well-known systematic effect (Kogut et al. 1996). In this way, the 31.5-GHz channels show a noise distribution on the sky which is substantially different from that at 53 and 90 GHz. Furthermore, the S/N ratio is quite poor; for 10° effective resolution smoothed maps, this is $\sim$0.5, 1.5 and 1 when combining A and B channels at 31.5, 53 and 90 GHz respectively, and it becomes $\sim$2 when combining together all the frequencies. However, no extra-smoothing is required before applying FastICA because the angular resolution is the same for each channel.

We have applied the FastICA algorithm to the *COBE*-DMR 4-yr data in HEALPix[1] format (Górski, Hivon & Wandelt 1999) with a resolution parameter $N_{\mathrm{side}} = 32$ corresponding to 12 288 pixels in the sky with size $\sim 1.83^\circ$. The A and B radiometers at each frequency are combined to obtain a sum map (A + B)/2.

Maino et al. (2002) have shown that the optimal CMB reconstruction, both in terms of frequency scaling and offset normalization, can be achieved by exploiting data at high galactic latitudes. We use the extended custom galactic cut (Banday et al. 1997) to select regions useful for this CMB reconstruction. Before applying the code, the best-fitting monopole, dipole and quadrupole – computed using only those pixels surviving the Galactic cut – are subtracted from the maps. FastICA 'naturally' removes a monopole from the input sky maps but does not do the same for dipole and quadrupole distributions. These modes are removed as we expect that structures on the largest angular scales are contaminated most significantly by a galactic emission which is characterized by a steeply falling power spectrum. Because one figure of merit for the separation performed by FastICA is the frequency scaling, this procedure minimizes the contamination of the derived CMB scaling by foreground emission.

We work with all the non-quadratic functions described before, $g(u) = u^3$, $g(u) = \tanh(u)$ and $g(u) = u\exp(-u^2)$ identified by p, t and g, respectively.

Finally, we make use of the S/N ratio and frequency scaling of each reconstructed component to assess its physical nature.

3 IMPACT OF NOISE DISTRIBUTION AND AMPLITUDE

One of the assumptions of FastICA is that the noise is Gaussian, stationary and uniformly distributed on the sky. The latter is manifestly not true for DMR. In order to validate the impact of noise amplitude and distribution on the reconstruction procedure, we performed $\sim$3000 simulations of fake CMB skies (characterized by a Harrison–Zel'dovich initial power-law spectrum with an rms quadrupole normalization of $\sim 18\ \mu$K) and instrumental noise. Specifically we consider three cases:

(i) DMR nominal noise amplitude and distribution;
(ii) uniform noise distribution (i.e. each pixel is observed the same number of times) with the noise amplitude matched to the

[1] See http://www.eso.org/science/healpix/.

Table 1. Reconstructed CMB frequency scaling between 90 and 53 GHz for our 3000 MC simulations for three different models of the noise distribution as explained in the text. There is a clear degradation of the result when the 31.5-GHz channel is included in the analysis.

Combination	i	ii	iii
31.5:53:90	0.871 ± 0.477	0.879 ± 0.188	0.876 ± 0.019
53:90	0.871 ± 0.205	0.878 ± 0.192	0.876 ± 0.019

mean noise amplitude in the actual DMR data (0.322, 0.100 and 0.139 mK at 31.5, 53 and 90 GHz, respectively);

(iii) uniform noise distribution with amplitude reduced by a factor of 4 with respect to the nominal values.

For each realization we performed a component separation on the cut sky with the *p*, *g* and *t* functions. We have considered the cases where all three frequencies are included, or just the 90- and 53-GHz data (as we know that the 31.5-GHz channel shows a different noise pattern on the sky). In each case, we evaluate the frequency scaling (between 90 and 53 GHz) of the reconstructed CMB component as well as its angular power spectrum on the cut sky, exploiting the technique described in Górski (1994) and in Górski et al. (1996, hereafter G96).

Table 1 reports the results for the derived frequency scaling for the *p* function (similar results are obtained with *g* and *t*). The results are almost identical for noise cases (ii) and (iii) with either two or three DMR frequencies, although naturally, with higher sensitivity, a higher accuracy is achieved for the derived CMB frequency scaling. However, the situation is quite different for the nominal noise distribution and amplitude (case i). Although the mean value is almost the same and consistent with the expected theoretical value, the rms is degraded to some extent, being 0.477 and 0.205 for three and two frequency channel separation, respectively. Therefore, a first indication of the relevance of the noise distribution is derived from an inspection of the frequency scalings.

A similar situation is found in the analysis of the reconstructed CMB component power spectrum as shown in Fig. 1. The CMB component recovered when all three DMR sum maps are inputs to the FastICA algorithm shows an angular power spectrum (dot-dashed line) that is suppressed with respect to that derived from only 90- and 53-GHz inputs (solid line), although they do agree at the 1-σ level (dark grey and light grey shaded regions for all DMR frequencies and only 90 and 53 GHz, respectively). The reason for this behaviour is probably associated with the distinct noise distribution of the 31.5-GHz channel relative to the other two frequencies, which have considerably more consistent noise patterns. This drives the FastICA algorithm to identify two components: one which has the expected CMB frequency scaling and a second with an unphysical (negative) frequency scaling (and interestingly a higher S/N ratio) into which part of the CMB power has been aliased.

