

Application of the Kolmogorov-Smirnov test to CMB data: Is the universe really weakly random? (Research Note)

S. K. Næss

Institute of theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, 0315 Oslo, Norway
e-mail: sigurdkn@astro.uio.no

Received 25 May 2011 / Accepted 19 September 2011

ABSTRACT

A recent application of the Kolmogorov-Smirnov test to the WMAP 7 year *W*-band maps claims evidence that the CMB is “weakly random”, and that only 20% of the signal can be explained as a random Gaussian field. I here repeat this analysis, and in contrast to the original result find no evidence for deviation from the standard Λ CDM model. Instead, the results of the original analysis are consistent with not properly taking into account the correlations of the Λ CDM power spectrum.

Key words. cosmic background radiation

1. Introduction

In astronomical data analysis, it is often useful to be able to test whether a set of data points follows a given distribution or not. For example, many analysis techniques depend on instrument noise being Gaussian, and to avoid bias, one must check that this actually is the case. There are many different ways in which two distributions can differ, and correspondingly many different ways to test them for equality. The simplest ones, such as comparing the means or variances of the distributions, suffer from the problem that there are many ways in which distributions can differ that they cannot detect no matter how many samples are available. For example, samples from a uniform distribution can easily pass as Gaussian if one only considers the mean and variance.

The popular Kolmogorov-Smirnov test (K-S test) resolves this problem by considering the cumulative distribution functions (CDF) instead: construct the empirical CDF of the data points and find its maximum absolute difference K from the theoretical CDF. Due to the limited number of samples, the empirical CDF will be noisy, and K will therefore be a random variable with its own CDF, which in the limit where the number of samples goes to infinity is given by

$$P(x < K) = F_{KS} \left(\sqrt{N_{\text{obs}}} K \right) \quad (1)$$

with

$$F_{KS}(x) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 x^2}. \quad (2)$$

In contrast with the simplest tests, this test can detect any deviation in the distributions, but may require a large number of samples to do so, especially in the tails of the distribution.

Recently, a series of papers (Gurzadyan et al. 2011, 2010; Gurzadyan & Kocharyan 2008) has applied this test to WMAP’s cosmic microwave background (CMB) maps, resulting in the remarkable claim that the CMB is “weakly random”, with

only 20% of the CMB signal behaving as one would expect from a random Gaussian field. This result went on to be used in a much discussed series of papers (Gurzadyan & Penrose 2010a,b, 2011) claiming a strong detection of concentric low-variance circles in the CMB, which was taken as evidence for conformal cyclic cosmology. Other groups failed to significantly detect the circles (Wehus & Eriksen 2010; Moss et al. 2011; Hajian 2010). The difference in significance was due to different CMB models: Wehus & Eriksen (2010); Moss et al. (2011); Hajian (2010) used realizations of the best-fit Λ CDM power spectrum, while Gurzadyan & Penrose (2010b) used a “weakly random” CMB model.

Both in order to resolve this issue, and because a weakly random universe would be a strong blow against the Λ CDM model in its own right, it is important to test this result.

2. Method

Before applying the K-S test, one must be aware of its limited area of validity: Eq. (1) requires an infinite number of independently identically distributed samples, while CMB maps actually consist of a limited number of correlated samples. However, both the correlations and number of samples can be compensated for, as we shall see.

2.1. Application of the K-S test to correlated data

Though the K-S test is not immediately applicable to a correlated data set, it is possible to perform an equivalent test on a transformed set of samples. The question we are trying to answer with the K-S test is “Do the samples follow the theoretical distribution?”. The truth or falseness of this is preserved if we apply the same transformation to both the samples and the distribution we test them against, and to be able to use the K-S test, the logical transformation to use is a *whitening* transformation, which results in an independent, identical distribution for the samples.

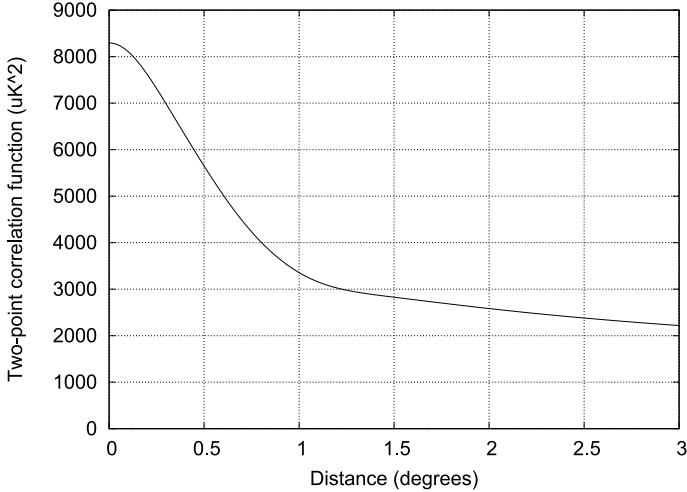


Fig. 1. ΛCDM two point correlation function after applying the WMAP W -band beam and the HEALPix (Górski et al. 2005) inside 512 pixel window.

