

Residual foreground contamination in the WMAP data

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 J. Phys.: Conf. Ser. 484 012028

(<http://iopscience.iop.org/1742-6596/484/1/012028>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 14.139.159.162

This content was downloaded on 18/03/2014 at 09:33

Please note that [terms and conditions apply](#).

Residual foreground contamination in the WMAP data

P Chingangbam¹ and C Park²

¹ Indian Institute of Astrophysics, Koramangala II Block, Bangalore 560034, India

² Korea Institute for Advanced Study, Hoegiro 85, Dongdaemun-Gu, Seoul 130-722, Korea

E-mail: prava@iiap.res.in, cbp@kias.re.kr

Abstract. We have studied whether there is any residual foreground contamination in the foreground reduced WMAP-7 data for the differential assemblies (DAs) Q, V and W. We have calculated the correlation between the foreground map, from which long wavelength correlations have been subtracted, and the foreground reduced map for each DA. We have found positive correlations for *all* the channels. The statistical significance of the resulting values has been tested by comparing with correlations between the cleaned CMB maps and 1000 simulated Gaussian maps to which instrumental effects have been added. We have found high statistical significance of the observed correlations, implying the presence of residual contamination in the cleaned data, and found that, for Q and V channels, a large fraction of the contamination comes from pixels, where the foreground maps have positive values larger than three times its rms value, which shows the presence of unresolved point sources that contribute significantly to the contamination.

1. Introduction

The experimentally measured CMB temperature fluctuations is composed of the true CMB signal and foreground contamination, coming from astrophysical sources that emit photons in the frequency ranges spanned by the experiments. The major part of the contamination comes from emissions in our galaxy, and a small fraction comes from point sources such as radio galaxies and dusty star-forming galaxies. It is necessary to know the contaminating component in order to get the true signal, which can be used to extract cosmological information [1]. The galactic foreground is usually estimated from a combination of theoretical modelling and observations, and then subtracted from the measured data [2]. Further, sky regions where there are possibility of residual contamination, such as the directions along the plane of our galaxy, and the locations of localized sources that have been identified are masked. The correct removal of the foreground contamination is extremely crucial, because any residual contamination will bias the extraction of cosmological information from CMB data.

In this paper, we have investigated whether there is small residual foreground contamination in the cleaned WMAP data by using a method that is simple, but has proved to be quite effective. Our method is based on calculating correlations between the foreground field, which has been processed so as to remove long wavelength correlation of the galaxy emissions, and the cleaned CMB data. Our basic premise is that if there is no residual contamination in the cleaned data, we should obtain no correlation. However, we have found statistically significant positive



correlation for WMAP-7 data for the Q, V and W differential assemblies (DAs)¹. Further, we have found that a big fraction (as big as 30% for Q channel) comes from regions where the foreground map has large positive values, which indicates unresolved point sources. This is not surprising, in light of the fact that the PLANCK early data release has identified point sources of the order of 15000 [3]. These results imply the presence of residual contamination in the cleaned data. This work has been done with the motivation to quantify the contaminant fraction, and understand the nature of non-Gaussian deviations they introduce into the CMB. The earlier studies on point sources contamination on the CMB have been included [4, 5, 6, 7].

This paper is organized as follows: In Section 2, we have outlined our procedure for calculating the correlations between the foreground and cleaned CMB maps, and then presented our correlation results and their statistical significance. We have given the concluding remarks in Section 3.

2. Quantifying residual foreground contamination

Let f_1 and f_2 be any two random fluctuation fields having zero mean values, defined on the surface of a two dimensional sphere. Let their rms values be given by σ_1 and σ_2 respectively. Re-scaling them as $\nu_1(i) \equiv f_1(i)/\sigma_1$ and $\nu_2(i) \equiv f_2(i)/\sigma_2$, where i denotes the pixel number, we can define a correlation parameter r as:

$$r \equiv \langle \nu_1(i) \nu_2(i) \rangle, \quad (1)$$

where the bracket denotes average over all pixels. We expect r to be zero if the two fields are uncorrelated, and non-zero otherwise. (In practice, we will always get a small, but non-zero value of r even though the two fields are known to be totally uncorrelated, due to the finite number of pixels.)

The observed WMAP data f^{obs} is a sum of the true CMB signal and foreground contamination. The foreground component is estimated using a combination of galaxy observation and theoretical modelling [2]. We have called this field the ‘apparent’ foreground field, denoted by f^{appfg} , keeping in mind that there may be small error in its estimation. This field is then subtracted pixel by pixel from f^{obs} to leave behind the ‘cleaned’ CMB signal, which we have denoted by f^{cleaned} . By definition, f^{cleaned} has zero mean. If f^{appfg} has been correctly estimated, then we expect it to have negligibly small correlation with f^{cleaned} , since they come from totally different physical processes. However, if the estimation is on the right track but not fully correct, then we should expect some residual contamination in the signal field. This should show up as non-zero correlation between f^{cleaned} and f^{appfg} . In what follows, we have described our results for the correlations.

