

High-Accuracy Method for the Removal of Point Sources from Maps of the Cosmic Microwave Background

© A.T. Bajkova

Main (Pulkovo) Astronomical Observatory, St.Petersburg, Russia
Email: bajkova@gao.spb.ru

Abstract. A new method for removing point radio sources and other non-Gaussian noise is proposed as a means of improving the accuracy of estimates of the angular power spectrum of the cosmic microwave background (CMB). The main idea of the method is to reconstruct fluctuations of the CMB in places contaminated by such emission, while traditional methods simply exclude these regions from consideration, leading to the appearance of "holes" in the resulting maps. The fundamental possibility of reconstructing the CMB signal in such holes follows from the analytical properties of a function with a finite spatial spectrum (the Silk damping frequency). A two-dimensional median filter is used to localize the point radio sources. Results of simulations of the method for maps of modest size are presented. The efficiency of applying the method to reconstruct the CMB from data with limited resolution and contaminated by appreciable pixel noise is investigated. The fundamental possibility of applying the method to reconstruct the CMB distribution in the region of the Galaxy is also established.

1. Introduction

In connection with numerous experiments on measuring the anisotropy of the cosmic microwave background (CMB) with the aim of deriving high-accuracy estimates of the main cosmological parameters of the Universe [1], it is currently of interest to find new approaches to data reduction that could increase the accuracy with which various background contaminating signals of Galactic and extragalactic origin can be removed from radio observations of the CMB. This problem has been widely discussed in the literature [2--4].

We will restrict our consideration to the removal of point-like radio sources and other non-Gaussian noise, which most strongly distorts the angular power spectrum of the CMB at high spatial harmonics (multipoles). This problem has been considered by a number of authors [5--9]. The currently available methods differ primarily in the means used to localize the contaminating signal, with the aim of subsequently removing the corresponding sections from the measured maps.

The following three strategies for the removal of point-like sources are all natural: (1) reconstructing the point sources, then subtracting them from the measured CMB maps; (2) excluding contaminated pixels from the maps; and (3) reconstructing the true values of the CMB in contaminated pixels.

Experience shows that the first strategy leads to larger errors in the angular power spectrum of the CMB than the other two methods, since it is not possible to completely accurately reconstruct the brightness distribution of the point sources when the spatial spectra of the individual distinguished signals overlap [10]. For this reason, maps obtained with this method are characterized by the presence of appreciable non-Gaussian residuals. The second approach, which requires only the determination of the locations of contaminated pixels and their elimination before estimating the CMB power spectrum, is substantially more effective. Precisely this method is currently the most widely used.

However, this second method for the removal of contaminating point sources is likewise not free of drawbacks. First, the total effective area of the map that is used to estimate the angular power spectrum is reduced, leading to appreciable errors for more distant sections of the CMB. Second, the simple elimination of contaminated pixels does not fully restore the Gaussian statistics of the CMB, since it leaves "holes" in the maps that are used. In addition, the sharp boundaries of the cutout sections lead to so-called Gibbs phenomena in the behavior of the spectrum, which are especially strongly manifest at high spatial harmonics, and can appreciably complicate studies of the secondary anisotropy of the CMB.

Note as well that increasing the resolving power of a system used to measure the CMB anisotropy will lead to the detection of a larger number of weak sources [11], so that an increasingly large number of pixels in the CMB map are contaminated by these signals. A special problem is also presented by the strong contamination due to Galactic emission. Simple removal of the zone of Galactic emission, which comprises an appreciable fraction of the overall area of the celestial sky, unavoidably leads to appreciable errors when estimating the angular power spectrum compared to the use of completely uncontaminated data.

Obviously, if the third strategy is realizable in practice, which depends primarily on the noise characteristics of the apparatus, it will be free of the backs of the first two methods. Therefore, we

propose and develop here a new method for the removal of high-multipole non-Gaussian noise, including both the localization of such noise and reconstruction of the true values of the CMB anisotropy in the contaminated regions. For the specific conditions considered in this paper, this could lead to an increase in the accuracy with which the angular spectrum of the CMB can be constructed, compared to the simple elimination of contaminated map pixels.

