Cross-correlation of the phases of the CMB and foregrounds derived from the WMAP data

P. D. Naselsky,1,2,3⋆ A. G. Doroshkevich1 and O. V. Verkhodanov1,4
1Theoretical Astrophysics Centre, Juliane Maries Vej 30, DK-2100, Copenhagen, Denmark
2Niels Bohr Institute, Blegdamsvej 17, DK-2100, Copenhagen, Denmark
3Rostov State University, Space Research Department, Zorge, 5, 344091, Russia
4Special Astrophysical Observatory, Nizhnij Arkhyz, Karachaj-Cherkesia, 369167, Russia

Accepted 2003 December 10. Received 2003 December 3; in original form 2003 October 21

ABSTRACT

We present circular and linear cross-correlation tests and the ‘friend-of-friend’ analysis for phases of the Internal Linear Combination Map (ILC) and the WMAP foregrounds for all $K − W$ frequency bands at the range of multipoles $\ell \lesssim 100$. We also compare Tegmark, de Oliveira-Costa & Hamilton and Naselsky et al. cleaned maps with corresponding foregrounds. We have found significant deviations from the expected Poissonian statistics for all the cleaned maps and foregrounds. Our analysis shows that, for a low multipole range of the cleaned maps, power spectra contain some of the foregrounds residuals mainly from the $W$ band.

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

1 INTRODUCTION

The recently published Wilkinson Microwave Anisotropy Probe (WMAP) data sets (see Bennett et al. 2003a,b,c; Hinshaw et al. 2003a, 2003b) open a new epoch for the cosmic microwave background (CMB) investigation. These data serve for the development of more refined technique for future high resolution measurements and for a choice of the realistic cosmological model.

The WMAP has observed the full sky in five frequency bands: $K$ (centered frequency 22.8 GHz), $Q$ (33.0 GHz), $V$ (40.7 GHz), $K_a$ (60.8 GHz), and $W$ (93.5 GHz) and produced five maps represented in $n = 12 \times 512^2$ HEALPix (Górski, Hivon & Wandelt 1999) pixels. Then, using smoothing of the maps to $1\degree$ and performing cleaning of the combined multifrequency map by minimization of the rms variance for each pixel, the WMAP team produced the Internal Linear Combination (ILC) map taking 12 optimization coefficients $w_i, i = 1, 2, \ldots, 5$ for 12 disjoint sky regions. Thus, the ILC map has a minimal variance corresponding to the minimization of the Galaxy and foreground contamination at the range of multipoles $\ell \lesssim 100$.

Tegmark et al. (2003) (hereafter TOH) have suggested to use ideologically similar but technically different optimization scheme based on Tegmark & Efstathiou (1996) (hereafter TE) method. The main goal of the TOH method is to combine the $K − W$ band maps into the Foreground Cleaned Map (FCM) using variable weighting coefficients $w_i(\ell)$ for each frequency band $i = K − W$. In order to minimize the total unweighted power from the Galaxy, foregrounds and noise separately for each harmonic $\ell$, TOH subdivided the sky into 9 disjoint regions and performed cleaning of the 5 frequency maps in each region. Unlike the ILC map, the TOH FCM has smaller power at the multipole range $\ell \lesssim 12$. Both the ILC and TOH FCM are available from the web and can be used for analysis of the foregrounds and the statistical properties of the cleaned maps. For example, Komatsu et al. (2003) tested the non-Gaussianity of the WMAP CMB signal using Minkowski functionals as a statistic. Chiang et al. (2003) have also tested the statistics of the phases for the ILC and the TOH FCM maps and discussed non-Gaussianity of the TOH FCM map. Dineen & Coles (2004) have presented a diagnostic of the Galactic synchrotron contamination and have discovered the cross-correlation between Faraday rotation measure and the TOH FCM map (see also Coles et al. 2004).

Recently Naselsky et al. (2004) proposed the Phase Cleaning Method (PCM) for decomposition of the CMB signal and foregrounds using the WMAP data. The main idea of the PCM is to minimize the variance of the combined $K − W$ maps and cross-correlations of the CMB and foreground phases simultaneously. The result of PCM in its application to the WMAP data reproduces well the power spectrum of the best-fitting WMAP $\Lambda$CDM model at the range $\ell < 50$. The main target of our paper is to extend the resolution of the derived PCM CMB map up to $\ell \sim 100$, to perform the the cross-correlation analysis (circular and linear statistics and ‘friend-of-friend’ analysis) of the CMB and foreground phases and to estimate their interconnection. We have included the ILC, TOH FCM and the PCM maps in our analysis and have discovered significant deviation of the phase statistics from the expected for uniformly random phases. We have shown that such a kind of non-Gaussianity arises owing to the contamination of the foregrounds signal in the ILC, TOH FCM and the PCM maps.

