Higher Criticism statistic: detecting and identifying non-Gaussianity in the WMAP first-year data

L. Cayón, J. Jin and A. Treaster

1Department of Physics, Purdue University, 525 Northwestern Avenue, West Lafayette, IN 47907-2036, USA
2Department of Statistics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067, USA

Accepted 2005 June 8. Received 2005 January 26

ABSTRACT

Higher criticism is a recently developed statistic for the detection of non-Gaussianity. It was proposed by Donoho & Jin, who showed it to be effective at resolving a very subtle testing problem: whether \( n \) normal means are all zero versus the alternative that a small fraction is non-zero. Higher Criticism is also useful in the detection of the non-Gaussian convolution component of cosmic strings in the cosmic microwave background (CMB) (see Jin et al.). In this paper, we study how well the anisotropies of the CMB fit with the homogeneous and isotropic Gaussian distribution predicted by the standard inflationary model. We find that Higher Criticism is useful for two purposes. First, Higher Criticism has competitive detection power, and non-Gaussianity is detected at the level of 99 per cent in the first-year Wilkinson Microwave Anisotropy Probe (WMAP) data. We generated 5000 Monte Carlo Gaussian simulations of the CMB maps. By applying the Higher Criticism statistic to all of these maps in wavelet space, we constructed confidence regions of Higher Criticism at levels of 68, 95 and 99 per cent. We find that the Higher Criticism value of the WMAP data is outside the 99 per cent confidence region at a wavelet scale of 5° (99.46 per cent of Higher Criticism values based on simulated maps are below the values for WMAP). Secondly, Higher Criticism offers a way to locate a small portion of data that accounts for the non-Gaussianity. This property is not immediately available for other statistical tests such as the widely used excess kurtosis test. Using Higher Criticism, we have successfully identified a ring of pixels centred at \( l \approx 209°, b \approx -57° \), which seems to account for the observed detection of non-Gaussianity at the wavelet scale of 5°. After removal of the ring from the WMAP data set, no more prominent deviation from Gaussianity was found. Note that the detection was achieved in wavelet space first. Secondly, it is always possible that a fraction of pixels within the ring might deviate from Gaussianity even if they do not appear to be above the 99 per cent confidence level in wavelet space. The location of the ring coincides with the cold spot detected by Vielva et al. and Cruz et al.

Key words: methods: statistical – cosmic microwave background – cosmology: miscellaneous.

1 INTRODUCTION

The standard inflationary model solves the horizon, the flatness, and the monopole problems, and provides a framework for the formation of structure in the Universe (Guth 1981; Guth & Pi 1982). Regarding the latter, the standard inflationary model predicts the existence of quantum density fluctuations that were amplified during the inflationary period and that grew, through gravitational instabilities, into the galaxies and clusters that populate our Universe. These primordial density fluctuations are predicted to form a homogeneous and isotropic Gaussian field. The predicted statistical distribution of the cosmic microwave background (CMB) temperature fluctuations reflects that of the primordial density fluctuations. Testing this prediction has been the aim of many works in the literature. In particular, since the release of the first year of data collected by the Wilkinson Microwave Anisotropy Probe (WMAP) satellite (Bennett et al. 2003), a considerable number of papers have presented different statistical analyses based on data in real space, spherical harmonics space and wavelet space (Park 2004; Eriksen et al. 2004a,b, 2005; Larson & Wandelt 2004; Hansen, Banday & Görski 2004; Prunet et al. 2005; Vielva et al. 2004; Cruz et al. 2005; Mukherjee C

© 2005 RAS
Higher criticism statistic

& Wang 2004; McEwen et al. 2005). All these works have claimed
the detection of deviations from the predictions of the standard
inflationary model in the WMAP data set optimal for CMB stud-
ies (see Komatsu et al. 2003). Several other works have presented
statistical methods shown to be very powerful in detecting deviations
from the standard inflationary model in the so-called internal linear
combination (ILC) map (Chiang et al. 2003; Coles et al. 2004; Copi,
Huterer & Starkman 2004; Naselsky, Doroshkevich & Verkhodanov
2003; Chiang & Naselsky 2004). In this case, the most convincing
source of deviations is foreground-related.