Thus, our aim is to present and develop a new method for eliminating the contamination of a CMB signal by point sources and other sources of high-multipole noise. We estimate the internal accuracy of the method and its stability to the input noise, and also investigate its application to data obtained with finite resolution for various levels of the instrumental pixel noise, thereby evaluating possibilities for applying the method to real systems. The following sections of the paper present a statistical description of the CMB anisotropy; describe our model for the observed map, the method used to localize the point sources, and the method used to reconstruct the CMB values in noise-contaminated regions.

2. Model for an observed map of the CMB

The distribution of the CMB temperature over the celestial sphere can be presented by the following expansion in spherical harmonics $Y_\ell^m(\theta, \phi)$ [12]:

$$\frac{\Delta T}{T} = \sum_{\ell, m} a_{\ell m} Y_\ell^m(\theta, \phi),$$

where T and ΔT are the mean temperature and temperature fluctuations of the CMB and the $a_{\ell m}$ are the coefficients of the expansion. The angular power spectrum of the fluctuations C_ℓ is determined as the mean square of the coefficients $a_{\ell m}$:

$$C_\ell = \langle |a_{\ell m}|^2 \rangle,$$

where ℓ is the multipole number.

If the fluctuations in the early Universe satisfy Gaussian statistics, as is expected in most cosmological theories, each coefficient $a_{\ell m}$ should be statistically independent. Thus, the power spectrum C_ℓ provides a full statistical description of the CMB anisotropy, which is a fundamental characteristic of the Universe that can be obtained directly from observations via a spherical-harmonic analysis.

For Gaussian fields, the expansion coefficients $a_{\ell m}$, $\ell \neq 0$, likewise represent Gaussian fields with random phases, zero means, and the dispersions

$$\langle a_{\ell m} a_{\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell, \quad \ell \neq 0.$$

Here, we consider sections of the sky that are modest in size. In this case, it is expedient to operate with the Gaussian fields in a flat, two-dimensional space, and to replace the spherical-harmonic analysis with a Fourier analysis, which appreciably simplifies the testing of proposed methods. The CMB fluctuations $\Delta T = T - \langle T \rangle$ can then be generated via the simple calculation of the Fourier series [12]:

$$\frac{\Delta T(\theta_x, \theta_y)}{T} = \sum_{n_u=0}^{N_u-1} \sum_{n_v=0}^{N_v-1} D(n_u, n_v) \exp\left[i \frac{2\pi}{L} (n_u \theta_x + n_v \theta_y)\right], \quad (1)$$

where L is the linear size of the region considered in radians, (θ_x, θ_y) are Cartesian coordinates on the sky (in the spatial region), and (n_u, n_v) is the number of the Fourier component $D(n_u, n_v)$ in the region with spatial frequencies u and v .

The amplitudes of the Fourier components $D(n_u, n_v)$ obey a Gaussian distribution with zero mean and the dispersion

$$\langle |D(n_u, n_v)|^2 \rangle = C_\ell, \quad \ell = \frac{2\pi}{L} \sqrt{n_u^2 + n_v^2}, \quad (2)$$

while the phases are uniformly distributed in the interval $(0, 2\pi)$. Here, C_ℓ is the angular power spectrum of the CMB temperature when it is expanded in spherical harmonics.

Relation (2) describes the circular symmetry of the power spectrum, i.e., its independence of the azimuthal number m , which we will use below as additional *a priori* information when reconstructing the CMB fluctuations in noise-contaminated regions.

Another, more important, characteristic of the power spectrum is its finite spatial extent, which follows from the existence of the so-called Silk damping frequency (ℓ_D) [13], above which the power-spectrum fluctuations fall off sharply and the contribution of the CMB to the total observed signal becomes negligibly small with increasing frequency. The finiteness of the spectrum enables us to apply the theory of analytical functions [14] to describe the CMB fluctuations, which implies the possibility of reconstructing functions over an entire region based on knowledge for part of the region or at some set of points [15]. This principle lies at the basis of the proposed algorithm for reconstructing the CMB fluctuations in regions contaminated by noise.