⋆E-mail: naselsky@tac.dk

© 2004 RAS

Downloaded from https://academic.oup.com/mnras/article-abstract/349/2/695/958689 by guest on 28 July 2018
2 PHASE CLEANING METHOD AND THE CMB EXTRACTION

The fluctuations of measured CMB plus foregrounds radiation on a sky sphere can be expressed as a sum over spherical harmonics:

$$\Delta T(\theta, \varphi) = T(\theta, \varphi) - \langle T \rangle = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta, \varphi),$$  

(1)

where $a_{\ell m}$ are the coefficients of expansion, $\langle T \rangle = 2.73$ K, and $\Delta T = 0$. Homogeneous and isotropic CMB Gaussian random fields (GRFs), as a result of the simplest inflation paradigm, possess $a_{\ell m}$ modes whose real and imaginary parts are independently distributed. The statistical properties of this field are completely specified by its angular power spectrum $C_{\ell}$,

$$\left(a_{\ell m}^{\text{cmb}}(\theta) a_{\ell m'}^{\text{cmb}}(\theta)\right) = C_{\ell}^{\text{cmb}} \delta_{\ell \ell'} \delta_{m m'},$$  

(2)

and random phases

$$\psi_{\ell m}^{\text{cmb}} = \tan^{-1} \frac{2}{\Re \left(a_{\ell m}^{\text{cmb}}(\theta)\right)},$$  

(3)

which are uniformly distributed at the range 0, $2\pi$. For the foregrounds, the signal is obviously non-Gaussian. So, for the combined CMB + foregrounds signal we define

$$a_{\ell m} = |a_{\ell m}| \exp(i\psi_{\ell m}),$$  

(4)

where $|a_{\ell m}|$ is the modulus and $\psi_{\ell m}$ is the phase of each $\ell, m$ harmonic. In practice, each of the WMAP $K - W$ maps is decomposed into set of $a$ where $a \equiv a_{\ell m}$ and index $j = 1, 2, ..., 5$ corresponds to $K, Ka, ...$, $W$ bands, using the HEALPix code (Gorski et al. 1999).1

The basic idea of the PCM is to generalize the TOH and TE minimization scheme including also the minimization of the cross-correlations between derived CMB signal and foregrounds (Naselsky et al. 2004). The method does not require any galactic cut-offs or disjoint regions. Following Naselsky et al. (2004), we consider the combinations of the WMAP maps: $Ka - Q$, $Ka - V$ and $Q - V$, for which the higher correlations between foreground phases takes place.2 For each pair of the maps, we introduce weighting coefficients $w_{\ell}$ (similar to TOH and TE methods and minimize the variance of the derived map per each mode $\ell$. Neglecting the instrumental for $\ell \leq 100$ noise and taking into account the beam shape deconvolution, we consider the signal as a superposition of the CMB and the foregrounds.

From all $a_{\ell}$, we have found the phases $\psi_{\ell} = \psi_{\ell}(\theta, \varphi)$ (see equations 2–4). Each set of the $a_{\ell}^{(j)}$ coefficients is defined by a combination of the different foreground coefficients $\sum_{j} a_{\ell}^{(j)} = G_{\ell}$ and the CMB signal $a_{\ell}^{\text{cmb}}$, where index $j$ marks the synchrotron, free–free and dust emission. For each combination of the maps, the derived CMB map is defined as

$$a_{\ell m}^{\text{cmb}} = \frac{1}{2} \sum_{j=1}^{5} w_{\ell}(j)a_{\ell m}^{(j)} = a_{\ell m}^{\text{cmb}} + \frac{1}{2} \sum_{j=1}^{5} w_{\ell}(j)G_{\ell},$$  

(5)

where

$$w_{\ell}^{(1)}(\ell) = \frac{1}{2(2\ell + 1)} \sum_{m} |a_{\ell m}^{(2)}|^{2} \left\{\left|\cos(\psi_{\ell}^{(2)} - \psi_{\ell}^{(1)})\right| - \left|\cos(\psi_{\ell}^{(2)} + \psi_{\ell}^{(1)})\right|\right\} - \frac{2}{C_{\ell}} \int_{0}^{2\pi} d\xi \left|\psi_{\ell}^{(2)} - \psi_{\ell}^{(1)}\right| \int_{0}^{2\pi} d\xi \left|\psi_{\ell}^{(2)} + \psi_{\ell}^{(1)}\right| \rightarrow \min$$  

(6)

$$w_{\ell}^{(2)}(\ell) = \frac{1}{2(2\ell + 1)} \sum_{m} |a_{\ell m}^{(2)}|^{2} \left\{\left|\cos(\psi_{\ell}^{(2)} - \psi_{\ell}^{(1)})\right| - \left|\cos(\psi_{\ell}^{(2)} + \psi_{\ell}^{(1)})\right|\right\} - \frac{2}{C_{\ell}} \int_{0}^{2\pi} d\xi \left|\psi_{\ell}^{(2)} - \psi_{\ell}^{(1)}\right| \int_{0}^{2\pi} d\xi \left|\psi_{\ell}^{(2)} + \psi_{\ell}^{(1)}\right| \rightarrow \min$$  

(7)

Note that minimization of the weighting variance equation (8) is equivalent to minimization of the error bars of the CMB reconstruction $\sum_{m} |a_{\ell m}^{\text{cmb}} - a_{\ell m}^{\text{cmb}}|^{2} \rightarrow \min$, where $a_{\ell m}^{\text{cmb}}$ is the true CMB (Naselsky et al. 2004).