Deviations from the predictions of the standard inflationary
model can have a cosmological origin. Non-Gaussianity can be
generated under different conditions. A review on the predictions
from several alternative scenarios, including multifield inflation,
inhomogeneous reheating, non-linearities in the gravitational poten-
tial and the curvaton-based model, can be found in Bartolo et al.
(2004). However, these deviations can also be caused by system-
atic effects or by noise associated with the experiment, as well as
by foregrounds (Galactic or extra-Galactic). The first year of data
collected by the WMAP satellite has undergone careful character-
ization and examination by the WMAP team in an attempt to un-
derstand the data set fully (Page et al. 2003; Hinshaw et al. 2003;
Bennett et al. 2003b; Barnes et al. 2003; Jarosik et al. 2003). The
detection of deviations indicated in the previous paragraph shows
the presence of some asymmetry between the Northern and South-
ern hemispheres. Moreover, analyses in wavelet space in different
regions (north, south, northeast, northwest, southeast, southwest of
the Galactic plane) performed by Vielva et al. (2004) and Cruz et al.
(2005) indicate that the source of deviations might be a cold spot lo-
cated at (l ≈ 209°, b ≈ −57°). The nature of the observed deviations
is still not clear.

The implementation and development of new statistical meth-
ods are indispensable for improving our understanding of the deviations
from the standard inflationary model predictions, in particular of
those observed in the WMAP data. There are many kinds of non-
Gaussianity, and each type may be sensitive to some statistical tests,
but immune to others. We need more types of tests, such as the
Higher Criticism (HC) statistic was first proposed in Donoho
& Jin (2004) and Jin (2004). Given n independent observations of a distribution which is
thought to be slightly deviated from the standard Gaussian, one can compare the fraction of observed significances at a given α-level (i.e. the number of observed values exceeding the upper-α quantile of the standard Gaussian) to the expected fraction under the standard Gaussian assumption:

\[ \sqrt{n} \left\{ \frac{\text{(Fraction of significance at } \alpha)}{\alpha(1-\alpha)} \right\}^{0.5}. \]

The HC statistic is then defined as the maximum of the above quantities over all significance levels 0 < α < 1. Given n independent observations \( X_i \) from a distribution which is thought to be symmetric and slightly deviated from the standard Gaussian, there is a simpler equivalent form of HC defined as follows. First, we convert the individual \( X_i \)s into individual p values: \( p_i = P\{N(0, 1) > |X_i|\} \), then we let \( p_{(1)} < p_{(2)} < \ldots < p_{(n)} \) denote the p values sorted in ascending order. Define

\[ HC_{n,i} = \sqrt{n} \left| \frac{i/n - p_{(i)}}{\sqrt{p_{(i)}(1-p_{(i)})}} \right|; \]

the HC statistic is then

\[ HC_n^* = \max_i HC_{n,i}; \]

or, in a modified form,

\[ HC_n^+ = \max_{i: 1/n \leq p_{(i)} \leq (1-1/n)} HC_{n,i}; \]

we let \( HC_n \) refer to either \( HC_n^* \) or \( HC_n^+ \) whenever there is no confusion.

The above definition is slightly different from that in Donoho & Jin (2004), but the ideas are essentially the same.

HC is useful in non-Gaussianity detection when \( X_i \)s are truly from \( N(0, 1) \), with the result that \( HC_{n,i} \) is approximately distributed as \( N(0, 1) \) for almost every i. Thus an unusually large \( HC_n^* \) or \( HC_n^+ \) value strongly implies non-Gaussianity. Moreover, \( HC_{n,i} \) also

\[ \text{http://www.rssd.esa.int/index.php?project=PLANCK} \]
provides localized information on deviations from Gaussianity. We can track down the source of non-Gaussianity by studying which portion of the data gives unusually large HC$_{n,i}$. 