With original samples \mathbf{d} with covariance matrix \mathbf{C} , the whitened (uncorrelated with unit variance) samples \mathbf{r} are given by:

$$\mathbf{r} = \mathbf{C}^{-\frac{1}{2}}\mathbf{d} \quad (3)$$

Thus, to test whether the data points $\mathbf{d} \leftarrow N(0, \mathbf{C})$, we can test the equivalent hypothesis $\mathbf{r} \leftarrow N(0, 1)$.

In the case of CMB maps, both the data itself and the noise is expected to be Gaussian, so the obvious theoretical distribution here is $N(0, \mathbf{S} + \mathbf{N})$, where the CMB signal covariance matrix \mathbf{S} is given by the two-point correlation function:

$$S_{ij} = \sum_{l=0}^{\infty} \sqrt{\frac{2l+1}{4\pi}} C_l B_l P_l(\cos(|\mathbf{p}_i - \mathbf{p}_j|)) \quad (4)$$

$P_l(x)$ are the Legendre polynomials normalized to $\frac{1}{2\pi}$, and \mathbf{p}_i and \mathbf{p}_j are the direction vectors for pixel i and j in the disk. C_l is the ΛCDM angular power spectrum, while B_l accounts for the beam and pixel window. \mathbf{N} is instrument dependent, but for the WMAP W -band CMB map we will use here, the noise is nearly diagonal, and given by the corresponding W -band rms map.

2.2. Application of the K-S test with few samples

The other problem we need to account for is our finite number of samples. In this case Eq. (1) is only approximate. For most uses of the test, this approximation is good enough, especially when employing analytical expressions for improving the quality of the approximation for low numbers of samples (von Mises 1964). For example, when performing a single test to accept or reject a test distribution, a bias of a few percent in the confidence with which the hypothesis is rejected is not important.

However, when making statistics for a large number of such test results, such a bias may make the results ambiguous. Given a set of experiments with a corresponding set of maximum deviations $\{K_i\}$, the corresponding probabilities $\{p_i = P(x < K_i)\}$ should be uniformly distributed if the samples actually follow the theoretical distribution, and a histogram of $\{p_i\}$ should therefore be flat. Deviations from this indicate that the theoretical distribution does not accurately describe the samples. However, the approximate Eq. (1) also introduces a small non-uniformity in $\{p_i\}$ even if the samples actually do follow the distribution. To

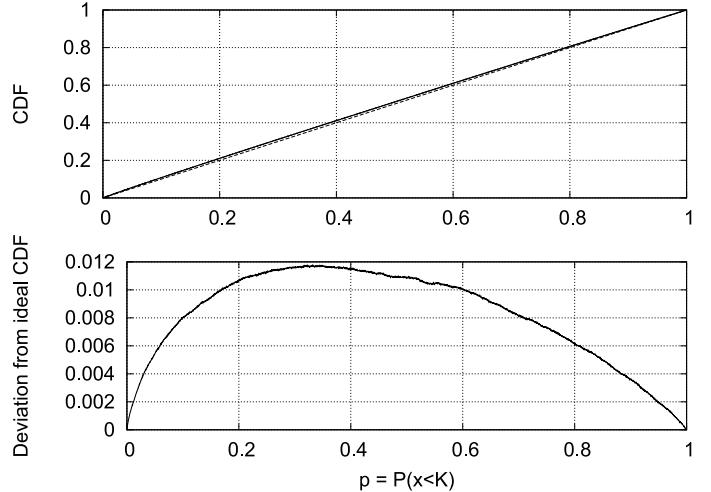


Fig. 2. When applying the K-S test to samples known to come from the correct distribution, the resulting values $\{p_i = P(x < K_i)\}$ should be uniformly distributed, but when working with a limited number of samples, the Kolmogorov distribution is only approximate, and the actual CDF of the results, $G(p)$, differs from the ideal $G_\infty(p) = p$. This is shown in the upper panel for the case of 540 samples per experiment, where $G(p)$ is the solid line and $G_\infty(p)$ is dashed. The lower panel shows the deviation between the two, which is of the order of 1% in this case (but larger with fewer samples).

avoid the ambiguity this causes, we will instead compute a numerical correction function mapping the approximate p to the true p' ¹.