When modeling the background of point sources (PS), we assume that they are randomly distributed over the sky in accordance with a Poisson law. The instrumental pixel noise is "white" Gaussian noise with zero mean. The measured CMB map can be represented by the model

$$CMB_{meas} = (CMB + PS) * BEAM + N,$$

where N represents noise, $BEAM$ the antenna beam, and $*$ a linear convolution.

The required data processing consists of solving an equation of the form of the convolution (3) for the CMB signal. The sequence of operations we have used during the reconstruction of the CMB is the following:

- (1) The noise is filtered using a Wiener filter or a modification of such a filter that does not distort the form of the angular power spectrum [16]; this yields an estimate of the signal $(CMB + PS) * BEAM$.
- (2) This quantity is deconvolved from the antenna beam using a regularized inversion filter [17,18], in order to derive an estimate of the signal $CMB + PS$.
- (3) Positions contaminated by point sources are localized using a median filter (see Section 4).
- (4) The CMB signal is reconstructed at the contaminated positions using the method proposed in the following section.

3. Reconstruction of the CMB fluctuations at contaminated positions

The proposed method for reconstructing the fluctuations in holes in a CMB map is a modification of the algorithm of Fienup [19], intended for the reconstruction of an image of an object with a finite carrier from the amplitude of its Fourier spectra (the phase problem). In our modified version, we determine the limitation on the spectral region using information about the finiteness and circular symmetry (2) of the amplitude spectrum, rather than the amplitude spectrum itself, and determine the limitation on the spatial region using the known map values.

The algorithm is iterative, and consists of the following sequence of operations.

- (1) An initial approximation for the map is made. We recommend the use of the initial map with zero brightness in specified locations (holes) as a first approximation.
- (2) The Fourier transform of the initial approximation is calculated, bringing about a transformation to the spatial-frequency domain.
- (3) A condition for the spatial limitation of the fluctuation spectra, which comes about because the values of the Fourier components derived in the first step with numbers $\ell > \ell_D$ have been set equal to zero, is imposed. An additional constraint on the region of spatial frequencies that appreciably speeds up the convergence of the algorithm is the circular symmetry of the power spectrum of the fluctuations. To satisfy this last condition, the form of the Fourier components is modified so that the power spectrum has a form that is consistent with relation (2). For each value of ℓ , the squared amplitudes of the spectra measurements are averaged in radius over a length ℓ , after which these measurements are replaced by values with the derived squared amplitudes, retaining the phases.

(4)The inverse Fourier transform of the spectrum obtained in the previous step is taken, bringing about a transformation to the domain of the CMB map.

(5)The constraints on the spatial region of the map are imposed. The brightness values outside the holes in the map are replaced by the known values, while the values inside holes are not changed.

(6)The Fourier transform of the map obtained in step 5 is taken.

(7)Return to step 3 until the image obtained in step 5 ceases to change in accordance with a chosen convergence criterion.

Simulations show (see Section 6) that this modified Fienup algorithm leads rather rapidly to the correct solution. For this to be the case, the number of unknown values of the CMB fluctuations should be approximately half the number of knowns, if the discretization frequency of the map is twice its upper spatial frequency. If the discretization frequency is increased, a larger number of map values can be reconstructed, since the insufficient information in the spatial domain can be compensated by information in the spatial-frequency domain.

Obviously, a high computational speed for the method can be reached via the application of Fast Fourier Transform algorithms.

4. Method for localizing contaminated sections

The proposed method for localizing the point sources is based on applying a two-dimensional median filtration, which is a row-by-row column algorithm consisting of one-dimensional n -point median filtrations [20,21]. The number n is taken to be odd. If $n=2k+1$, one-dimensional median filtration consists of ascribing to the current value for the sequence the mean of the series that is obtained when the $(2k+1)$ -point sequence is placed in increasing order, with the first k values located to the left and the last k values to the right of the given value. As a result of applying this operation to a one-dimensional series, all impulsive noise is eliminated. In our case, such noise corresponds to point sources.