Previous works have claimed the detection of deviations from the predictions of the standard inflationary model in the first-year WMAP data, using the $\kappa$ statistic in wavelet space (Vielva et al. 2004; Mukherjee & Wang 2004; McEwen et al. 2005). In order to establish a comparison between this statistic and the HC we will include calculations of $\kappa$ in our analysis. For completeness, we will also use the so-called Max statistic. The definitions for these two statistics are provided below. A discussion on the theoretical power of the three statistics to detect deviations from Gaussianity is included in the following subsection.

**Kurtosis, $\kappa$.** For a (symmetric) random variable $X$, $\kappa(X)$ is $\kappa(X) = \{E[X^4]/(E[X^2])^2\} - 3$, which uses the fourth moment to measure the departure from Gaussianity. $\kappa$ is useful in the detection of non-Gaussianity because $\kappa(X_1, X_2, \ldots, X_n) \approx \kappa(X)$ for large $n$.

Max. The largest (absolute) observation is a classical statistic:

$$M_n = \max(|X_1|, |X_2|, \ldots, |X_n|).$$

Max is useful in detection of non-Gaussianity because $M_n \approx \sqrt{2\log n}$ when $X_i$ are truly from $N(0, 1)$; thus a significant difference between $M_n$ and $\sqrt{2\log n}$ implies non-Gaussianity.

### 2.1 Comparison of Higher Criticism, maximum and excess kurtosis statistics

The aim of this section is to establish a simple theoretical comparison among the three statistics applied in this paper. We show the power of the various statistical tests in detecting the distortion generated by a faint non-Gaussian signal (modelled as a function of the decaying rate of the tail of the distribution) superposed on a Gaussian signal. Similar to in the analysis presented in Jin et al. (2004), the superposed image can be thought of as a Gaussian signal. Similar to in the analysis presented in Jin et al. (2004), the superposed image can be thought of as $Y = N + G$, where $Y$ is the observed image, $N$ is the non-Gaussian component, and $G$ is the Gaussian component [assumed to have mean zero and dispersion one, $N(0, 1)$]. We study the power of the three statistics in testing whether $N = 0$ or not.

One can do such a test either in real space (the space of the observations) or, as in this paper, in wavelet space. For sufficiently fine resolution, the wavelet coefficients $X_i$ of $Y$ can be modelled as

$$X_i = \sqrt{1-\lambda} \cdot z_i + \sqrt{\lambda} \cdot w_i, \quad 1 \leq i \leq n,$$

where $n$ is the number of observations, $1 \geq \lambda \geq 0$ is a parameter, $z_i \sim N(0,1)$ are the transform coefficients of the Gaussian component $G$, and $w_i \sim W$ are the transform coefficients of the non-Gaussian component $N$. $W$ is some unknown distribution with vanishing first- and third-order moments (this constraint is imposed only in order to simplify the possible range of non-Gaussian distributions, and it will be representative of a Gaussian distribution modified by a few large values contributing to a tail). Without loss of generality, we assume that the standard deviations of both $z_i$ and $w_i$ are 1. Phrased in statistical terms, the problem of detecting the existence of a non-Gaussian component is equivalent to discriminating between the null hypothesis and the alternative hypothesis, where

$$H_0 : \quad X_i = z_i,$$

$$H_1 : \quad X_i = \sqrt{1-\lambda} \cdot z_i + \sqrt{\lambda} \cdot w_i, \quad 0 \leq \lambda \leq 1.$$

$N \equiv 0$ being equivalent to $\lambda \equiv 0$.

In order to obtain a quantitative estimate of the ability of the three statistics to detect the non-Gaussian component, we parametrize the tail probability of $W$ as follows.

$$\lim_{x \to \infty} x^{\alpha} P( |W| > x ) = C_\alpha, \quad C_\alpha \text{ is a constant}.$$  

To model increasingly challenging situations as the number of observations increases, we calibrate $\lambda$ to decay with $n$ as $\lambda = \lambda_n = n^{-r}$.