To build up the correction, we simulate a large number² of experiments, each with the same number of samples as the actual data set, but drawn directly from the theoretical distribution. Thus, for these, $\{p_i\}$ should be uniform, with a CDF of $G_\infty(p) = p$. However, since Eq. (1) is inexact, for small numbers of samples, the actual CDF is $G(p) \neq G_\infty(p)$. The mapping between the approximate p and true p' is given by $G(p) = G_\infty(p') \Rightarrow p' = G_\infty^{-1}(G(p)) = G(p)$. Thus, for a limited number of samples

$$P(x < K) = G(F_{\text{KS}}(\sqrt{N_{\text{obs}}}K)). \quad (5)$$

Figure 2 illustrates the correction function for 5×10^6 simulations of 540 each. For this many samples, the correction is only of the order of 1%.

3. Does ΛCDM fail the K-S test?

With this in hand we can finally apply the K-S test on CMB data. Following Gurzadyan et al. (2011), we randomly pick 10 000 disks with a radius of 1.5 degrees from the WMAP 7 year W -band map (Jarosik et al. 2011), with the region within 30 degrees from the galactic equator excluded. Each disk contains on average 540 pixels, which are whitened using Eq. (3). A typical disk before and after the whitening operation can be seen in

¹ What we do here is essentially replacing the analytical Komolgorov distribution (Eq. (1)) with a numerical distribution. This could also be done without using the analytical distribution as a basis, at a small cost in clarity.

² The number necessary depends on the level of accuracy desired. The noise in the estimate of $G(p)$ propagates to the final results. To make this a subdominant noise contribution, the number of simulations should be at least as large as the number of actual experiments, preferably much higher.

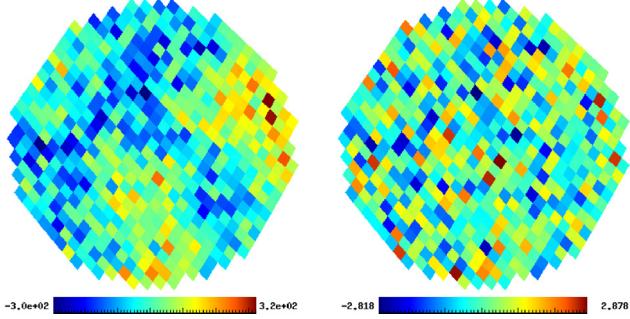


Fig. 3. A randomly selected disk before (left) and after (right) the whitening operation. The samples are strongly correlated and thus unsuitable for the K-S test before the transformation, but afterwards no correlations are visible and the variance is 1. Note that whitening the data does not mean that we are “forcing” the K-S test to pass. The whitened data will only end up matching $N(0, 1)$ after whitening if they followed our theoretical distribution $N(0, \mathbf{C})$ before.

Fig. 3. After whitening, the values should follow the distribution $N(0, 1)$ if our model is correct.

The histogram of resulting probabilities $\{p_i = P(x < K_i)\}$ from of applying Eq. (5) to the hypothesis $r \leftarrow N(0, 1)$ is shown in Fig. 4, together with the 68% and 95% intervals from 300 simulations. The data and simulations are consistent, and follow a uniform distribution as expected³: the CMB map is fully consistent with $\Lambda\text{CDM} + \text{WMAP}$ noise as far as the K-S test is concerned.

This is dramatically different from the curve found by Gurzadyan et al. (2011), which was strongly biased towards low values. Low values of $P(x < K)$ would mean that the empirical CDF of the samples matches the theoretical one *too well*, i.e. even better than samples drawn directly from the theoretical distribution.

What could cause Gurzadyan et al. to get results so different from ours? One way biasing $P(x < K)$ low is by basing the parameters of your test distribution on the values themselves. However, even without doing this, it is possible to get low values if the values used in the K-S test are *correlated*. This is also consistent with the presentation given by Gurzadyan et al. (2011) who apparently applied the K-S test directly to the raw samples \mathbf{d} , or equivalently, that they model the pixel values as coming from a 1-dimensional distribution. To check this, I repeated the analysis, this time using the theoretical distribution $d \leftarrow N(\mu, \sigma^2)$, where μ and σ^2 are the measured mean and variance of the samples in the disk. The result is also shown in Fig. 4. This time, the bias towards low values is clearly recreated.