Practical implementation of the PCM includes the iteration scheme in which the first step corresponds to minimization of the variance of the derived map in each mode $\ell$ with the choice of coefficients equation (6). This step of optimization reconstructs the pre-CMB and pre-foreground map for each pair of channels where the foreground maps are a simple subtraction $a_{\ell m}^{\text{cmb}}$ coefficients of the $K - W$ signals and derived pre-CMB $a_{\ell m}$. Then, using the moduli and phases of the pre-foregrounds, we perform the next iteration, using equations (9) and (5) and so on. This iteration scheme is stable and reproduce quite well the CMB signal from each pair of $K - W$ maps after two steps of iteration.

However, the CMB maps reconstructed from each pair of $K - W$ maps are slightly different because of residues of the foregrounds. To minimize such residues, as the last step of the PCM we use the so-called MIN-MAX filter for each pixels of the map. The MIN-MAX filter compares signals, $\Delta T_{p}$, in each pixel $p$ in a set of maps. In this method, we chose the minimal amplitude of $\Delta T_{p}$ in the pixel as the CMB signal, i.e. $\Delta T_{p} = \arg \min \{\Delta T_{p}^{(i)}\}$, where $|\Delta T_{p}^{(i)}| = \min \{|\Delta T_{p}^{(i)}|\}$. To estimate residues from point sources, we find the maximal amplitude of the signal in each pixel $|\Delta T_{p}^{(i)}| = \Delta T_{p}^{(i)}$, where $|\Delta T_{p}^{(i)}| = \max \{|\Delta T_{p}^{(i)}|\}$, for all the pre-CMB maps produced by the method equation (9). For true CMB, the difference $\Delta T_{p}^{(i)} = \Delta T_{p}^{(i)}$ is equal to zero and we consider $\Delta T_{p}^{(i)}$ as a measure of deviation from the CMB map.

The reason for such a filter is quite obvious. The signal in each pixel is a superposition of the CMB signal and (small) residues of the foregrounds. If correlations between the CMB signal and foregrounds are minimal then we expect that all the deviations of the $\Delta T_{p}^{(i)}$ in the pixel are caused by the residues. In our case, two

© 2004 RAS, MNRAS 349, 695–704

1 In addition we used the glesp pixelization scheme (Doroshkevich et al. 2004a) in order to control the accuracy of the $a_{\ell}^{(j)}$ harmonics estimation.

2 We did not include $W$ and $K$ bands to the separation procedure because of the peculiar phases (Naselsky et al. 2004).
approximations of the CMB map restored by $K_a - V$ and $Q - V$ channels (Naselsky et al. 2004) have been used and $\Delta T_{\text{min}}$ has been taken as the CMB map.

The result of the PCM application to the WMAP data at the range $\ell \leq 100$ is shown in Fig. 1 for the CMB signal. In this figure, the differences between the ILC and PCM maps (top right), TOH FCM and PCM (bottom left) and between TOH Wiener and PCM maps (bottom right) are also shown. As is shown in Fig. 1, the main visual difference is connected with sources in the Galactic plane and the harmonic $a_{51}$ which is probably related to the galactic emission. The difference shows foreground components being absent in the PCM map. Actually, difference of the maps contains a set of low multipoles but a visual structure parallel to the galactic plane is due to the $a_{51}$-harmonic.

For the simplest estimator of the power spectrum, $C(\ell) = (2\ell + 1)^{-1} \sum_{m} |a_{\ell m}|^2$, we show in Fig. 2 the power spectrum for the ILC, FCM, PCM and TOH Wiener filtered maps. As one can see from Fig. 2 the PCM map reproduces well the spectrum for the TOH Wiener filtered map.

![Figure 2](image-url)  
Figure 2. The power spectrum $C(\ell)$ in mK$^2$ for the PCM reconstructed CMB (thick solid line), ILC (dash line), TOH FCM (dash dot line) and TOH Wiener filtered map (thin solid line).

### 3 CORRELATIONS OF THE CLEANED SIGNALS AND FOREGROUNDS FOR THE ILC, FCM AND PCM MAPS

Three cleaned maps (ILC, FCM and PCM) are produced from different methods of the foreground component separation. For the perfect separation of the CMB signal and foregrounds we can expect small random correlations between their phases. However, these methods provide the approximate separation only. Therefore, these remaining cross-correlations between phases of the cleaned signal, $\phi_s$, and the foregrounds, $\psi_f$, can be used to characterize the degree of separation achieved.

