# Non-Gaussian statistics and the microwave background radiation 

Peter Coles and John D. Barrow Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH

Accepted 1987 April 14. Received 1987 April 1; in original form 1987 February 20


#### Abstract

Summary. We show how to calculate statistical properties of non-Gaussian random fields. We apply this method to determine the mean size and frequency of occurrence of high and low level excursions of the Rayleigh, Maxwell, Chisquared, lognormal, rectangular and Gumbel type I random fields. These results permit us to calculate the expected size and frequency of fine-scale hotspots and coldspots expected in the microwave background distribution on the sky under the assumption that it possesses non-Gaussian statistics of the above-mentioned types. This generalizes and extends previous studies which confined attention to the simpler case in which the microwave background radiation was assumed to be a Gaussian random field. We also discuss whether it will be possible to determine observationally whether the underlying statistics of the temperature fluctuations in the microwave background are indeed Gaussian as predicted by the standard theory of inflation.


## 1 Introduction

Observational evidence that rich clusters of galaxies are more strongly clustered than galaxies (Bahcall \& Soneira 1983; Klypin \& Kopylov 1983; Ling, Frenk \& Barrow 1986) led to the introduction of the idea of 'biased' galaxy formation (Kaiser 1984) wherein the distribution of luminous matter is not assumed to be a faithful representation of the total mass distribution. Since luminous galaxies are most likely to form in the highest density regions and the maxima of a Gaussian field are more strongly clustered than the underlying field this may provide a natural explanation for the strong clustering of Abell clusters. Davis et al. (1985) recognized that the identification of galaxies only with $3 \sigma$ peaks in the mass distribution could reconcile the results of simulated galaxy clustering with the observed covariance function in cosmological models possessing a present density close to the critical density. Subsequently, various ideas were put forward to show how a bias could be introduced into the galaxy formation process by physical processes occurring in the Universe after recombination (Rees 1985; Couchman \& Rees 1986; Silk 1985; Dekel \& Silk 1986). A priori it is by no means obvious that biasing is inevitable. One could easily imagine a form of 'anti-biasing' in which the densest regions of the Universe undergo very rapid evolution through a luminous phase and are now dark voids.

Motivated by these questions a number of detailed mathematical analyses have been made concerning the statistical properties of regions of high density in Gaussian random fields (Kaiser 1984; Politzer \& Wise 1984; Peacock \& Heavens 1985;Otto, Politzer \& Wise 1986; Jensen \& Szalay 1986; Bardeen et al. 1986; Coles 1986; Couchman 1986, 1987) following the earlier introduction of this general idea in a pioneering paper of Doroshkevich (1970), [see also Sunyaev 1972 and Marochnik, Nasel'skii \& Zabotin 1980). All these analyses make use of the extensive literature on the statistics of Gaussian random fields, the most notable surveys being those of Rice (reprinted in Wax 1954), Adler (1981) and Vanmarcke (1983).

The importance of considering the maxima of random fields has also been recognized (Sazhin 1985a, b; Zabotin \& Nasel'skii 1985) in connection with predicting the map of fine-scale microwave temperature fluctuations expected over the celestial sphere as a result of the gravitational potential fluctuations in the Universe that gave rise to galaxies and clusters. A comprehensive study of this problem has been completed by Bond \& Efstathiou (1987) and Vittorio \& Juszkiewicz (1987). All these analyses of the microwave background assume that the temperature fluctuations are a Gaussian random field.