To obtain the series of removed sources itself, we subtract the output map from the input map. Obviously, since the nonlinear transform is applied to the useful signal together with the noise, the resulting difference map obtained from the subtraction contains additional noise, whose magnitude is appreciably lower than the desired impulsive noise when n is chosen correctly. Sections of the resulting difference map that are contaminated by point sources can be identified by applying a cutoff to the levels of this map. The lower the cutoff level, the more pixels are subject to distortion. In our case, the cutoff level was usually chosen to be fairly low (10% of the peak value of the CMB fluctuations), in order to avoid missing genuine places that are contaminated by weak radio sources. On the other hand, the number of identified points must not be too large to ensure convergence of the subsequent reconstruction of the CMB-fluctuation values at these points.

In contrast to numerous methods that have been proposed in the literature (wavelet analysis [7,8], clean and maximum-entropy deconvolution [10], optimal linear filtration [9], etc.), the proposed method for localizing the point sources is distinguished by its simplicity and the speed with which it operates, which is important in the reduction of large datasets obtained over the entire celestial sphere.

Obviously, this method will operate more reliably the brighter the point sources and the fewer the number of point sources in the map being analyzed. We have investigated the method for the cases of both relatively bright (resolved) and relatively weak (unresolved) radio sources, whose number grows with the resolution of the instrument used [11].

5. Filtration of the instrumental noise and deconvolution

We used a modification of a Wiener filter that preserves the form of the angular power spectrum [16], called a power filter, for the filtration of the instrumental white noise. A power filter can be applied in two ways. In the first, the dispersion of the instrumental (pixel) white noise is taken to be known *a priori*, whereas in the second, it is not. We present a preliminary estimate of this noise based on applying a median filter to the input signal. The dispersion of the instrumental noise was determined by subtracting the output signal of the median filter from the measured map, CMB_{meas} . We present here results obtained

using the second approach. Reconstruction of the map taking into account the antenna beam was carried out using a regularized inversion filter containing a regularizing parameter that depends on the level of the residuals after filtration of the noise [17, 18].

6. Simulation results

Here we consider a model for the observed CMB map that corresponds to expression (3), which includes the effect of the limited resolution of the system and the input pixel noise. In the following set of experiments, the map size was 128x128 pixels, the pixel size was 3.51', and the full width at half maximum of the antenna beam 10.53'. For an antenna with a diameter of 1m, this resolution corresponds to an observing frequency of 98 GHz. These system parameters are close to those for one of the channels of the PLANCK mission [11].

We carried out two experiments based on these parameters, without instrumental pixel noise and with significant level of white pixel noise. The results are shown in Fig.1, where the numbers (columns) indicate the experiment number and the letters (rows) indicate the physical meaning of the map: (a) measured CMB map satisfying (3); (b) reconstructed CMB map with holes in contaminated pixels, including the effect of the antenna beam; (c) residual map, equal to the difference of the given CMB map and the reconstructed map shown in row "b"; (d) reconstructed CMB map obtained using our method to interpolate functions in the holes; (e) residual map, equal to the difference of the given CMB map and the reconstructed map shown in row "d". Obviously, each column of maps presented in the figure corresponds to a particular value for the pixel-noise level.

The results presented in Fig.1 were obtained using pixel-noise filtration based on applying a median filter to derive a preliminary estimate of the noise dispersion (see Section 5). The effect of the antenna beam in the reconstructed map was included using a regularized inversion filter.