We conclude that differences in the noise patterns between the DMR channels are a potential source of unphysical components reconstructed with FastICA, and can have a notable effect on the frequency scaling of the reconstructed CMB component and its angular power spectrum. Therefore, in what follows we restrict analysis to the 90- and 53-GHz channels only.

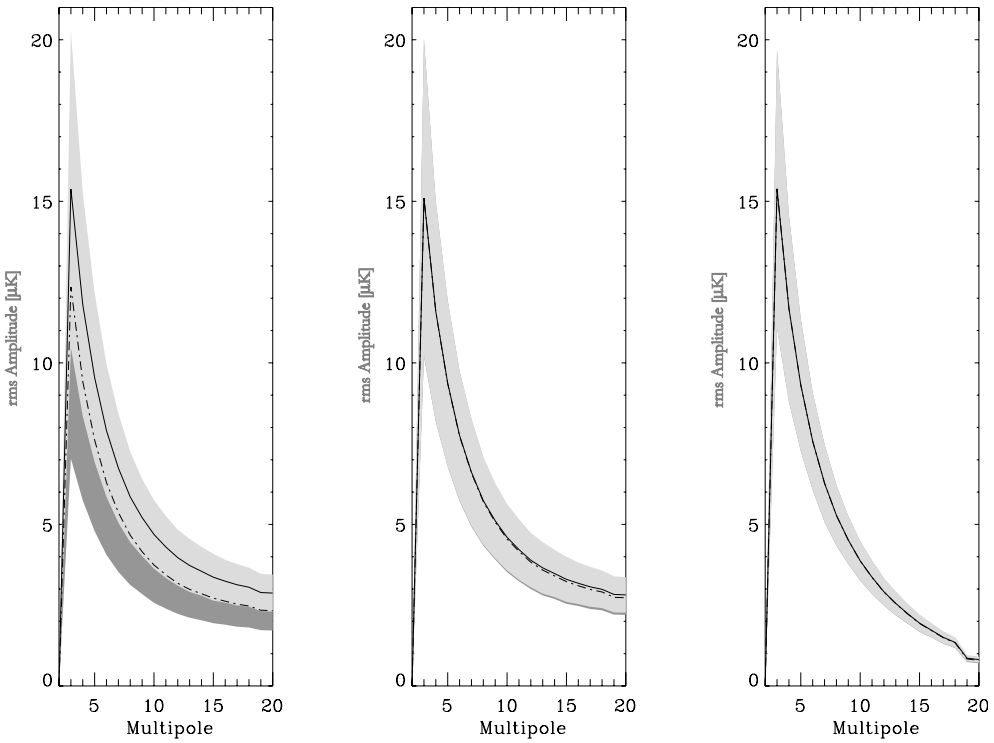

Figure 1. Reconstructed CMB angular power spectrum from our ∼3000 Monte Carlo (MC) realizations of fake CMB skies and instrumental noise. The left panel shows results from the nominal noise case, the middle panel is for uniform noise with nominal mean amplitude and the right panel is for uniform noise with reduced amplitude. Solid lines refer to the 53- and 90-GHz combination while dashed lines include 31.5 GHz in the analysis. The light grey shaded area refers to the 1-σ limit for 53- and 90-GHz channels while dark grey is the 1-σ limit also with 31.5 GHz. These shaded areas nearly fully overlap in all cases except for that with nominal noise. Including the 31.5-GHz channel in the analysis leads to power spectra that are suppressed with respect to the case in which it is excluded. Small deviations are visible also for the uniform noise and mean nominal amplitude case.

4 RESULTS – CMB

We have applied FastICA to the 90- and 53-GHz frequency channels. The results are almost identical for the three non-quadratic functions considered. Table 2 summarizes the reconstructed CMB frequency scalings and the percentage deviation from the theoretical CMB frequency scaling. Results from the *p* and *g* functions are consistent in terms of the correct frequency scaling and S/N ratio. Results determined using the *t* non-quadratic function reveal a non-optimal frequency scaling for the CMB component. Nevertheless the correct normalization can be derived.

Fig. 2 shows the angular power spectrum of the reconstructed CMB component for the *p*, *g* and *t* functions compared with that computed from the minimum-variance noise combination of the 53- and 90-GHz channels. All results are consistent although the *t* function shows a small excess of power on all scales. At large ℓ this corresponds to a larger noise contribution.

Using the same linear combination coefficients that produce the reconstructed CMB components from the input 90- and 53-GHz maps, it is possible to derive the rms noise properties per pixel for the reconstructions. This noise prescription is used as an input to the likelihood analysis (see Górski 1994, and G96) solving for $Q_{\rm rms-PS}$ and *n*. Table 3 reports the results of these fits, whilst Fig. 3 plots the corresponding likelihood contours.

The results are consistent for *p* and *g* while larger $Q_{\rm rms-PS}$ and *n* values are recovered for *t*. Nevertheless, all are consistent with the fit derived from the optimal minimum-variance combination of the 90- and 53-GHz sky maps without foreground corrections applied.