To do this, here we will use the simplest phase–phase circular (Fisher 1993) correlation coefficients. We will use also the linear correlation coefficients and more refined ‘friend-of-friend’ statistics discussed, for example, by Roeder (1992) neglecting circularity of the phases. Indeed, equivalence of the phases $\phi_s$ and $\phi_s \pm 2\pi$ and $\psi_f$ and $\psi_f \pm 2\pi$, changes their correlation functions and cumulants and the results obtained with the ‘friend-of-friend’ statistics. To suppress this effect we compared our result with 200 random realizations of Poissonian process prepared in the same manner (see Section 3.3).

For five frequency channels $K - W$, the maps are taken from the WMAP web site, and all the phases are obtained by the spherical harmonics decomposition using the HEALPix (Górski et al. 1999) and GLESP (Doroshkevich et al. 2004a) codes. We consider separately two ranges of multipoles, $2 \leq \ell \leq 50$ and $51 \leq \ell \leq 100$. For the ILC, FCM and PCM CMB maps, the derived foregrounds are determined as differences between the signal ($S$) and cleaned map ($C$) $F = S - C$ for each $(\ell, m)$. In addition, for the ILC cleaned map, WMAP foregrounds given in the same website were also used. Examples of difference of $S$, ILC and WMAP foregrounds for $V$ and $W$ WMAP bands are shown on Fig. 3.
3.1 Circular cross-correlation of the phases

Following Fisher (1993) we define the statistics

\[ M_p = \frac{1}{\ell_{\text{max}}} \sum_{m=1}^{\ell_{\text{max}}} \exp[ip(\phi_m - \langle \phi_m \rangle)] , \]

\[ M_\delta = \frac{1}{\ell_{\text{max}}} \sum_{m=1}^{\ell_{\text{max}}} \exp[i(\psi_m - \langle \psi_m \rangle)] , \]

where \( \langle \phi_m \rangle = \tan^{-1} \left( \sum_{m=1}^{\ell_{\text{max}}} \sin \phi_m / \sum_{m=1}^{\ell_{\text{max}}} \cos \phi_m \right) , \)

\( \langle \psi_m \rangle = \tan^{-1} \left( \sum_{m=1}^{\ell_{\text{max}}} \sin \psi_m / \sum_{m=1}^{\ell_{\text{max}}} \cos \psi_m \right) , \)

\[ R_{\ell}(\ell) = \frac{1}{\ell_{\text{max}} - \ell_{\text{min}} + 1} \sum_{\ell=\ell_{\text{min}}}^{\ell_{\text{max}}} \cos(\phi_m - \psi_m) , \]

\[ (R_{\ell}) = \frac{1}{\ell_{\text{max}} - \ell_{\text{min}} + 1} \sum_{\ell=\ell_{\text{min}}}^{\ell_{\text{max}}} R_{\ell}(\ell) , \]

where \( M_p \) and \( M_\delta \) are the pth trigonometric moments of the samples, \( \phi \) and \( \psi \) are corresponding mean directions, \( R_{\ell}(\ell) \) is the circular cross-correlation coefficient in each mode \( \ell \) and \( r_{\ell} \) is the mean circular cross-correlation coefficient for all phases. For \( m = 0 \) and for all \( \ell \) phases \( \phi(\ell, 0) = \psi(\ell, 0) = 0 \) and here we neglect them.

For the \( K - W \) bands the circular coefficients, \( R_{\ell}(\ell) \), are plotted in Fig. 4 for the ILC cleaned signal and its own and derived foregrounds and for PCM signal and derived foregrounds. For the first three channels, these coefficients are quite moderate and do not exceed the random scatter (1σ) obtained from 200 random realizations. Note that, for all the bands, the shapes of the functions \( R_{\ell}(\ell) \) are quite similar to each other. This fact reflects strong correlation of phases in all the foregrounds (Naselsky et al. 2004).

From Fig. 4, for the \( V \) and \( W \) channels the cross-correlations of both the PCM and ILC phases with the derived foregrounds seems to be quite significant.

The same tendency is seen from the estimations of the mean coefficients, \( R_{\ell} \), for \( 2 \leq \ell \leq 50 \) listed in Table 1 for all three cleaned signal and the own WMAP and three derived foregrounds. For first three channels these coefficients are small but they become significant for channels \( V \) and \( W \) and derived foregrounds.

Correlations of phases of different foregrounds with the ILC, PCM and FCM cleaned signals are plotted in Figs 5–7. For the free–free foregrounds, correlations are quite moderate and they exceed the random scatter only for the 35–36 harmonics. However, for the synchrotron and dust foregrounds we see the significant (95 per cent) correlations for the 21–23 harmonics for all cleaned signal maps. In all the cleaned maps this range corresponds to minima of the power (see Fig. 2).