Analysis of the results presented in Fig.1 shows that the main advantage of the proposed method for removing point sources is that the residual maps shown in row "e" are free from non-Gaussian features, and resemble residual Gaussian noise, whose amplitude depends on the level of the input Gaussian noise. As we can see in map row "c", the traditional method, which simply excludes contaminated pixels from the CMB map, does not fully eliminate the contribution of the point sources, which are manifest as small non-Gaussian features determined by the magnitude of the CMB fluctuations in the contaminated locations. In addition, the sharp edges of the cutout sections lead to high-frequency noise that extends far beyond the frequency edge, which corresponds in our case to $\ell_{\max} = 1536$.

We illustrate the effect of applying our method in pure form in Fig.2, which shows (a) the angular power spectra of the convolved map $CMB * BEAM$ without (solid curve) and with holes, whose size is determined by the size of the antenna beam (dashed curve), and (b) the same spectra for the input CMB map. We can see from the curves that the presence of holes in the maps leads to errors of about 13% in the angular power spectrum of the CMB for multipoles $\ell = 200 - 300$ and errors of about 30% for multipoles $\ell = 450 - 550$. The relative errors are even higher for higher ℓ values.

Note that, in spite of the fact that the radio sources considered are point sources (i.e., they are essentially δ functions), we have taken the contaminated locations to include a region surrounding the source coordinates, determined by the characteristics of the inverse filter, which does not provide an ideal reconstruction of a δ function. The presence of even a small amount of input noise requires regularization of the algorithm, which leads to a solution with finite resolution. To ensure reliable exclusion of all contaminated pixels, we have assumed that the region distorted by each point source occupies an area equal to the area of the base of the antenna beam, which exceeds somewhat the total area of the excluded CMB measurements. However, this is not a serious concern for our method, since we reconstruct the intrinsic CMB values in the resulting holes.

Obviously, the maximum effect from reconstructing the CMB signal in contaminated locations is achieved in the absence of pixel noise. The quantitative effect is lowered as the pixel-noise level is increased, although the qualitative effect of a complete elimination of non-Gaussian features from the CMB maps is maintained even in the presence of fairly high noise levels.

Figures 3a and 3b present the angular power spectra of the signals corresponding to experiments 1 and 2, respectively. The solid curve shows the angular power spectrum of the initial CMB map, the dotted curve the measured CMB map, which satisfies (3), the dashed curve the angular spectrum of the reconstructed CMB map with the pixels distorted by point sources zeroed, and the dash-dotted curve the CMB map with reconstructed values for the fluctuations at the distorted locations.

We can see from Fig.3a that it is possible to reconstruct the CMB map nearly perfectly in the absence of pixel noise. The error in the reconstruction is determined only by the error in the inversion filtration. The effect of applying our method in this case is close to that shown in Fig.2..