We should point out that there is particular interest in the possible non-Gaussian nature of fluctuations in the Universe. Although fluctuations generated during an inflationary de Sitter phase are expected to be Gaussian (with a scale-invariant spectrum), those generated during an inflationary stage which is not exactly exponential will not, in general, be Gaussian (nor will they be scale invariant). The initial fluctuations present at recombination are necessarily non-Gaussian fluctuations in vacuum string theories for the origin of large-scale structure. It is also clear that initially Gaussian fluctuations become non-Gaussian under the influence of non-linear gravitational clustering. Although the mathematical analysis is considerably more complicated in the non-Gaussian case there exist a number of physically realistic non-Gaussian fields which can be treated analytically using various mathematical tricks. All these examples can be related in some functional way to the Gaussian so that certain aspects of the Central Limit Theorem, can be exploited. We can avoid much of the analytical complexity involved in the treatment of twodimensional random fields by constructing them from one-dimensional random processes (this trick was used by Coles (1986) to obtain results for a three-dimensional Gaussian field in connection with the problem of rich-cluster correlations). We find that it is possible to deduce the expected number of regions where the field exceeds some level and also the expected area of such regions. We apply these results to the microwave background sky and compare them with those obtained earlier for the Gaussian case. The examples we consider exhibit a wide range of probabilistic behaviours, from those which are very similar to the Gaussian to some which are qualitatively different.
The layout of the paper is as follows: in Section 2 we demonstrate how to calculate the required statistical properties of non-Gaussian random fields, using the mean upcrossing rate of a random process. In Section 3 we apply the method to various non-Gaussian fields and in Section 4 we discuss the various normalizations required in order to compare Gaussian and non-Gaussian behaviour. In Section 5 we apply these results to the case of microwave background temperature fluctuations and compare the results for Gaussian and non-Gaussian fluctuations. Finally, in Section 6, we discuss briefly whether observations of the microwave background might allow us to determine whether the spectrum of fluctuation in the Universe possesses Gaussian or nonGaussian statistics.

## 2 Average properties of upcrossings

In this section we shall determine mean properties of the peaks of random fields from the rate of upcrossing of the field in its excursions above some general level. First, we establish a useful
expression for the upcrossing rate of a one-dimensional random process $X(\theta)$.
We introduce notations
$X(0) \equiv x$,
$y_{q} \equiv \frac{1}{q}[X(q)-X(0)]$.
Now define the joint probability density function of $x$ and $y_{q}$ by $g_{q}\left(x, y_{q}\right)$ and assume that
$\operatorname{Lt}_{q \rightarrow 0} g_{q}\left(x, y_{q}\right)=p(x, y)$
where $y=d X / d \theta$. Define the upcrossing of the level $u$ to be an event such that
$X(0)<u<X(q) ; \quad q \geqslant 0$.
Using (2.2), we see this event is equivalent to the simultaneous requirement that $X(0)<u$ and $y_{q}>[u-X(0)] / q$. Hence, we consider the associated joint probability

$$
\begin{align*}
J_{q}(u) & =\frac{1}{q} P\left[X(0)<u \text { and } y_{Q}>[u-X(0)] / q\right]  \tag{2.5}\\
& =\frac{1}{q} \int_{-\infty}^{u} d x \int_{(u-x) / q}^{\infty} g_{q}(x, y) d y \tag{2.6}
\end{align*}
$$

If we introduce new variables $(u, v)$ such that $(x, y)=(u-q z v, z)$ then (2.6) becomes
$J_{q}(u)=\int_{0}^{\infty} d z \int_{0}^{1} g_{q}(u-q z v, z) d v$.
Hence, if there exists some function $h(z)$ such that, for any $u, q$
$\int_{0}^{\infty} z h(z)<\infty \quad$ and $g_{q}(u, z) \leqslant h(z)$,
then,
$\operatorname{Lt}_{q \rightarrow 0} J_{q}(u)=\int_{0}^{\infty} z p(u, z) d z$.
The function $p(x, y)$ is the joint distribution function of $X(\theta)$ and $X^{\prime}(\theta)$ and for Gaussian processes it will be a bivariate Gaussian. In what follows we shall denote the upcrossing rate of level $u$, (2.9), by $N_{u}$; that is:
$N_{u}=\int_{0}^{\infty} z p(u, z) d z$.
The one-dimensional upcrossing formula (2.10) can be used to calculate the mean number and size of peaks above a given threshold for two-dimensional random fields, whether or not they are Gaussian. In Section 3 we shall evaluate $N_{u}$ for a number of non-Gaussian fields.

We denote the duration of excursions of a one-dimensional random field $X(\theta)$ above a level $X(\theta)=u$ by $\theta_{u}$, and the duration of the intervening excursions below this level by $\theta_{u}^{\prime}$ as illustrated in Fig. 1.