It therefore seems likely that Gurzadyan et al.’s reported “weak randomness” is the result of not properly taking the CMB’s correlations into account. One is, of course, free to use whatever distribution one wants as the theoretical distribution in a K-S test, even a model where the CMB pixels are independently identically distributed, with no correlations at all. The problem lies in the interpretation of the test results. For Gurzadyan et al. (2011), the K-S test results are clearly not uniform, indicating that the chosen theoretical distribution has been disproven. However, Gurzadyan et al. then go on to create a set of simulations (linear combinations of 20% Gaussian and 80%

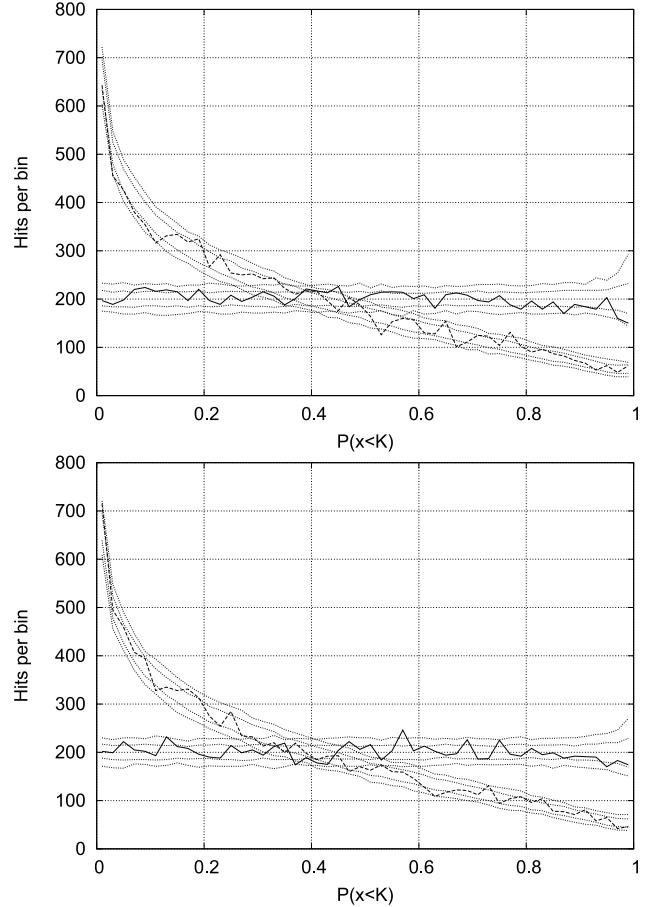


Fig. 4. Histogram of results of the K-S test. Each panel compares the results from properly taking the correlations into account (solid line) with those one gets from ignoring them (dashed line), together with 68% and 95% intervals (dotted lines) from simulations. *The upper panel* corresponds to using samples further than 30 degrees away from the galactic equator, while *the lower panel* instead uses the WMAP KQ85 analysis mask. In both cases, both the map and the simulations pass the K-S test when taking the correlations into account, while if the are ignored, the K-S test fails in the same way Gurzadyan et al. (2011) reported.

static signal) that fail the test in the same way as the WMAP map does. But having two sets of samples fail the K-S test the same way does not prove that they have the same properties. It simply means that the chosen test distribution was a poor choice.

4. Kolmogorov maps

While Gurzadyan et al.’s Kolmogorov statistics are biased by not taking the correlations into account, the approach of making sky maps of K-S test results introduced in Gurzadyan et al. (2009) is still an interesting way to search for regions of the sky that do not follow the expected distribution. Making an unbiased Kolmogorov map straightforwardly follows the procedure in Sect. 3, with the main difference being the selection of pixels. Instead of randomly selecting disks, we now systematically go through n_{side} 16 pixels, using the $1024 n_{\text{side}} 512$ subpixels inside each one as the samples. These are then tested against $N(0, \mathbf{C})$ by whitening them via Eq. (3) and then comparing the whitened samples to $N(0, 1)$.

The result is the n_{side} 16 map of $P(x < K)$ shown in Fig. 5. Regions that pass the test have a value uniformly

³ It should be noted that the histogram bins are not completely independent for two reasons: firstly, some disks are going to overlap, meaning that the same samples enter into several different K-S tests, and secondly, while our transformation has made the samples within each disk independent, the correlation between different disks is still present.