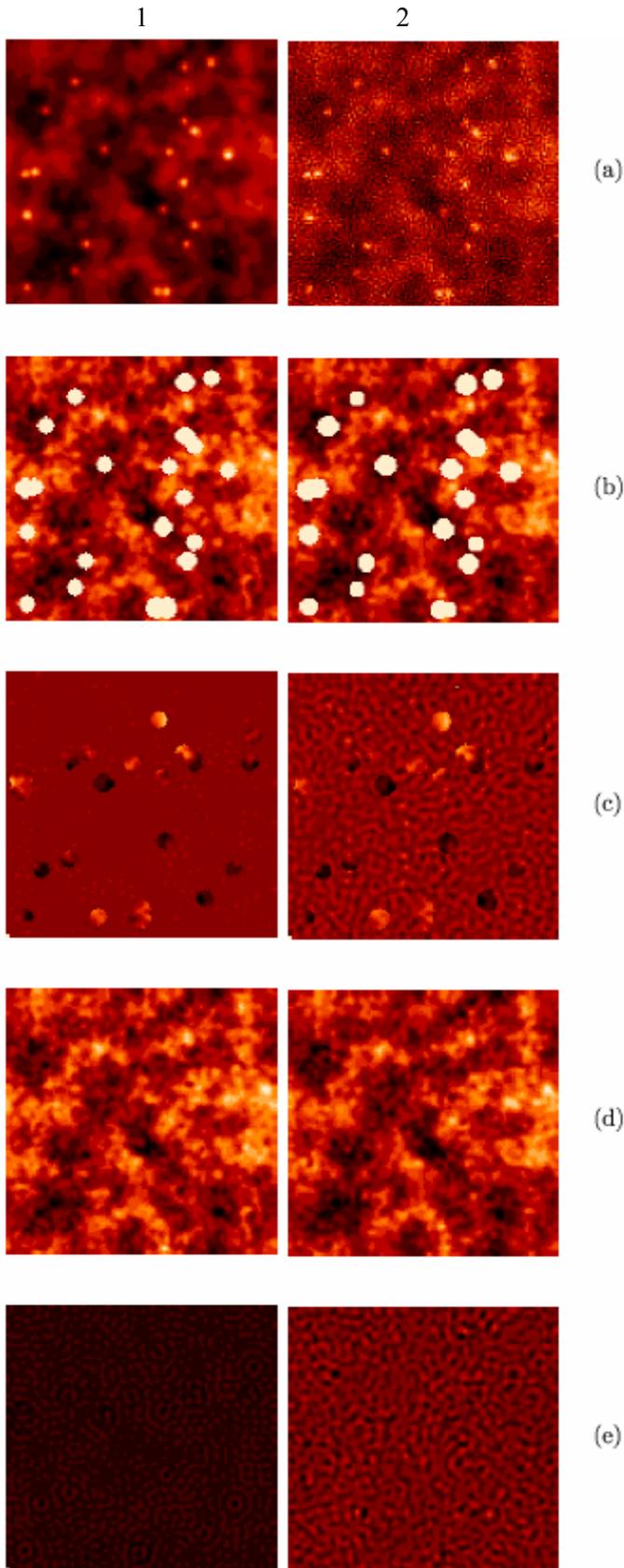


Fig.1

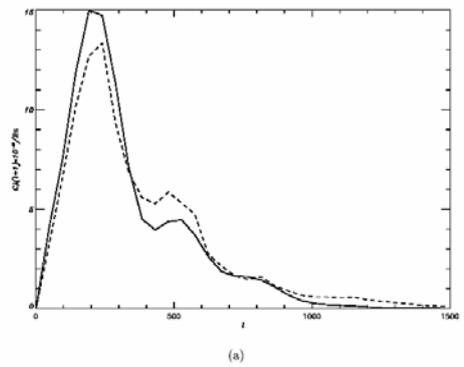
(a)

(b)

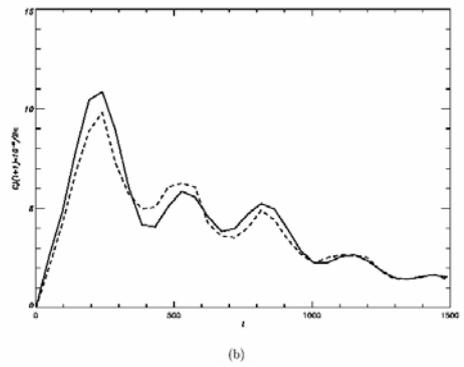
(c)

(d)

(e)