Table 2. DMR 90–53 GHz reconstructed CMB frequency scaling and S/N ratio. Also reported are the deviations from the theoretical expected CMB frequency scaling.

ICA form	Scaling	Δ	S/N
p	0.82 ± 0.28	6.6 per cent	1.069
g	0.91 ± 0.24	3.0 per cent	1.065
t	0.64 ± 0.23	30 per cent	1.072

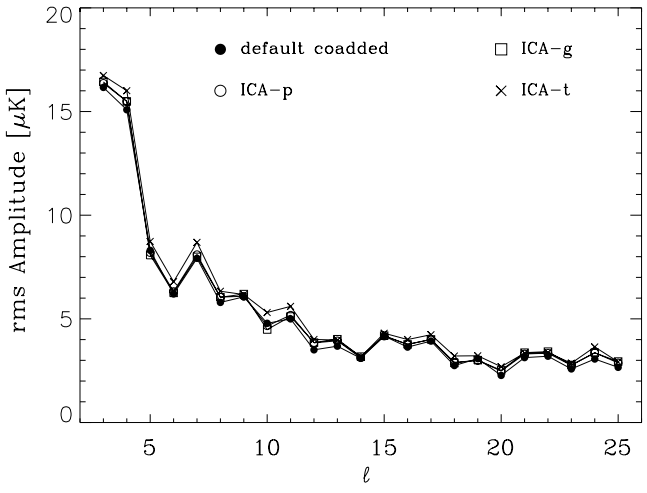

Figure 2. Angular power spectra from the 90- and 53-GHz combination of DMR data (default coadded) and from the reconstructed CMB ICA components with *p*, *g* and *t* functions.

Table 3. $Q_{\rm rms-PS}$ and *n* as derived from the CMB component reconstructed with the FastICA algorithm. We report both 1- and 2-σ confidence intervals and the corresponding results from G96 determined without galactic correction.

Map	$Q_{\rm rms}$ (μK)	$n_{\rm s}$
G96	$15.63^{+3.19+7.67}_{-2.56-4.48}$	$1.25^{+0.22+0.45}_{-0.25-0.53}$
p	$15.63^{+3.19+7.35}_{-2.56-4.48}$	$1.27^{+0.25+0.45}_{-0.27-0.55}$
g	$15.63^{+3.19+7.67}_{-2.24-4.48}$	$1.27^{+0.27+0.45}_{-0.27-0.57}$
t	$16.27^{+3.51+7.67}_{-2.56-4.80}$	$1.29^{+0.25+0.48}_{-0.25-0.54}$

5 ADDING 'PRIORS' ON FOREGROUNDS

In the previous section we detailed the results of our analysis of FastICA as applied to the 53- and 90-GHz sky maps alone, and stressed the consistency of the cosmological results with those determined in G96 without the application of any galactic foreground correction. The relevance of this comparison can be understood as follows. With only two input sky maps, the FastICA algorithm can only reconstruct two outputs, and we have determined that these correspond to CMB together with an unphysical, noise-related component, i.e. the method is not able to reconstruct the galactic emission at high galactic latitudes. Indeed, we have verified by means of simulations – which include a galactic contribution modelled on a DIRBE 140-μm template (scaled to the DMR frequencies according to the correlation coefficients from G96) – that the S/N ratio of the galactic emission in the DMR data is too low to allow FastICA to recover it. Only when the noise rms is reduced by a factor of 10 relative to the actual DMR values can some galactic emission be separated by FastICA, although the separation remains less than optimal.

In order to attempt to extract and better separate the CMB signal from that foreground emission which is certainly present in the data, we have proceeded to add additional channels (priors) using foreground templates such as the Haslam map of the diffuse galactic radio emission at 408 MHz (Haslam et al. 1982) and the DIRBE map of galactic diffuse infrared emission at 140 μm. When applying FastICA to the 53- and 90-GHz DMR data in combination with the Haslam and DIRBE templates, we have assumed that the noise contribution to the template maps is small compared to that in the DMR data, which is certainly the case. Because we now work with four inputs, FastICA returns four derived components. One is again almost consistent with noise (S/N lower that 1), one has the CMB frequency scaling and the expected S/N ratio while the other two show features of the Haslam and DIRBE maps, respectively. It is important to note that FastICA does not completely mix the Haslam and DIRBE signals into a single foreground component, but is able to keep them largely separated (see Section 6). As before, we fit the CMB component for $Q_{\rm rms-PS}$ and *n*. Table 4 reports the results. There is a general agreement between our results, especially those determined with the *p* and *g* functions, and those reported in G96 after galactic emission has been taken into account using the template Haslam and DIRBE maps. The left panel of Fig. 4 shows the CMB map as derived from the DMR 53- and 90-GHz data plus the Haslam and DIRBE 140-μm maps after smoothing to an effective $\sim 10^\circ$ angular resolution to facilitate a direct comparison with previous work in Bennett (1996) and more recently in Barreiro et al. (2003). In fact, such visual inspection of the extracted CMB component does not indicate any obvious differences from the CMB component derived when only the DMR data are used as input to the FastICA method. However, differences are visible in their power