To constrain the detectability of such a non-Gaussian distribution superposed on a Gaussian one, we search the $r$ and $\alpha$ parameter space. We define the regions of this space that will be detectable under the various statistical tests. It was shown in Jin (2004) that there is a curve in the $r-\alpha$ plane that separates the detectable regions of parameters from the undetectable regions. That curve is given by

$$r = \begin{cases} 2/\alpha, & \alpha \leq 8, \\ 1/4, & \alpha > 8. \end{cases}$$

In Fig. 1, we compare the results of the HC, the Max and the $\kappa$ statistics. When it is possible to detect, HC or Max are better than $\kappa$ when $\alpha \leq 8$, $\kappa$ is better than HC or Max when $\alpha > 8$.

To conclude this section, we remark that the performance of HC, Max, and $\kappa$ depends on different types of non-Gaussianity. Intuitively, Max is designed to capture evidence of unusual behaviour of the few most extreme observations against the Gaussian assumption. The HC statistic is able to capture unusual behaviour of the few most extreme observations, as well as an unusually large number of moderately high observations. Thus HC is better than Max, in general. However, when the evidence against the Gaussian assumption truly lies in the most extreme observations, HC and Max are almost equivalent. In contrast, the $\kappa$ statistic is designed to capture the evidence hidden in the fourth moment. Therefore this statistic depends on the bulk of the data, rather than on a few extreme observations or a small fraction of relatively large observations. Finally, the HC statistic offers an automatic way to find the area of the data accounting for the non-Gaussianity, while the Max and $\kappa$ statistics do not have this capability.

![Figure 1](https://i.imgur.com/3EoG.png)  
**Figure 1.** Detectable regions in the $\alpha-r$ plane. When $(\alpha, r)$ is in the white region on the top or the undetectable region, all possible statistics fail asymptotically for detection. When $(\alpha, r)$ is in the white region on the bottom, both $\kappa$ and Max/HC are able to detect reliably. In the shaded region to the left, Max/HC is able to detect reliably, but $\kappa$ completely fails, and in the shaded region to the right, $\kappa$ is able to detect reliably, but Max/HC completely fail asymptotically.
3 WMAP FIRST-YEAR DATA AND SIMULATIONS

The data collected by the WMAP satellite during the first year of operation are available at the Legacy Archive for Microwave Background Data Analysis (LAMBDA) website. Following Komatsu et al. (2003), the analysis presented in this work is based on a data set obtained by the weighted combination of the released foreground-cleaned intensity maps at bands $Q, V$ and $W$. Each of these maps can be downloaded in the HEALPIX nested format at resolution $nside = 512$ (the total number of pixels being $12 \times nside^2$). The co-added map is obtained by the following combination:

$$T(i) = \frac{\sum_{r=3}^{10} T_r(i) w_r(i)}{\sum_{r=3}^{10} w_r(i)}.$$  

The temperature at pixel $i$, $T(i)$, results from the ratio of the weighted sum of temperatures at pixel $i$ at each radiometer divided by the sum of the weights of each radiometer at pixel $i$. The radiometers $Q_1, Q_2, V_1, V_2, W_1, W_2, W_3$ and $W_4$ are sequentially numbered from 3 to 10. The weight at each pixel, for each radiometer $w_r(i)$, is the ratio of the number of observations $N_r(i)$ divided by the square of the receiver noise dispersion $\sigma_r$. This results in a co-added map at resolution $nside = 512$. This map is downgraded to resolution $nside = 256$ before the analyses are performed. In order to remove Galactic and point-source emission, we applied the most conservative mask provided by the WMAP team, the Kp0 mask. This follows the procedure first presented by Vielva et al. (2004).

In order to detect any possible deviations from the predictions of the standard inflationary model we compared the values of several statistics in the WMAP data set described above with those obtained from 5000 Monte Carlo simulations. The temperature at a given pixel $i$ (pointing towards a direction characterized by polar angles $\theta_i$ and $\phi_i$) can be expressed as an expansion in spherical harmonics $Y_{lm}(\theta_i, \phi_i)$:

$$T(\theta_i, \phi_i) = \sum_{l,m} a_{lm} Y_{lm}(\theta_i, \phi_i).$$

The Monte Carlo simulations were performed assuming a Gaussian distribution $N(0, C_l)$ for the spherical harmonic coefficients $a_{lm}$, where $C_l$ represents the power spectrum that best fits the WMAP, CBI and ACBAR CMB data, plus the 2dF and Lyman-alpha data. Beam transfer functions as well as number of observations and receiver noise dispersion were taken into account when simulating data taken by each of the receivers (all of these, as well as the best-fitting power spectrum, are provided by the WMAP team in the LAMBDA website).