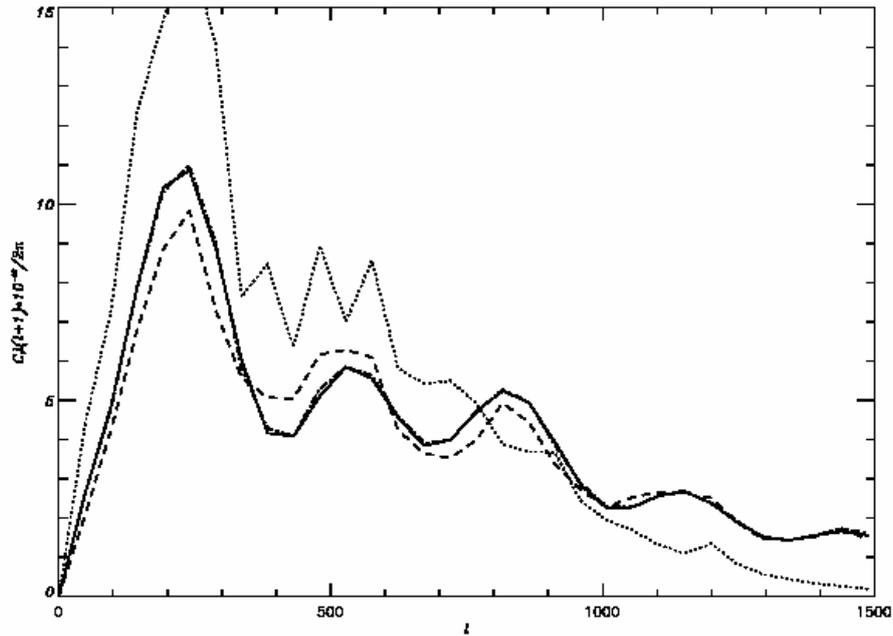


(a)

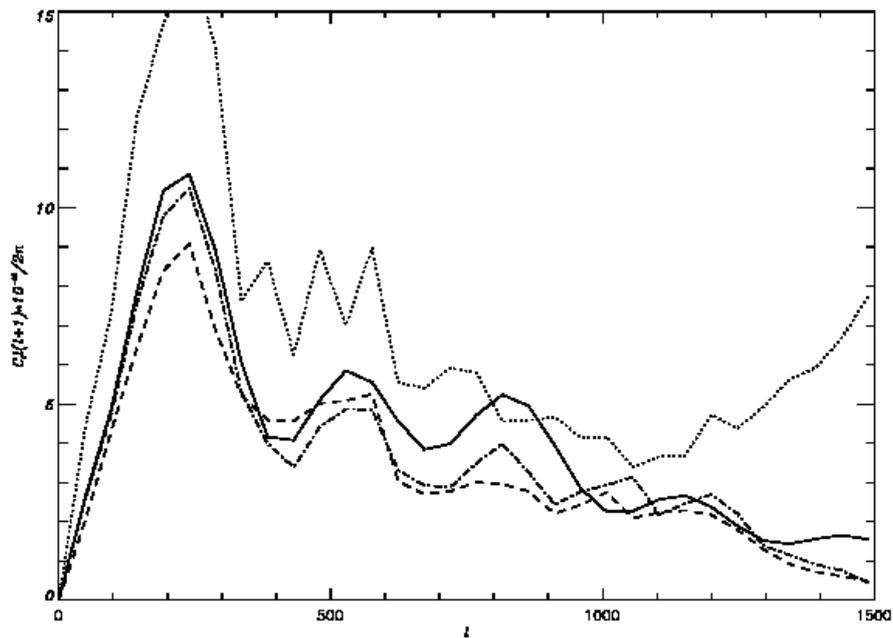


(b)

Fig.2



(a)



(b)

Fig.3

Figure 3b shows that a high level of pixel noise leads to a substantial loss in the accuracy of the reconstructed CMB angular power spectrum, for both a simple exclusion of contaminated sections and the reconstruction of the CMB signal in holes. However, the accuracy is nonetheless higher in the latter case - quite appreciably for multipoles $\ell = 24 - 500$. The dispersion of the residual-noise map for the specified interval of the input pixel noise is 25-27% lower due to the full removal of point sources from the CMB maps. Although the simulations show that the effect of applying our method decreases with growth in the pixel-noise level in a natural way, it nonetheless remains fairly high for a system with noise

characteristics similar to those planned for the PLANCK mission [11].

7. Conclusion

We have investigated the fundamental possibility of improving the accuracy of estimates of the angular power spectrum of the CMB by reconstructing the CMB anisotropy in regions contaminated by point radio sources and other high-multipole noise. A series of examples have been used to demonstrate the effect of applying this approach, compared to the currently standard strategy of simply excluding contaminated sections of CMB maps.