2.1. Peak field

Our analysis is done using the 7 years data from the eight DAs of WMAP [8], namely, Q₁, Q₂, V₁, V₂, W₁, W₂, W₃, and W₄. For each DA, we have obtained f^{appfg} as:

$$f^{\text{appfg}} = f^{\text{obs}} - f^{\text{cleaned}}. \quad (2)$$

Since, we want to remove the long distance correlations of the emissions from our galaxy, we have first defined what we call the *peak* field as:

$$f^{\text{peak}} \equiv \left(f^{\text{appfg},\theta_s} - f^{\text{appfg},3\theta_s} \right) - \langle f^{\text{appfg},\theta_s} - f^{\text{appfg},3\theta_s} \rangle, \quad (3)$$

where θ_s and $3\theta_s$ are FWHM values, at which we have smoothened the field. By definition, f^{peak} has zero mean. Figure 1 shows the peak field for Q1 channel.

¹ We have done similar analysis using WMAP-5 data and obtained similar results.

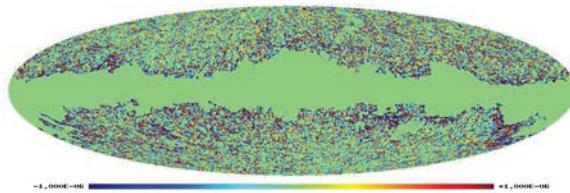


Figure 1. Peak field for Q_1 DA.

	Q_1	Q_2	V_1	V_2	W_1	W_2	W_3	W_4
r_c for $s_{\text{mask}} = 0.89$	0.026	0.025	0.020	0.019	0.010	0.008	0.009	0.006
	0.018	0.018	0.017	0.016	0.009	0.008	0.007	0.007
r_c for $s_{\text{mask}} = 0.91$	0.025	0.025	0.020	0.019	0.009	0.007	0.008	0.006
	0.018	0.017	0.016	0.016	0.008	0.008	0.006	0.006
r_c for $s_{\text{mask}} = 0.93$	0.025	0.024	0.019	0.018	0.008	0.007	0.007	0.004
	0.017	0.017	0.016	0.015	0.007	0.007	0.005	0.006

Table 1. r_c values for the eight DAs are shown. For each DA and s_{mask} , the upper value gives r_c calculated using all unmasked pixels, while the lower value has been calculated after excluding pixels having $\nu^{\text{peak}} > 3$ also. The sky fractions for the three s_{mask} values from top to bottom are roughly 62%, 60% and 58%.

We have used the KQ75y7 galaxy mask and point sources given in the 7 years data release of WMAP. Rather than using the galaxy plus point sources mask directly, we have first smoothed it with $\text{FWHM} = 3\theta_s$. The pixels near the mask boundaries have been obtained values between zero and one. To mask f^{cleaned} and f^{peak} , we have chosen some threshold value, s_{mask} (between zero and one), and throw away pixels of f^{cleaned} and f^{peak} , if the corresponding smoothed mask pixels have values less than s_{mask} . This method of masking makes it easy to control how far away we stay from the mask boundary of f^{cleaned} and f^{peak} in order to minimize the boundary effect. Staying 2σ (where σ is the rms of the field under consideration) away from the boundary, corresponds to choosing pixels with $s_{\text{mask}} > 0.89$. As we have chosen larger values of s_{mask} , we stay further away from the boundaries, and the sky fraction decreases.

2.2. Correlation between peak and cleaned CMB fields

Let us denote:

$$\nu^{\text{cleaned}}(i) \equiv \frac{f^{\text{cleaned}}(i)}{\sigma^{\text{cleaned}}}, \quad \text{and} \quad \nu^{\text{peak}}(i) \equiv \frac{f^{\text{peak}}(i)}{\sigma^{\text{peak}}}, \quad (4)$$

and define:

$$r_c \equiv \langle \nu^{\text{cleaned}} \nu^{\text{peak}} \rangle_{\theta_s}, \quad (5)$$

where the suffix θ_s is to remind us that we do the calculation for a choice of FWHM at which f^{cleaned} has also been smoothed. σ^{cleaned} and σ^{peak} are of the orders of 10^{-5} and 10^{-7} respectively.

Table 1 summarizes the main results for r_c . We have used $\theta_s = 35'$. Two values of r_c are shown for each DA and s_{mask} . The upper value is the case where r_c is calculated using all unmasked pixels, while the lower value is the case where pixels with $\nu^{\text{peak}} > 3$ have also been excluded. The first observation we have made is that *all* r_c values are positive. For Q channels,

	Q_1	Q_2	V_1	V_2	W_1	W_2	W_3	W_4
N for $s_{\text{mask}} = 0.89$	0	0	5	5	121	156	139	224
	14	16	107	113	262	270	311	305
N for $s_{\text{mask}} = 0.91$	0	0	6	10	143	186	162	247
	15	16	110	118	276	284	332	320
N for $s_{\text{mask}} = 0.93$	0	0	11	15	169	218	196	291
	16	21	116	127	282	292	351	342

N_0 for $s_{\text{mask}} = 0.89$	0
	5
N_0 for $s_{\text{mask}} = 0.91$	0
	6
N_0 for $s_{\text{mask}} = 0.93$	0
	9