Figure 1. Definition of the quantities $\theta_{u}$ and $\theta_{u}^{\prime}$ introduced in the text. Their values generate the distributions of the durations of the excursions of a random process $X(\theta)$ above and below the threshold $X(\theta)=u$ respectively.

Then, we have the following expectation values if the probability density function (pdf) of the field is $f(x)$ :
$E\left[\theta_{u}+\theta_{u}^{\prime}\right]=\frac{1}{N_{u}}$
$\frac{E\left[\theta_{u}\right]}{E\left[\theta_{u}+\theta_{u}^{\prime}\right]}=\int_{u}^{\infty} f(x) d x \equiv Q(u)$
where $Q(u)$ is the complementary cumulative distribution function (ccdf). Hence
$E\left[\theta_{u}\right]=\frac{Q(u)}{N_{u}}$.
We can obtain expectation values for the area $A_{u}$ of a two-dimensional isotropic random field above some threshold level, $u$, by considering two orthogonal and independent one-dimensional processes $X^{(1)}(\theta)$ and $X^{(2)}(\theta)$. Thus,
$E\left[A_{u}\right]=E\left[\theta_{u}^{(1)}\right] E\left[\theta_{u}^{(2)}\right]$,
so, by (2.13),
$E\left[A_{u}\right]=\left(\frac{Q(u)}{N_{u}}\right)^{2}$.
We should emphasize that since the shapes of the areas above some level crossing are not known $a$ priori and can only be computed in the case of a Gaussian field, the 'areas' computed here are representative two-dimensional quantities equal to the square of a characteristic length scale. This procedure produces the correct result for the Gaussian case and will give reliable comparative results since the same criterion is used in each non-Gaussian case.

Consider a large plane area $A$. We can express the subarea within it at which the field exceeds a certain level $u$ in two different ways and equate them to obtain
$E\left[A_{u}\right] \mu_{u} A=Q(u) A$
where $\mu_{u}$ is the mean number of regions above the level $u$ per unit area. Hence,
$\mu_{u}=\frac{Q(u)}{E\left[A_{u}\right]}=\frac{N_{u}^{2}}{Q(u)}$.

The formulae (2.15) and (2.17) enable us to calculate mean sizes and frequencies of excursions of any two-dimensional random field for which we can calculate $N_{u}$ and $Q(u)$ in simple form. In the next section we shall do this for six non-Gaussian distributions and compare the results with those obtained for the two-dimensional Gaussian random field.

## 3 The calculation of $N_{u}$ and $Q(u)$

We shall be interested in a group of non-Gaussian processes which can be derived from the standard Gaussian process. We denote the second spectral moment of these processes by $\sigma_{1}^{2}$; hence, if the processes possess a covariance function $\xi(\theta)$ then*
$\sigma_{1}^{2}=-\xi^{\prime \prime}(0)$.
In this section $X(\theta)$ will denote a one-dimensional Gaussian random processes with zero mean and unit variance. For comparative reference we give the results for a Gaussian field first.
(i) Gaussian

The joint distribution function $p(x, y)$ is a bivariate normal distribution with
$p(x, y)=\frac{1}{2 \pi \sigma_{1}} \exp \left(-\frac{x^{2}}{2}-\frac{y^{2}}{2 \sigma_{1}^{2}}\right)$,
where $Y(\theta)=X^{\prime}(\theta)$. Hence, from (2.10), we have
$N_{u}=\frac{\sigma_{1}}{2 \pi} \exp \left(-\frac{u^{2}}{2}\right)$
and the ccdf, $Q(u)$, is a standard tabulated function (Abramowitz \& Stegun 1965, table 26.5) for the one-dimensional Gaussian:

$$
\begin{equation*}
Q(u)=\frac{1}{\sqrt{2 \pi}} \frac{\Phi(u)}{u} \exp \left(-\frac{u^{2}}{2}\right) \tag{3.4}
\end{equation*}
$$

where $\Phi(u)$ is the infinite series
$\Phi(u)=1-\frac{1}{u^{2}}+\frac{1.3}{u^{4}}-\frac{1.3 .5}{u^{6}}+\ldots$.
This case is particularly simple because if $X$ and $Y$ are stationary Gaussian processes then they are necessarily independent and hence the joint distribution (3.2) is a separable function of $x$ and $y$.
(ii) Rayleigh