We have shown that, in the absence of instrumental noise, the proposed method ensures the removal of point sources and other non-Gaussian multipole noise with very high accuracy, and is fairly stable to variations in the noise level in the input data.

The main advantage of the method is that it is able to fully remove non-Gaussian features from CMB maps, while the simple exclusion of contaminated locations does not achieve this goal. The full removal of residual point sources from the maps makes it possible to appreciably lower the dispersion of the residual noise, even in the presence of a relatively high level of input pixel noise. Our simulations show that the effect of applying our method decreases in a natural way as the input noise level grows, but nonetheless remains fairly high for noise characteristics corresponding to those planned for the PLANCK mission.

We have also shown the fundamental possibility of enhancing the accuracy of estimates of the angular power spectrum of the CMB by reconstructing the CMB anisotropy in the zone of the Galaxy where the strongest level of background contamination is observed.

References

- 1.E. Gawiser and J. Silk, Phys. Rep. 333, 245 (2000).
- 2.M. Tegmark and G. Efstathiou, Mon. Not. R. Astron. Soc. 281, 1297 (1996).
- 3.M. Tegmark, A.de Oliveira-Costa, and A. Hamilton, Phys. Rev. D 68,123523 (2003).
- 4.V. Stolyarov, M.P. Hobson, M.A.J. Ashdown, *et al.*, Mon. Not. R. Astron. Soc. 336, 97 (2002).
- 5.M. Tegmark and A.de Oliveira-Costa, Astrophys. J. 500, L83 (1998).
- 6.M.P. Hobson, R.B. Barreiro, L. Toffolatti, *et al.*, Mon. Not. R. Astron. Soc. 306, 232 (1999).
- 7.J.L. Sanz,R.B. Barreiro,L. Cayon, *et al.*, Astron. Astrophys. 140, 99 (1999).
- 8.P. Vielva, E.. Martinez-Gonzalez, L. Cayon, *et al.*, Mon. Not. R. Astron. Soc. 326, 181 (2001).
- 9.R. Vio, L. Tenorio, and W. Wamsteker, Astron.Astrophys. 391, 789 (2002).
- 10.A.T. Bajkova, Izv. Vyssh. Uchebn. Zaved., Radiofiz. 45, 909 (2002); Radiophys. Quantum Electron. 45, 835 (2002).
- 11.M. Bersanelli *et al.*, *COBRAS/SAMBA, ESA Report D/SCI* (1996).
- 12.J.R. Bond and G. Efstathiou, Mon. Not. R. Astron. Soc. 281, 655 (1987).
- 13.J. Silk, Nature 215, 1155 (1967).
- 14.Ya.I. Khurgin and V.P. Yakovlev, *Finite Functions in Physics and Technology* (Nauka, Moscow, 1971) [in Russian].
- 15.G.I. Vasilenko and A.M. Taratorin, *Image Reconstruction* (Radio i Svyaz', Moscow, 1986) (in Russian)..
- 16.K.M. Gorski, astro-ph/9701191 (1997).
- 17.A.N. Tikhonov and V.Ya.Arsenin, *Solutions of Ill-Posed Problems* (Halsted, New York, 1977; Nauka, Moscow, 1986).
- 18.A.V. Goncharkii, A.M. Cherepachshuk, and A.G.Yagola, *Numerical Method for Solving Inverse Problems in Astrophysics* (Nauka, Moscow, 1978) (in Russian)..
- 19.J.R. Fienup, Opt. Lett. 3, 27 (1978).
- 20.B. Yustusson, in *Two-Dimensional Digital Signal Processing II.Transforms and Median Filters*, Ed. by T.S.Huang (Springer-Verlag, New York, 1984; Radio i Svyaz', Moscow, 1984), p. 156.
- 21.S.G.Tyan, in *Two-Dimensional Digital Signal Processing II. Transforms and Median Filters*, Ed. by T. S. Huang (Springer-Verlag, New York, 1984; Radio i Svyaz', Moscow, 1984), p. 191.