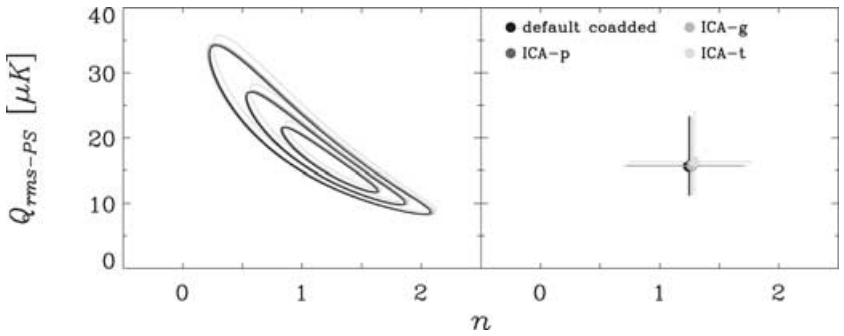

Figure 3. Likelihood contours in the $Q_{\rm rms-PS}$, n plane for the reconstructed CMB component with FASTICA. Also plotted is the 'classic' result for the minimum noise combination of the 90- and 53-GHz DMR channels.

Table 4. $Q_{\rm rms-PS}$ and n determined from an analysis of the CMB component reconstructed by FASTICA when additional foreground channels (Haslam and DIRBE maps – HD) are included. We report both 1- and 2-σ limits and the results from G96 after galactic correction.

Map	$Q_{\rm rms}$ (μK)	$n_{\rm s}$
G96+HD	$14.99^{+3.19+7.35}_{-1.92-4.16}$	$1.27^{+0.20+0.43}_{-0.27-0.55}$
p+HD	$14.03^{+2.56+6.39}_{-2.56-4.16}$	$1.27^{+0.25+0.47}_{-0.27-0.57}$
g+HD	$14.67^{+2.87+7.03}_{-2.24-4.16}$	$1.29^{+0.20+0.46}_{-0.27-0.54}$
t+HD	$11.15^{+2.24+5.44}_{-1.60-3.19}$	$1.34^{+0.18+0.47}_{-0.25-0.53}$

spectrum, as indicated by the right panel of the figure. The power spectrum for the DMR-only case, as compared to that derived when 'priors' from the Haslam and DIRBE maps are included, shows a clear difference in power up to $\ell \sim 14$, mostly due to the subtraction of foreground signal. At higher ℓ, where noise dominates, the power spectrum obtained when including 'priors' is slightly lower in amplitude than the DMR-only case. This is possibly due to the fact that with more input signals FASTICA has a larger number of degrees of freedom to access. Therefore, some of the less Gaussian properties of the noise (related to the non-uniform noise pattern on the sky and perhaps even small mismatches between the 53- and 90-GHz noise patterns) project into the other derived components. Indeed, the extracted foreground components (see Fig. 5) seem to contain some small noise contributions absent from the input foreground maps.

It is clear that p and g almost give the same results while large differences are observed with the t form. This may be due to the low S/N ratio of the DMR data; Maino et al. (2002) did not observe such behaviour when working with higher fidelity *Planck* simulations. Inspection of the extracted components shows that the frequency scalings of the CMB component are almost the same as when only the DMR data are used as inputs (variations are $\lesssim$1 per cent) and that the extracted signals also have the same S/N ratio. We have also computed the final rms in the extracted maps and compared them with the previous DMR-only case. The rms figures are almost the same for p and g (deviations within 8 per cent) while for the t form we observe a variation in the map rms of about 30 per cent. This implies that t is unable to adequately separate the underlying components and identify them as CMB and galactic signals.

6 RESULTS – FOREGROUND EMISSIONS

We have analysed the two additional signal components extracted by FASTICA. These appear to closely resemble the Haslam and DIRBE maps, but this is not optimal (the Spearman correlation coefficient $r_{\rm S} \sim 90$ per cent). Furthermore, visual inspection reveals that some structures which appear in the input Haslam map template appear as 'ghosts' (with negative signal) in the reconstructed dust map, and vice versa as shown in Fig. 5. Such behaviour almost certainly reflects a more complex spectral signature of the foreground components than the method (with only four input data sets) is able to disentangle. Undoubtedly, the significant large angular scale spatial correlation between foregrounds in various regions of the sky and