3.2 Linear cross-correlation of the phases

For \( \ell_{\text{min}} \leq \ell \leq \ell_{\text{max}} \), the mean linear correlation coefficient between phases of the cleaned signal, \( \phi \), and foregrounds \( \psi \), is defined as follows:

\[ r_{\ell}(\ell_{\text{min}}, \ell_{\text{max}}) = \frac{\langle \phi(\ell) \psi(\ell) \rangle - \langle \phi(\ell) \rangle \langle \psi(\ell) \rangle}{\sigma_{\phi} \sigma_{\psi}} , \]

\[ \langle \phi(\ell) \rangle = \frac{1}{N_{\ell}} \sum_{\ell_{\text{min}}}^{\ell_{\text{max}}} \sum_{m=1}^{\ell} \phi(\ell, m) \approx \pi , \]
Cross-correlation of the phases of the CMB

$$\sigma_\phi^2 = \frac{1}{N_\ell} \sum_{\ell=\ell_{\text{min}}}^{\ell_{\text{max}}} \sum_{m=1}^l \phi_\ell^2 \approx \pi^2/3,$$

$$\langle \phi_\ell \psi_\ell \rangle = \frac{1}{N_\ell} \sum_{\ell=\ell_{\text{min}}}^{\ell_{\text{max}}} \sum_{m=1}^l \phi_\ell \psi_\ell,$$

$$N_\ell = (\ell_{\text{max}} - \ell_{\text{min}} + 1)(\ell_{\text{max}} + \ell_{\text{min}})/2.$$  \hspace{1cm} (11)

For the foreground phases, $\psi_\ell$, the mean value, $\langle \psi_\ell \rangle$, and the variance, $\sigma_\psi$, are defined with similar relations. Here $N_\ell$ is the number of phases in the sample under consideration. For $m = 0$ and for all $\ell$, phases $\phi_\ell(\ell, 0) = \psi_\ell(\ell, 0) = 0$ and here we neglect them.

The linear correlation coefficients for $2 \leq \ell \leq 50$ and $51 \leq \ell \leq 100$ are listed in Table 2 for the ILC, FCM and PCM cleaned signals and five frequency channels. First three columns characterize the correlation with the ILC foregrounds. Six other columns represent these coefficients for the same signals and derived foregrounds. For six samples the correlations between the signal and derived foregrounds, $r_{sf}(\ell_{\text{min}}, \ell_{\text{max}})$, increase with the channel frequency and are comparable for a given frequency. The random scatter of the coefficient determined by averaging of 200 random realizations is $r_{sf}(\ell_{\text{min}}, \ell_{\text{max}}) \approx 0.06$ at 68 per cent CL. As is seen from Table 1, for high-frequency channels $V$ and $W$ the measured $r_{sf}(\ell_{\text{min}}, \ell_{\text{max}})$ exceeds this value. This fact indicates the noticeable correlation between the cleaned signal and derived foregrounds in channels $V$ and $W$ and the limited precision of separation methods.

For all the channels and WMAP foregrounds, the coefficients $r_{sf}(\ell_{\text{min}}, \ell_{\text{max}})$ are less than the random value that indicates a high efficiency of corrections of foregrounds. However, this result depends upon the foregrounds rather than the cleaned signal. For the FCM and PCM cleaned signal and the same foregrounds, the coefficients $r_{sf}(\ell_{\text{min}}, \ell_{\text{max}})$ are also listed in the same table. For all samples these coefficients $r_{sf}(\ell_{\text{min}}, \ell_{\text{max}})$ do not exceed the random value. Therefore, the small correlations between the foregrounds of ILC and the cleaned signals is determined by the properties of foregrounds.

By definition, this coefficient is more sensitive to $\ell_{\text{max}}$ because the majority of phases used comes from $\ell \sim \ell_{\text{max}}$. However, it decreases when we correlate phases with different $\ell$.

3.3 Linear correlation coefficient of the phases for each multipole $\ell$

Linear correlation coefficient of phases of cleaned signal and foregrounds, $r_{\phi}(\ell)$, can also be found for each $\ell$ in the same manner as was done for the circular cross-correlation coefficient in Section 3.1. In this case, the coefficient is defined by the same relations...
(11) neglecting the summation over \( \ell \) and for \( N_\ell = \ell \). It allows one to characterize properties of each mode separately and to determine harmonics with maximal \( r_{cl}(\ell) \). However, the small sample statistics of phases at moderate \( \ell \) increases its random scatter significantly. Note also that, as is seen from the definition, this coefficient characterizes the sample used. Thus, in contrast with the circular cross-correlation coefficient, the direct averaging of \( r_{cl}(\ell) \) over \( \ell \) does not reproduce results obtained in previous Section.

The coefficients \( r_{cl}(\ell) \) are plotted in Fig. 8 for channels \( V \) and \( W \) for two foregrounds of the ILC cleaned signal and for the PCM cleaned signal. For each \( \ell \) the error bars are found for 200 random realizations and are well consistent with the expected behavior \( \approx \ell^{-1/2} \).