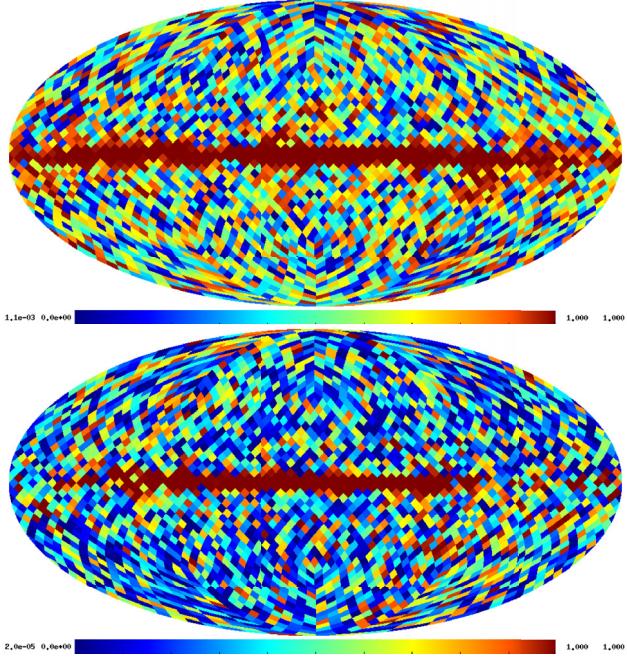


Fig. 5. nside 16 map of $P(x < K)$ based on pixels from the WMAP 7 year nside 512 W-band map. This time, the K-S test is performed on the 1024 nside 512 pixels inside each nside 16 pixel instead of on disks. The upper panel uses $N(0, S + N)$ as the theoretical distribution (by testing the whitened data against $N(0, 1)$), while the lower panel ignores the correlations, instead using $N(\mu, \sigma^2)$, where μ and σ are the measured mean and standard deviation of the samples. The former passes the test outside the galactic plane, while the latter fails everywhere, being biased low outside the galaxy.

distributed between 0 and 1, and we see that this applies to the CMB-dominated areas of the sky, while areas dominated by the galaxy fail the test as expected.

For comparison, Fig. 5 also includes the result of making the same map while ignoring correlations. In this case, the whole sky fails the test: The CMB-dominated areas are biased low, while the galaxy is biased high. This map is similar to the map in Gurzadyan et al. (2009), which is also too low outside the galaxy,

and too high inside, which is, again, consistent with Gurzadyan et al. applying the K-S test directly to the raw samples.

5. Summary

The Kolmogorov-Smirnov test is a useful and general way of testing whether a data set follows a given distribution or not. However, it only applies to independently identically distributed samples. The CMB is strongly correlated, and thus not immediately compatible with the test. However, this can be resolved by the application of a whitening transformation, replacing the hypothesis $d \leftarrow N(0, C)$ with the equivalent $C^{-\frac{1}{2}}d \leftarrow N(0, 1)$. With this, we find that the Λ CDM passes the K-S test. This is incompatible with the original analysis by Gurzadyan et al. (2011), which claimed detection of an unknown non-random component making up 80% of the CMB based on the CMB failing the K-S test there. It turns out that this analysis did not take the CMB correlations into account, which we confirm by producing the same failure of the K-S test when we skip the whitening step. When the correlations are handled properly, there is no need for a weakly random universe.

Acknowledgements. The author would like to thank Hans Kristian Eriksen for useful discussion and comments.

References

- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
- Gurzadyan, V. G., & Kocharyan, A. A. 2008, A&A, 492, L33
- Gurzadyan, V. G., & Penrose, R. 2010a [arXiv1011.3706]
- Gurzadyan, V. G., & Penrose, R. 2010b [arXiv1012.1486]
- Gurzadyan, V. G., & Penrose, R. 2011 [arXiv1104.5675]
- Gurzadyan, V. G., Allahverdyan, A. E., Ghahramanyan, T., et al. 2009, A&A, 497, 343
- Gurzadyan, V. G., Kashin, A. L., Khachatryan, H. G., et al. 2010, Europhys. Lett., 91, 19001
- Gurzadyan, V. G., Allahverdyan, A. E., Ghahramanyan, T., et al. 2011, A&A, 525, L7
- Hajian, A. 2010 [arXiv1012.1656]
- Jarosik, N., Bennett, C. L., Dunkley, J., et al. 2011, ApJS, 192, 14
- Moss, A., Scott, D., & Zibin, J. P. 2011, JCAP, 4, 33
- von Mises, R. 1964, Mathematical Theory of Probability and Statistics (New York: Academic Press)
- Wehus, I. K., & Eriksen, H. K. 2011, ApJ, 733, L29