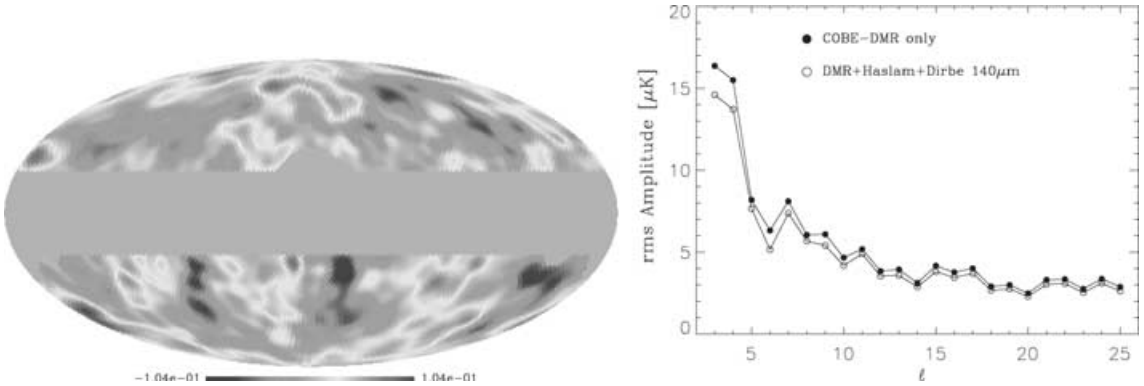

Figure 4. Left-hand panel: the CMB component extracted with FASTICA using the DMR 53- and 90-GHz data and including the Haslam and DIRBE maps. The map has been smoothed to an effective $\sim 10^\circ$ resolution. This map is visually difficult to distinguish from the corresponding component derived using DMR data alone. Right-hand panel: the power spectra of the DMR only case (filled circles) and the DMR, Haslam and DIRBE 140-μm maps (open circles). Differences are clearly visible in the power spectrum with lower power at low ℓ corresponding to the effect of foreground removal, and that at higher ℓ corresponding to a smaller noise contribution remaining in the analysis with 'priors'.

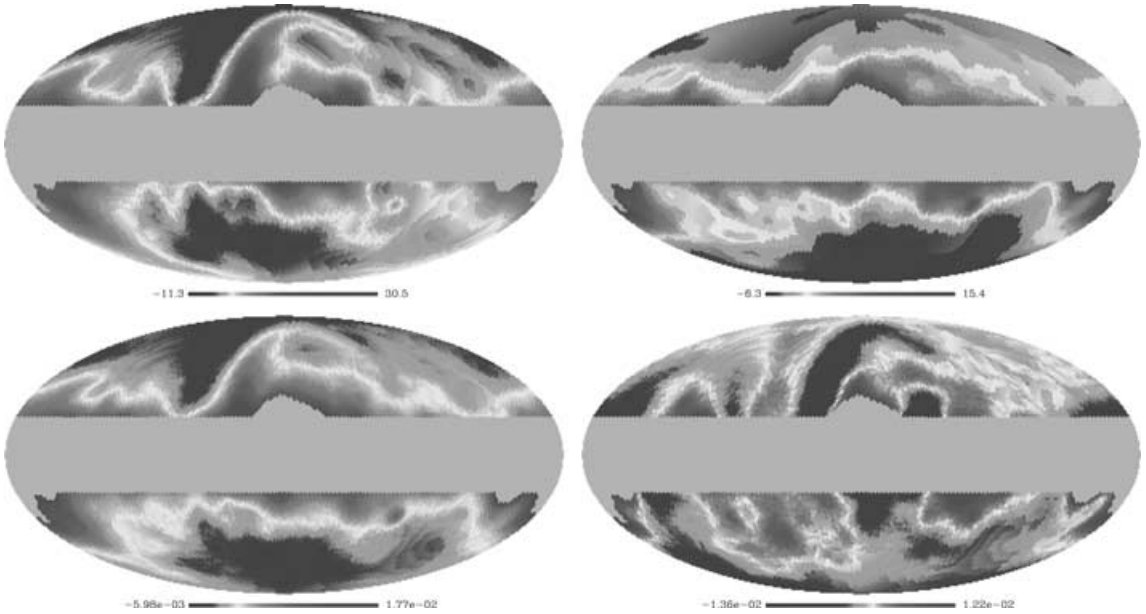

Figure 5. The top panels show the input Haslam 408-MHz (left) and DIRBE 140-μm (right) foreground maps while the bottom panels report the reconstructed ICA component maps derived with the *t* non-quadratic form. The ICA components show large noise and signal ghosts from the other component (e.g. the synchrotron North Polar Spur is visible with a negative amplitude in the reconstructed dust map and high latitude structures in the DIRBE map are aliased into the reconstructed synchrotron map). Units for Haslam and DIRBE maps are K and MJy sr^{-1}, respectively, and are the same for the ICA components (normalized at 90 GHz).

lower fidelity of the foregrounds, due to both the S/N and large angular smoothing of the DMR data, also play a role. In fact, the extent of the signal aliasing depends on the non-quadratic form considered in the analysis, and is particularly evident for the *t* form. The spectral indices for these galactic independent components derived are not robust to the choice of non-quadratic form and furthermore are not consistent with what is expected from previous studies (e.g. G96). However, the CMB component is reconstructed well and remains clean from galactic emission, as results in Table 4 show.

Interestingly, this means that FastICA *can* provide information on foregrounds, not directly from the (spectral) properties of the reconstructed galactic components, but rather from information gleaned in the reconstruction of the CMB component. Each ICA component is obtained by a weighted *linear* combination of the input data, even though the combination coefficients (actually the rows of the matrix **W**) are computed by a non-linear process. This means that we can interpret the derived weights effectively as correlation coefficients between the foreground templates and the DMR data. Indeed, for the previous analysis with the Haslam and DIRBE maps, the recovered values are similar to those found in G96 using an explicit (minimum χ^2) correlation analysis. Furthermore, the consistency of the results between *p*, *g* and, to some extent, *t*, suggest that the method allows a robust estimation of the correlation of the DMR data with the considered templates.