The analysis was carried out in both real space and wavelet space. We convolved the WMAP data set and the Monte Carlo simulations with the spherical Mexican hat wavelet (SMHW) at 15 scales, following the same procedure as presented in Vielva et al. (2004) and Cruz et al. (2005). The scales numbered from 1 to 15 correspond to 13.74, 25.0, 50.0, 75.0, 100.0, 150.0, 200.0, 250.0, 300.0, 400.0, 500.0, 600.0, 750.0, 900.0, and 1050.0 arcmin. In order to avoid border effects caused by the mask included in the WMAP data set and the simulations, analyses are performed outside extended masks defined at each scale. Given the extended mask is the Kp0 mask plus pixels near the Galactic plane that are within 2.5 times the scale. All of these a scale, extended masks were presented in fig. 2 of Vielva et al. (2004). As discussed in the next section, deviations are detected in wavelet space. This shows once again the value of wavelets in providing a space in which certain features that might be causing deviations from the predictions of the standard inflationary model are enhanced.

4 ANALYSIS

Based on 5000 simulations, we calculated the 68, 95 and 99 per cent confidence regions for each of the four statistics ($\kappa$, Max, and the two HC tests) at each of the 15 wavelet scales used. We used two-sided confidence regions for $\kappa$, as it is symmetric about 0 under the null hypothesis that the data are Gaussian. Max and HC statistics are defined so that the larger the statistic, the stronger the evidence against the null hypothesis. Therefore, we used one-sided confidence regions for these statistics.

Fig. 2 shows the results obtained based on $\kappa$. Crosses denote the value of this statistic for the WMAP data set. Bands outlined by dashed, dot–dashed, and solid lines correspond to the 68, 95 and 99 per cent confidence regions respectively (symbols and lines represent the same quantities in all figures included in this paper unless a comment is added). Clearly, the results agree with those presented in Vielva et al. (2004). Figs 3 and 4 show the new results based on the Max and the HC statistical tests. The WMAP data appear outside the 99 per cent confidence level obtained based on the predictions of the standard inflationary model at scale 9 (300 arcmin) for these three tests. In particular, $\approx 99.46$ per cent of the 5000 simulations have Max and HC/HC+ values below the one obtained from the WMAP data set. Therefore in this particular case these statistics are as competitive as the kurtosis in detecting deviations from the assumed null hypothesis.

4.1 Location of the outlying pixels provided by the Higher Criticism statistic

As discussed above, the HC statistical tests introduced in this paper are able to detect deviations from the standard inflationary model in the WMAP data at scales around 300 arcmin. Moreover, the power

Figure 2. Values of the $\kappa$ statistic for the analysed WMAP data set crosses. The bands outlined by dashed, dot–dashed and solid lines correspond to the 68, 95 and 99 per cent confidence regions respectively.
of this new test resides in its ability to provide a direct way to determine which pixels in the WMAP data set are generating these deviations. By comparing the values of the HC test at each pixel with the value that defines the 99 per cent confidence limit at scale 9, we were able to extract a total of 490 pixels that are causing the detected deviation. Fig. 5 shows the individual pixel values of the HC statistic for the WMAP data set at scale 9. The extracted pixels above the 99 per cent limit are plotted in the map shown in Fig. 6. As can be seen, these pixels define a ring centered at position \((l \approx 209^\circ, b \approx -57^\circ)\). It is important to note that the correlations introduced by the convolution with the wavelet need to be taken into account in order to interpret this result properly. Some pixels within the ring could also initially (in the map in real space) deviate from Gaussianity.

in this paper as well as in previous papers by Vielva et al. (2004), Cruz et al. (2005), Mukherjee & Wang (2004) and McEwen et al. (2005), disappeared as the pixels in the ring \((l \approx 209^\circ, b \approx -57^\circ)\) were removed. These pixels are part of the cold spot pointed out by Vielva et al. (2004) and Cruz et al. (2005) as being the source of the observed deviations. The Max and the HC values decrease as well (after removal of the pixels in the ring), fitting now within the 68 per cent confidence region. The HC statistic is useful in that it offers an automatic detection of the non-Gaussian portions of the data. This characteristic is not easily available for other tests, such as kurtosis.