For channels \( K-V \) and for all three samples, the correlation coefficients do not exceed the random scatter but the disposition of their maxima and minima are very similar. For the channel \( W \) we see significant differences between functions \( r_{cl}(\ell) \) for different samples. More pronounced correlations are seen for \( 4 \leq \ell \leq 10, \ell = 13-15, \ell = 18-27, \ell = 32-36 \) and \( \ell = 45-50 \). As is seen from the comparison with Fig. 2, some of these harmonics correspond to local minima in \( C_\ell \). This fact indicates that, perhaps, these peculiarities are caused by the same sources as the Galactic dust emission.

### Table 2

| channel   | W-ILC | W-FCM | W-PCM | ILC-50 | ILC-100 | FCN-50 | FCN-100 | PCN-50 | PCN-100 |
|-----------|-------|-------|-------|--------|---------|--------|---------|--------|---------|
| \( K \)   | 0.022 | 0.019 | 0.026 | 0.032  | 0.047   | 0.035  | 0.022   | 0.042  | 0.056   |
| \( K_a \) | 0.028 | 0.032 | 0.041 | 0.034  | 0.046   | 0.030  | 0.027   | 0.059  | 0.048   |
| \( Q \)   | 0.028 | 0.034 | 0.045 | 0.054  | 0.069   | 0.048  | 0.040   | 0.065  | 0.074   |
| \( V \)   | 0.023 | 0.028 | 0.051 | 0.073  | 0.116   | 0.081  | 0.088   | 0.113  | 0.121   |
| \( W \)   | 0.033 | 0.033 | 0.057 | 0.143  | 0.191   | 0.142  | 0.167   | 0.194  | 0.209   |

#### 3.4 ‘Friend-of-friend’ statistics

To compare the foregrounds and the foreground-cleaned signal with random Poissonian process we use a more refined technique, the ‘friend-of-friend’ approach, for the 2D phase diagram. In these diagrams each point has coordinates \( x_\ell = \phi_\ell(\ell,m), y_\ell = \psi_\ell(\ell,m) \) where \( \phi_\ell(\ell,m) \) and \( \psi_\ell(\ell,m) \) are the phases of the cleaned signal and the foregrounds, respectively.

For each channel, we find the fraction of clusters, \( f_M \), with a given richness, \( M \), in wide range of linking lengths, \( r_{\text{link}} \), which determines the maximal point separation within a cluster. Here we use the normalized linking length,

\[
l_{\text{link}} = \frac{\pi r_{\text{link}}^2}{2}, \quad (n) = N_{\text{phs}}/4\pi^2,
\]

where \( r_{\text{link}} \) is the actual distance between two points, and \( N_{\text{phs}} \) is the number of points used. As was discussed in White (1979) and Borgani (1995), the fractions \( f_M \) depend on the correlation functions (or cumulants) of all orders that provides high sensitivity to deviations of the sample from the Poissonian-like one. This sensitivity is weak for \( M = 1 \) and 2 and fast increases for larger \( M \). However, the random scatter of \( f_M \) increases for larger \( M \) as well, and the analysis becomes ineffective for \( M \geq 7 \).

For the truly Poissonian sample of points, the correlation functions of all orders are equal to zero and these fractions can be approximately described as follows:

\[
f_M(l_{\text{link}}) \approx \left[1 - \exp(-l_{\text{link}})\right]^M \exp(-l_{\text{link}}),
\]

where the first term describes the probability to find \( M \) separation between points \( \leq l_{\text{link}} \) while the last term gives probability to find larger separation for the \( M+1 \) point. The scatter of \( f_M \) is determined by averaging of \( f_M \) over 200 random realizations of diagrams analysed in the same manner.

To find clusters with a given richness we use the Minimal Spanning Tree technique which, first, allows one to connect all points within a unique tree. Further on, rejection of edges of a tree larger than a chosen linking length transforms the tree to the system of clusters of all richness for a given linking length. The Minimal Spanning Tree technique was proposed in Barrow, Bhavsar & Sonoda (1985) and van de Weygaert (1991) and was applied for analysis of rich galaxy catalogues (see, e.g., Doroshkevich et al. 2001, 2004b).