Recently, Banday et al. (2003) have derived correlation coefficients between the DMR data and three foreground templates: the Haslam map at 408 MHz, the DIRBE map at 140 μm and a new map of Hα emission (Dickinson, Davies & Davis 2003) that could be used to determine a reliable free–free template. Furthermore, when thermal dust and free–free emission are properly accounted for, they found a strong anomalous correlation between data at 31.5 GHz and the DIRBE map. This component is found to be well fitted by a spectral shape of the form $\nu^{-\beta}$ with $\beta \sim 2.5$.

We attempt to verify here whether FastICA is able to arrive at similar results, and we therefore extend our analysis to include the Hα template. An initial consideration is that the non-quadratic forms of the FastICA algorithm are, in general, sensitive to different features of the astrophysical components, therefore it is possible that a given form may or may not be particularly well matched to the detection of certain signals on the sky. We therefore calibrate which form is best suited to our current studies, by using results from Banday et al. (2003) to construct a fake sky with CMB, foreground emissions (whose correlation coefficients are known a priori) and ∼10 000 realizations of the DMR instrumental noise. In each run, a single DMR channel is considered together with foreground emission templates and each of *p*, *g* and *t* is employed to evaluate the correlation coefficients. These simulations also provide an indication of the uncertainty in the reconstruction process. Simulations show that *p* and *g* provide consistent results.

Table 5 reports our estimates of the correlation coefficients at each of the DMR frequencies. Fig. 6 plots our results (filled circles) together with those from Banday et al. (2003) (open circles). There is good general agreement between the two sets of results. FastICA also clearly detects foreground emission correlated with the DIRBE dust map at 140 μm and identifies this as the major foreground emission at 31.5 GHz.

Table 5. FastICA correlation coefficients between each of the DMR frequencies and foreground templates. Results are reported for the best calibrated non-quadratic form in each case. Errors are derived from simulations. Units are μK X^{-1}, where *X* are template units.

Component	DMR map (GHz)		
	31.5	53	90
Dust	5.19 ± 1.55	2.58 ± 1.09	1.91 ± 1.14
Synchrotron	1.77 ± 0.90	-0.08 ± 0.67	-0.35 ± 0.61
Free–free	3.19 ± 1.48	0.86 ± 0.68	0.98 ± 0.86

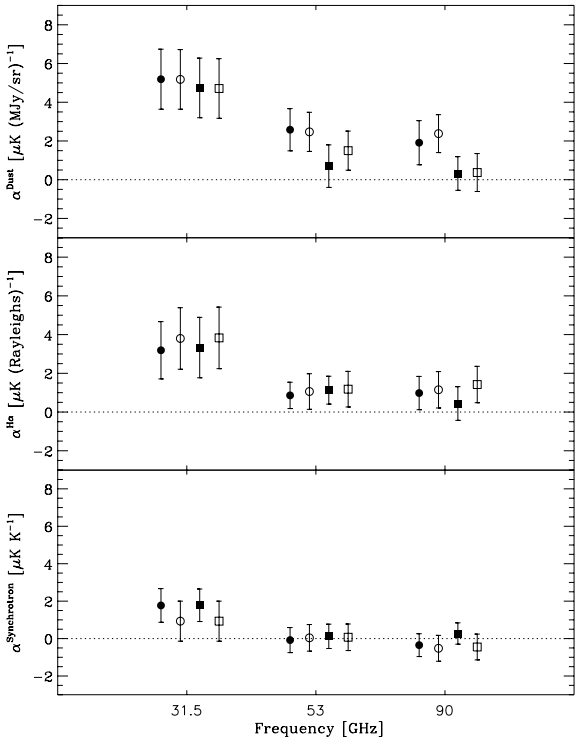

Figure 6. Derived correlation coefficients between DMR at 31.5, 53 and 90 GHz and the three foreground emission templates. Dust emission is traced by the DIRBE map at 140 μm, synchrotron emission by the Haslam map at 408 MHz and free–free emission by the Hα map. Filled symbols are results obtained with FastICA while open symbols are from Banday et al. (2003) using the cross-correlation technique. Circles refer to the test without subtraction of dust thermal model while squares are when model 7 of FDS has been subtracted from the DMR data.

In order to further probe the nature of this anomalous dust correlated component, we subtract from the DMR data model 7 of thermal dust emission from Finkbeiner et al. (1999, hereafter FDS). This comprises two dust components, one at a temperature of 9.6 K with dust emissivity $\alpha = 1.5$ and the other at 16.4 K with $\alpha = 2.6$. The results are shown in Table 6.

It is clear from Table 6 and Fig. 6 that, if the FDS model 7 provides an accurate prediction for the thermal dust emission at DMR frequencies, then it must account for essentially the entire dust correlation at 90 GHz. At 31.5 GHz a strong anomalous correlation with the DIRBE 140-μm dust template persists after the removal of the thermal dust model, whilst the results at 53 GHz seem to indicate the presence of both thermal and anomalous dust correlated components, which would have roughly equal contributions at a frequency of ∼60 GHz.

Table 6. FastICA correlation coefficients between each of the DMR frequencies, after thermal dust emission removal following model 7 in FDS, and foreground templates. Results are reported for the better 'calibrated' non-quadratic form in each case. Errors are derived from simulations. Units are μK X^{-1}, where X are template units.