If $X_{1}(\theta)$ and $X_{2}(\theta)$ are independent standard Gaussian processes then the Rayleigh process is defined by
$R(\theta)=\left[X_{1}^{2}(\theta)+X_{2}^{2}(\theta)\right]^{1 / 2}$
and it possesses a pdf
$f_{R}(u)=u \exp \left(-\frac{u^{2}}{2}\right)$.

[^0]Given that $X_{i}(\theta)=x_{i}$ we have
$R^{\prime}(\theta)=\frac{x_{1} X_{1}^{\prime}(\theta)+x_{2} X_{2}^{\prime}(\theta)}{\left[x_{1}^{2}+x_{2}^{2}\right]^{1 / 2}}$
Since $X_{1}(\theta)$ and $X_{2}(\theta)$ are independent and normally distributed, $R^{\prime}(\theta)$ will be normally distributed with zero mean and variance $\sigma_{1}^{2}$. [The latter deduction can be verified by noting that the variance of the right hand side of (3.8) is $\left.\left(x_{1}^{2} \sigma_{1}^{2}+x_{2}^{2} \sigma_{1}^{2}\right) /\left(x_{1}^{2}+x_{2}^{2}\right)=\sigma_{1}^{2}.\right]$

The distribution of $R^{\prime}(\theta)$ conditional on $X_{i}(\theta)=x_{i}$ is independent of $x_{i}$ and hence independent of $R$. Therefore

$$
\begin{align*}
p(u, y) & =\frac{u}{\sqrt{2 \pi \sigma_{1}}} \exp \left(-\frac{u^{2}}{2}-\frac{y^{2}}{2 \sigma_{1}^{2}}\right)  \tag{3.9}\\
& =\operatorname{Prob}\left[(R=u) \quad \text { and } \quad\left(R^{\prime}=y\right)\right] . \tag{3.10}
\end{align*}
$$

So, using (2.10), we have
$N_{u}=\int_{0}^{\infty} y p(u, y) d y=\frac{u \sigma_{1}}{\sqrt{2 \pi}} \exp \left(-\frac{u^{2}}{2}\right)$.
The ccdf for (3.7) is just
$Q(u)=\exp \left(-\frac{u^{2}}{2}\right) \equiv Q_{1}(u)$.
(iii) Maxwell

This is a generalization of (ii) defined by
$M(\theta)=\left[X_{1}^{2}(\theta)+X_{2}^{2}(\theta)+X_{3}^{2}(\theta)\right]^{1 / 2}$
where $X_{i}$ are independent standard Gaussians, and the pdf is
$f_{M}(u)=\sqrt{2 / \pi} u^{2} \exp \left(-\frac{u^{2}}{2}\right)$.
Proceeding as in Section 3.2 we find that
$N_{u}=\frac{\sigma_{1}}{\pi} u^{2} \exp \left(-\frac{u^{2}}{2}\right)$
and

$$
\begin{equation*}
Q(u)=2 Q_{1}(u)+\sqrt{2 / \pi} u \exp \left(-\frac{u^{2}}{2}\right) \tag{3.16}
\end{equation*}
$$