First, we are testing the hypothesis that the distribution of phases are Poissonian-like and correlations are negligible for both foregrounds and cleaned signal. The main results of our analysis are presented in Figs 9 and 10 where for all five frequency channels variations of fractions \( f_M(l_{\text{link}}) \), \( 1 \leq M \leq 7 \) are plotted versus the linking length, \( l_{\text{link}} \). As is expected, the functions \( f_M(l_{\text{link}}) \) are quite similar to Poissonian ones for \( M = 1 \) and \( M = 2 \), and even for \( M = 3 \) the deviations are moderate. However, for \( M \geq 4 \) the deviations...
Figure 9. Fraction of clusters with richness $M = 2$–$7$ versus dimensionless linking length, $l_{lnk}$ for the PCM reconstructed CMB (thick solid line), ILC (dash line), TOH FCM (thin solid line). The dots represent a set of random samples. The range of multipoles for the phases is $2 \leq \ell \leq 50$. 
Figure 10. Fraction of clusters with richness \( M = 1 \text{–} 6 \) for \( K \) and \( W \) channels versus dimensionless linking length, \( l_{\text{lnk}} \) for the ILC (top panels), TOH FCM (middle panels) and PCM (bottom panels) reconstructed CMB. The range of multipoles used is \( 2 \leq \ell \leq 50 \). The error bars correspond to 1σ level.
from the Poissonian samples become significant, especially for $V$ and $W$ channels. This result illustrates the expected non-Gaussian character of the cleaned signal and its correlations with the foregrounds.

This result is expectable because of non-Gaussian character of the foregrounds signals. However, comparing the phase diagram for the cleaned signal and foregrounds with 100 diagrams prepared for the foregrounds signals, we can quantify their divergences using $\chi^2$ statistics. These results are shown in Table 3. If for $M = 1$ and $M = 2$ we see small $\chi^2 \leq N$, where $N$ is the number of linking lengths used, then for $M \geq 3$ we have $\chi^2 \geq N$. Larger ratio $\chi^2/N$ is found again for higher frequency channels $V$ and $W$. This fact indicates that all the cleaned maps contain some residues from the dust emission.

### 4 CONCLUSION

In this paper we have proposed new methods of investigation for the statistical properties of the signal derived from the WMAP $K$ -- $W$ bands foreground cleaned maps. With this approach we show significant correlations between the phases of the cleaned maps and the foregrounds. We have compared three cleaned maps, ILC, TOH, FCM and PCM, and have investigated variations of the linear cross-correlation coefficients versus the multipole index, $\ell$. We suggest also more sophisticated 'friend-of-friend' statistics. All the methods described in Section 3 show significant cross-correlations between the phases of the cleaned maps and the foregrounds which manifest themselves more clearly for the $W$ band of the WMAP. We have pointed out that some of peculiarities of the ILC and PCM power spectrum, for example, local minima at $\ell \sim 6$, and $\ell \sim 20$--25 are accompanied by significant cross-correlations between phases of the foregrounds and cleaned signal. Such correlations are important indicators for investigation and detection of possible non-Gaussianity of the CMB signal which could be a byproduct of the component separation methods.

### ACKNOWLEDGMENTS

This paper was supported by Danmarks Grundforskningsfond through its support for the establishment of the Theoretical Astrophysics Centre. We thank Max Tegmark et al. for providing their processed maps and making them public with openness. We thank Igor Novikov and Lung-Yih Chiang for useful discussions. We thank TAC CMB collaboration for used GLESP code. We also acknowledge the use of HEALPix package (Górski et al. 1999)\(^4\) to produce $a_{l,m}$ from WMAP maps and some figures.