Component	DMR map (GHz)		
	31.5	53	90
Dust	4.74 ± 1.54	0.70 ± 1.10	0.32 ± 0.87
Synchrotron	1.78 ± 0.87	0.12 ± 0.65	0.27 ± 0.57
Free–free	3.33 ± 1.56	1.13 ± 0.72	0.44 ± 0.87

Table 7. Power-law spectral indices from pairs of frequencies and from a fit using data from 31.5 to 90 GHz. We fit for $\boldsymbol{A}_{\rm norm}$ $(\nu/\nu_0)^{-\beta}$ where ν_0 is taken as 31.5 GHz. We also report the 68 per cent confidence level errors.

$\beta_{31:53}$	3.67 ± 2.01
$\beta_{53:90}$	1.48 ± 2.95
Fitted β	$2.56^{+1.56}_{-0.98}$
Fitted $\boldsymbol{A}_{\rm norm}$	$3.90^{+1.54}_{-0.99}$

We investigate the spectral behaviour of this anomalous dust correlated component by fitting our data points (the correlation coefficients with DIRBE 140 μm after FDS model 7 subtraction) by a power law of the form $\boldsymbol{A}_{\rm norm}(\nu/\nu_0)^{-\beta}$. $\boldsymbol{A}_{\rm norm}$ has units of μK (MJy sr^{-1})$^{-1}$ and is the amplitude of the emission at the reference frequency ν_0 of 31.5 GHz. Table 7 shows the results from this analysis. It is clear that the frequency behaviour is consistent with a spectrum steeper that the typical value of ∼2.15 expected for pure free–free emission (although the results are still compatible with this type of emission).

Such conclusions are consistent with those of by Banday et al. (2003), and are particularly satisfactory due to the independence of the two techniques employed, and the fact that the ICA method makes no assumptions about the cosmological signal present in the data.

7 CRITICAL DISCUSSION AND CONCLUSION

In this paper we have applied, for the first time, the FastICA algorithm to real CMB data and specifically to the *COBE*-DMR 4-yr data.

One of the requirements for applicability of FastICA is that noise has to be Gaussian and uniformly distributed on the sky. Unfortunately, this not the case for the DMR observing strategy. Furthermore, some of the original data at 31.5 GHz were removed to eliminate potential contamination from a particular systematic effect (Kogut et al. 1996), resulting in a noise distribution that is notably different than that at 53 and 90 GHz. Therefore, before applying the technique, we have tested the effect of such a non-uniform noise distribution on both the frequency scaling and power spectrum reconstruction of the recovered CMB component. We run ∼3000 MC simulations and considered three scenarios of different noise distribution in the data: one nominal, one with uniform noise distribution and nominal mean amplitude and the third with uniform distribution and reduced amplitude. The results indicate that the different noise distribution of the 31.5-GHz channel results in the identification of a spurious non-physical component, and the degradation of the frequency scaling of the reconstructed CMB component. Furthermore, the angular power spectrum of the latter has suppressed amplitude with respect to the case in which only the 90- and 53-GHz channels are considered, resulting from the aliasing of CMB power into the non-physical component associated with the noise asymmetry.

Further analysis was therefore restricted to include the 90- and 53-GHz channels only. In this case, we obtained a CMB reconstruction with a frequency scaling consistent with the expected theoretical one. We fit the cosmological component with a power-law

model parametrized by the rms-quadrupole normalization $Q_{\rm rms-PS}$ and spectral index n, and we find that the results are robust against the choice of non-quadratic function, completely consistent with those derived in G96 from the optimal minimum-variance combination of the 90- and 53-GHz channels, without foreground corrections. This is a very important outcome, as equivalent results have been obtained with independent approaches.

Simulations with different noise rms properties have shown that foreground emission in the DMR data has too low a S/N ratio to be detected and separated with FastICA. Only with a S/N ratio a factor of 10 or better does it become possible to identify foreground emission. However, after adding two additional channels – the Haslam map or radio emission at 408 MHz and the DIRBE map of 140-μm galactic infrared emission – FastICA is now able to recognize foreground emission in the DMR data and to separate it from the CMB signal. The CMB frequency scaling for different FastICA forms are consistent with that found without priors on foregrounds. More importantly, we derived values for $Q_{\rm rms}$ and $n_{\rm s}$ for this new cleaned CMB reconstruction in very good agreement with those reported in G96 after foreground correction.

The situation with the foreground component reconstruction is less satisfying, however. We found that the foreground spectral indices derived from the ICA foreground components are not robust against the choice of the FastICA non-quadratic form and this could be due to several reasons. First of all, the different forms tend to preferentially discriminate between different morphological aspects of the foreground emission, with the residuals tending to be mixed together. Secondly, the S/N ratio for foregrounds at high galactic latitudes is too low for FastICA to obtain faithful information on these emissions. Finally, and quite realistically, it is likely that the true foreground emission at microwave frequencies is different from our priors, which are based on observations at considerably different wavelengths. It is worth noting, however, that Maino et al. (2002) have demonstrated that, for higher S/N observations and with an exact correspondence between 'priors' and foreground signals, the algorithm was able to properly extract the underlying foreground contributions.