5 CONCLUSIONS AND DISCUSSION

This paper presents an analysis of the compatibility of the distribution of the CMB temperature fluctuations observed by WMAP with
the predictions from standard inflation. The analysis was based on the recently developed HC statistic. This statistic is especially sensitive to two types of deviations from Gaussianity: an unusually large number of moderately significant values or a small fraction of unusually extreme values. In both of these cases the HC statistic is optimal (Donoho & Jin 2004, Jin et al. 2004). Moreover, the definition of the statistic naturally suggests a direct way to locate the possible sources of non-Gaussianity. Even if the wavelet transform is a linear process, the original distribution is slightly distorted as the scale of the wavelet, and the area covered by the mask, increases. We are working on improving the power of the HC statistic to be optimal in the detection of deviations from Gaussianity when applied in wavelet space.

We compared the performance of the HC statistic with that of $\kappa$ and Max. The comparison was based on the analysis of the first-year WMAP data, in real and in wavelet space. Distributions of the three statistics were built on 5000 simulations assuming the standard inflationary model predictions as well as the constraints coming from the WMAP observations. The three statistics provided comparable results, pointing to the presence of deviations at a wavelet scale of $5^\circ$. We made use of the HC statistic to identify automatically the pixels in the WMAP data set that were causing such deviations. A ring centred at ($l \approx 209^\circ$, $b \approx -57^\circ$) containing 490 pixels was shown to be the cause of the detected deviations. Removal of this set of pixels from the WMAP data made the data compatible with the predictions from the standard inflationary model. One should be cautious when interpreting this result. The detection is achieved in wavelet space. Even if only the pixels in the ring appear as outliers of the HC distribution at a wavelet scale of $5^\circ$, some other pixels within the ring could be involved in the observed deviation. Convolution with the wavelet introduces correlations between pixels that need to be properly taken into account in the definition of the statistic to allow a correct interpretation. It is important to note that the pixels selected as the source of the observed deviations are part of the cold spot pointed out by Vielva et al. (2004) and Cruz et al. (2005). A careful study of the possible influence of foregrounds, noise and beam distortion and assumption of a certain power spectrum was carried out in those two papers.

To conclude, we would like to remark on the power of the HC statistic for detecting and locating possible sources of non-Gaussianity. In particular, regarding analysis of the CMB, the HC statistic can be useful at several steps in the data processing and final analysis, from the study of distortions caused by systematic effects in the time-ordered data to the study of the compatibility of the statistical distribution of the observed temperature fluctuations with predictions from theories of the very early Universe.

ACKNOWLEDGMENTS

We thank P. Vielva for providing the extended masks used at the various wavelet scales. We thank Douglas Crabill for helping with some of the figures. We acknowledge the use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science. Some of the results in this paper have been derived using the healpix (Górski, Hivon & Wandelt 1999) package. LC and AT acknowledge the Indiana Space Grant Consortium for financial support. AT acknowledges the Spira Fellowship (summers of 2004 and 2005).

REFERENCES

Barnes C. et al., 2003, ApJS, 148, 51
Bartolo N., Komatsu E., Materrese S., Riotto A., 2004, Phys. Rept. 402, 103
Bennett C. L. et al., 2003a, ApJS, 148, 1
Bennett C. L. et al., 2003b, ApJS, 148, 97

Figure 7. Values of $\kappa$, Max and HC tests for the analysed WMAP data set after subtracting the pixels that were causing the deviations from the predictions of the standard inflationary model (crosses). The bands outlined by dashed, dot–dashed and solid lines correspond to the 68, 95 and 99 per cent confidence regions, respectively, for $\kappa$. The dashed, dot–dashed and solid lines correspond to the 68, 95 and 99 per cent confidence levels respectively for Max, HC and HC+. 