\(^4\)http://www.eso.org/science/healpix/
REFERENCES

Barrow J. D., Bhavsar S. P., Sonoda D. H., 1985, MNRAS, 216, 17
Bennett C. L. et al., 2003a, ApJ, 583, 1
Bennett C. L. et al., 2003b, ApJS, 148, 1
Bennett C. L. et al., 2003c, ApJS, 148, 97
Borgani S., 1995, in Bonometto S., Primack J., Provenzale A., eds, Dark Matter in the Universe, Varenna 1995. World Scientific Press, Singapore, 126
Chiang L.-Y., Naselsky P. D., Verkhodanov O. V., Way M. J., 2003, ApJ, 590, L65
Coles P., Dineen P., Earl J., Wright D., 2004, MNRAS, submitted (astro-ph/0310252)
Dineen P., Coles P., 2004, MNRAS, 347, 52
Doroshkevich A. G., Tucker D. L., Turchaninov V. I., Lin H., Christensen P. R., 2001, MNRAS, 322, 369
Doroshkevich A. G., Naselsky P. D., Verkhodanov O. V., Novikov D. I., Turchaninov V. I., Novikov I. D., Christensen P. R., 2004a, A&A, submitted (astro-ph/0305537)
Doroshkevich A. G., Tucker D. L., Allam S., Way M. J., 2004b, A&A submitted (astro-ph/0307233)
Fishier N. J., 1993, Statistical Analysis of Circular Data. Cambridge Univ. Press, Cambridge
Górski K. M., Hivon E., Wandelt B. D., 1999, in Banday A. J., Sheth R. S., Da Costa L., eds, Proc. MPA/ESO Cosmology Conf., Evolution of Large-Scale Structure. Print Partners, Ipskamp
Hinshaw G. et al., 2003a, ApJS, 148, 63
Hinshaw G. et al., 2003b, ApJS, 148, 135
Komatsu E. et al., 2003, ApJS, 148, 119
Naselsky P. D., Verkhodanov O. V., Chiang L.-Y., Naselsky P. D., Verkhodanov O. V., Novikov I. D., 2004b, A&A sub-
Coales P., Dineen P., Earl J., Wright D., 2004, MNRAS, submitted (astro-ph/0310252)
Dineen P., Coles P., 2004, MNRAS, 347, 52
Doroshkevich A. G., Tucker D. L., Turchaninov V. I., Lin H., Christensen P. R., 2001, MNRAS, 322, 369
Doroshkevich A. G., Naselsky P. D., Verkhodanov O. V., Novikov D. I., Turchaninov V. I., Novikov I. D., Christensen P. R., 2004a, A&A, submitted (astro-ph/0305537)
Doroshkevich A. G., Tucker D. L., Allam S., Way M. J., 2004b, A&A submitted (astro-ph/0307233)
Fishier N. J., 1993, Statistical Analysis of Circular Data. Cambridge Univ. Press, Cambridge
Górski K. M., Hivon E., Wandelt B. D., 1999, in Banday A. J., Sheth R. S., Da Costa L., eds, Proc. MPA/ESO Cosmology Conf., Evolution of Large-Scale Structure. Print Partners, Ipskamp
Hinshaw G. et al., 2003a, ApJS, 148, 63
Hinshaw G. et al., 2003b, ApJS, 148, 135
Komatsu E. et al., 2003, ApJS, 148, 119
Naselsky P. D., Verkhodanov O. V., Chiang L.-Y., Naselsky P. D., Verkhodanov O. V., Novikov I. D., 2004b, A&A submitted (astro-ph/0310253)
Roeder K., 1992, in Feigelson D., Babu G.J., eds, Statistical Challenges in Modern Astronomy. Springer-Verlag, Berlin, 106
Tegmark M., Efstathiou G., 1996, MNRAS, 281, 1297
Tegmark M., de Oliveira-Costa A., Hamlton A., 2003, Phys. Rev. D, 68, 123523
van de Weygaert R., 1991, PhD thesis, Univ. Leiden
White S. D. M., 1979, MNRAS, 186, 145

APPENDIX A

In this section we would like to introduce an analytical approach for investigation of the cross-correlation coefficient \( r_d(\ell) \) between foregrounds and derived CMB phases for each mode \( \ell \). We define \( r_d(\ell) \) as

\[
r_d^2(\ell) = \frac{3}{4\pi^2} \sum_{m=1}^{\ell} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \Psi_a \Psi_b \Psi_{a,m} \Psi_{b,m} d\Psi_a d\Psi_b
\]

(A1)

where \( k \) marks the foreground derived from the WMAP frequency band. Here we assume that the distributions of both phases are close to homogeneous ones with \( \langle \Psi \rangle = 0, \langle \Psi^2 \rangle \approx \pi^2/3 \) and we average a correlation coefficient over \( m \) for a given \( \ell \). For the phase of the cleaned CMB signal we will use equation (7), which reflects directly the cross-correlation between the derived CMB and foregrounds phases. Then, using the definition \( \Psi = \alpha \Psi', \) where \( \alpha \) is the cross-correlation coefficient between the foreground phases in \( j \)th and \( k \)th channels, from equation (A1) we obtain

\[
r_d^2(\ell) = \frac{3}{\sqrt{2\pi^2}} \sum_{m=1}^{\ell} \frac{1}{|w_{a,m}|} \sum_j w_{j,m} |G_{j'}|^2 J_2(\pi \alpha_{jk})
\]

(A2)

Let us assume that cross-correlation between different foreground phases is very strong and corresponding coefficients are \( \alpha_{jk} \approx 1 \). Then, from (A2) we obtain

\[
r_d^2(\ell) = \frac{3}{\sqrt{2\pi^2}} \sum_{m=1}^{\ell} \frac{1}{|w_{a,m}|} \sum_j w_{j,m} |G_{j'}|^2.
\]

(A3)

For the PCM method, \( w_{j,m} \) coefficients are the functions of \( \ell \), but not \( m \), while ratio \( |G_{j'}|/|a_{j,m}| \) depends on \( m \) for a given value \( \ell \).

From equation (5) we get

\[
|a_{j,m} - a_{j,m}| = \sum_j |w_{j,m} G_{j'}| \]

(A4)

and, because the foregrounds seem to be highly correlated,

\[
\sum_j |w_{j,m} G_{j'}| = \text{sign} \left( \sum_j w_{j,m} |G_{j'}| \right) \sum_j w_{j,m} |G_{j'}|.
\]

(A5)

Thus, the cross-correlation coefficient \( r_d^2(\ell) \) simply describes the accuracy of the reconstruction of the CMB signal.

5 see Naselsky et al. (2004) for details of the phase correlations.

This paper has been typeset from a \TeX/\LaTeX \ file prepared by the author.