Despite the fact that the spectral behaviour of the derived foreground components is not properly recovered, our results do still allow some information on the non-CMB emission to be determined. Because the ICA components are linear combinations of the input sky maps, the weightings between the various input data used to reconstruct the CMB sky can be interpreted as correlation coefficients between the DMR data at specific frequencies and the foreground templates, in the spirit of the analysis in G96. Indeed, the results in Table 4 are in substantial agreement with this. Following the recent work of Banday et al. (2003) we have further considered an extra channel using the Hα template provided by Dickinson, Davies and Davis (2003). Using these three templates, FastICA is able to derive the correlation coefficients between the DMR data and foreground templates in very good agreement with their results. We found that after subtraction of the dust thermal emission as accounted for by model 7 of FDS, a strong dust correlated anomalous component remains, and constitutes the strongest foreground emission at 31.5 GHz. This component is found to be well fitted by a power-law spectrum with $\beta \sim 2.56$.

At this point, we would like to comment on some aspects of our component separation relative to another method recently applied to the DMR data. Although the FastICA CMB reconstructed map is noisier than that derived by Barreiro et al. (2003) using a maximum-entropy (MEM) based approach (which results in an effective denoising of the derived physical components), the noise properties of the reconstructed components from our method are well understood and do allow for the derivation of cosmological parameters such as $Q_{\rm rms-PS}$ and n. This is not the case for MEM, where such analyses are currently limited by the considerably more complex noise properties of the derived components.

Blind component separation algorithms such as FastICA are promising and can be applied to real data even if their assumptions are not completely satisfied. The results obtained so far are encouraging and future application to CMB data sets (e.g. *WMAP* and *Planck*) are foreseen.

ACKNOWLEDGMENTS

It is a pleasure to thank E. Salerno for useful discussion. DM thanks A. J. Banday for warm hospitality at MPA and AJB thanks D. Maino for kind hospitality at OAT in Trieste. We thank Clive Dickinson for providing us with the Hα template prior to publication. Some of the results in this paper have been derived using the HEALPix (Górski et al. 1999) package.

REFERENCES

Baccigalupi C. et al., 2000, MNRAS, 318, 769
Banday A. J., Gorski K. M., Bennett C. L., Hinshaw G., Kogut A., Lineweaver C., Smoot G. F., Tenorio L., 1997, ApJ, 475, 393
Banday A. J. et al., 2003, ApJ, submitted
Barreiro R. B., Hobson M. P., Banday A. J., Lasenby A. N., Stolyarov V., Vielva P., Gorski K. M., 2003, MNRAS, submitted (astro-ph/0302091)
Bennett C., 1996, BAAS, 28, 1391
Bouchet F., Gispert R., 1999, New Astron., 4, 443
de Oliveira-Costa A., Tegmark M., 1999, eds, ASP Conf. Ser. Vol. 181, Microwave Foregrounds. Astron. Soc. Pac., San Francisco
Dickinson C., Davies R. D., Davis R. J., 2003, MNRAS, 341, 369
Finkbeiner D. P., Davis M., Schlegel D. J., 1999, ApJ, 524, 867
Górski K. M., 1994, ApJ, 430, L85
Górski K. M., Banday A. J., Bennett C. L., Hinshaw G., Kogut A., Smoot G. F., Wright E. L., 1996, ApJ, 464, L11
Górski K. M., Hivon E., Wandelt B. D., 1999, in Banday A. J., Sheth R., Da Costa L. N., eds, Evolution of Large Scale Structure: from Recombination to Garching. Print Partners Ipskamp, the Netherlands
Haslam C. G. T., Stoffel H., Salter C. J., Wilson W. E., 1982, A&ASS, 47, 1
Hobson M. P., Jones A. W., Lasenby A. N., Bouchet F., 1998, MNRAS, 300, 1
Hobson M. P., Barreiro R. B., Toffolatti L., Lasenby A. N., 1999, MNRAS, 309, 125
Hyvärinen A., 1999, IEEE Signal Process. Lett., 6, 145
Hyvärinen A., Oja E., 1997, Neural Computation, 9, 1483
Hyvärinen A., Oja E., 2000, Neural Networks, 13, 411
Kogut A., 1997, in Bouchet F. R., Gispert R., Guilderdoni B., Tran Thanh Van J., eds, Proc. XVIth Moriond Astrophys. Meeting, Microwave Background Anisotropy. Gif-sur-Yvette, p. 445
Kogut A. et al., 1996, ApJ, 470, 653
Maino D. et al., 2002, MNRAS, 334, 53
Prunet S., Teyssier R., Scully S. T., Bouchet F. R., Gispert R., 2001, A&A, 373, L13
Stolyarov V., Hobson M. P., Ashdown M. A. J., Lasenby A. N., 2001, MNRAS, 336, 97
Tauber J. A., 2000, in Harwitt M., Hauser M. G., eds, IAU Symp. 204, The Extragalactic Infrared Backgounds and its Cosmological Implications. Astron. Soc. Pac., San Francisco, p. 93
Tenorio L., Jaffe A. H., Hanany S., Lineweaver C. H., 1999, MNRAS, 310, 823

This paper has been typeset from a TEX/LATEX file prepared by the author.