Table 2. *Left:* Number of maps N , having $r_g > r_c$ for individual DAs, out of 1000 Gaussian maps. As in Table 1, upper values are for all unmasked pixels included, while lower values are for the case when pixels with $\nu^{\text{peak}} > 3$ have also been excluded. *Right:* Number of Gaussian maps N_0 , having $r_g > r_c$ simultaneously for all DAs.

we get considerably larger correlation when we keep all unmasked pixels, larger by about 30%. This indicates that there is non-trivial correlation arising from the pixels with $\nu^{\text{peak}} > 3$. For V channels, the difference is about 20%, while W channels do not seem to be affected. We have repeated these calculations for the case where ν^{cleaned} are also excluded, but not found any significant effect. The sky fractions for the three s_{mask} values are roughly, 62%, 60% and 58% respectively. For Q and V channels, as we stay further away from the mask boundaries, there is small but systematic decrease of r_c .

2.3. Statistical significance of r_c values

We have investigated how likely it is to get the observed r_c values given above by comparing with correlations between the peak field and Gaussian CMB simulations. For this purpose, we have simulated 1000 Gaussian CMB maps with WMAP-7 parameter values, and added pixel window effect, beam smearing and WMAP-7 noise characteristics. Next, we have smoothed by FWHM 35' and mask in exactly the same way as we did when calculating r_c , and calculated the correlation with the peak field. We have denoted the correlation value by r_g . The Gaussian fields are uncorrelated with the signal field, and we should get small value of r_g . This exercise has told us what is the typical value of 'small' r_c that we can approximate it to be zero for the number of pixels under consideration, and how likely are our observed r_c values to occur by random fluctuation and not due to a true correlation.

We have counted, out of the thousand r_g values, how many are greater than r_c . The results are shown in Table 2. The left table shows the number N , of Gaussian maps having $r_g > r_c$ for each individual DA, for the three s_{mask} values used earlier for calculating r_c , and including/excluding the pixels having $\nu^{\text{peak}} > 3$. When all unmasked pixels are included, we get $N = 0$ for all s_{mask} values for Q channel, for V channels, N lies between 5 and 15, while for W channels, N lies between 121 and 291. These numbers imply that the r_c values for Q and V are statistically significant, whereas, the values for W channels have much lower significance. When pixels with $\nu^{\text{peak}} > 3$ are also excluded, we have found a reduction of N for all the channels. The table on the right side of Table 2 shows the number N_0 , of Gaussian maps having $r_g > r_c$ *simultaneously for all DAs*. These values are again significant. Therefore, we have concluded that the cleaned WMAP data, particularly Q and V channels, contain small but statistically significant amount of residual foreground contamination.

3. Conclusion

We have shown that the cleaned WMAP data contains small but statistically significant amount of residual foreground contamination. The level of contamination is quantified by the level of correlation obtained between the cleaned map and the peak field. For Q and V DAs, we have found that a big fraction of the contamination could be coming from unresolved point sources. Our result has important implications for the extraction of cosmological data using the cleaned WMAP data. For example, it is known that constraints on the local primordial non-Gaussianity parameter f_{NL} using Fourier space method such as the bispectrum, and those use pixel space methods, like the Minkowski Functionals (MFs) do not agree with each other. The bispectrum gives positive best fit value of f_{NL} [9], while the MFs give negative values [10]. Our analysis in this work was motivated by an attempt to understand whether this conflict is due to residual foreground contamination. Our next goal is to understand the effect of the residual contamination on the measurement of the MFs, and we are currently pursuing this line of investigation.

Acknowledgement

We thank Korea Institute for Advanced Study for providing computing resources (KIAS Center for Advanced Computation Linux Cluster System QUEST), where a part of the computation was carried out. We acknowledge the use of the HEALPIX package [11, 12], and the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science.

References

- [1] Komatsu E *et al* 2011 *Ap. J. Supp.* **192** 18
- [2] Gold B *et al* 2011 *Ap. J. Supp.* **192** 15
- [3] Planck early results VII. *The Early Release Compact Source Catalogue* 2001 *A. & A.* **536** A7
- [4] Argueso F, Gonzalez-Nuevo J and Toffolatti L 2003 *Ap. J.* **598** 86
- [5] Boughn S P and Partridge R B 2008 *PASP* **120** No. 865
- [6] Babich D and Pierpaoli E 2008 *Phys. Rev. D* **77** 123011
- [7] Lacasa F, Aghanim N, Kunz M and Frommert M [arXiv:1107.2251[astro-ph.CO]]
- [8] <http://lambda.gsfc.nasa.gov/>
- [9] Komatsu E *et al* (WMAP Collaboration) 2009 *Ap. J. Supp.* **180** 330
- [10] Hikage C, Matsubara T, Coles P, Liguori M, Hansen F K and Matarrese S 2008 *MNRAS* **389** 1439
- [11] Gorski K M, Hivon E, Banday A J, Wandelt B B, Hansen F K, Reinecke M and Bartelmann M 2005 *Ap. J.* **622** 759 [arXiv:astro-ph/0409513]
- [12] <http://healpix.jpl.nasa.gov>