where $Q_{1}(u)$ is defined by (3.12).
(iv) Chi-squared $\chi_{n}^{2}$

If we have $n$ independent standard Gaussian processes $X_{i}(\theta), i=1, \ldots n$ then the chi-squared process of order $n$ is defined by
$\chi_{n}(\theta)=\sum_{i=1}^{n} X_{i}^{2}(\theta)$
and possesses the pdf
$f_{x}(u)=\frac{1}{u}\left(\frac{u}{2}\right)^{n / 2} \frac{\exp (-u / 2)}{\Gamma(n / 2)}$.
From (3.17) it follows that
$\Psi_{n}(\theta) \equiv \frac{d \chi_{n}(\theta)}{d \theta}=2 \sum_{i=1}^{n} X_{i}(\theta) X_{i}^{\prime}(\theta)$.
We can rewrite (2.10) as
$N_{u}=E\left(\Psi_{n}^{+} \mid \chi_{n}=u\right) f\left(\chi_{n}=u\right)$
where
$\Psi_{n}^{+}=\max \left[0, \Psi_{n}(0)\right]$.
First consider
$E\left[\Psi_{n}^{+} \mid X_{i}=X_{i}(0)=x_{i}\right.$ and $\left.\chi_{n}=u\right]$.
Now the distribution of $\Psi_{n}$ given that $X_{i}=x_{i}(i=1,2, \ldots n)$ and $\chi_{n}=u$ is normal with mean zero and variance equal to $4 u \sigma_{1}^{2}$. Thus
$E\left(\Psi_{n}^{+} \mid X_{i}=x_{i}\right.$ and $\left.\chi_{n}=u\right)=\left(\frac{2 u \sigma_{1}^{2}}{\pi}\right)^{1 / 2}$,
and since the right-hand-side is independent of the individual $x_{i}$ we have
$E\left(\Psi_{n}^{+} \mid \chi_{n}=u\right)=\left(\frac{2 u \sigma_{1}^{2}}{\pi}\right)^{1 / 2}$.
So, using (3.18), we find the mean upcrossing rate to be (see Sharpe 1978 and Hasofer 1974 for a more laborious derivation of this result)
$N_{u}=\frac{\sigma_{1}}{\sqrt{\pi}}\left(\frac{u}{2}\right)^{(n-1) / 2} \frac{\exp (-u / 2)}{\Gamma(n / 2)}$
and the ccdf reduces to one- and two-point gamma functions:
$Q(u)=\frac{\Gamma(n / 2, u / 2)}{\Gamma(n / 2)} \equiv Q_{2}(u)$.
$Q_{2}(u)$ is also a tabulated function (Abramowitz \& Stegun 1965, table 26.8).
(v) Lognormal

If $X(\theta)$ is the standard Gaussian process then
$Y(\theta)=\exp [X(\theta)]$,
$Y^{\prime}(\theta)=X^{\prime}(\theta) \exp [X(\theta)]$
defines the lognormal process with pdf
$f_{L}(u)=\frac{1}{u \sqrt{2 \pi}} \exp \left[-\frac{1}{2}(\log u)^{2}\right]$.

Since
$P(X=u)=\frac{1}{\sqrt{2 \pi}} \exp \left(-\frac{u^{2}}{2}\right)$
we have
$p(u, y)=\frac{1}{2 \pi \sigma_{1}} \exp \left(-\frac{u^{2}}{2}-\frac{y^{2}}{2 \sigma_{1}^{2}}\right)$.
The joint distribution of $Y$ and $Y^{\prime}, \Pi\left(Y, Y^{\prime}\right)$, is related to $P\left(X, X^{\prime}\right)$ by
$\Pi\left(Y, Y^{\prime}\right)=P\left(X, X^{\prime}\right)\|J\|$
where $\|J\|=1 / Y^{2}$ is the Jacobian of the transformation $\left(X, X^{\prime}\right) \rightarrow\left(\log Y, Y^{\prime} / Y\right)$ defined by (3.27) and (3.28). Hence
$\Pi\left(Y=\theta, Y^{\prime}=\psi\right)=\frac{1}{\theta^{2}} \frac{1}{2 \pi \sigma_{1}} \exp \left[-\frac{(\log \theta)^{2}}{2}-\frac{\psi^{2}}{2 \sigma_{1}^{2} \theta^{2}}\right]$
and we obtain, after some algebra, that for the lognormal process
$N_{u}=\frac{\sigma_{1}}{2 \pi} \exp \left[-\frac{(\log u)^{2}}{2}\right]$.
That is, $Y$ crosses the level $u$ as frequently as $X$ crosses the level $\log u$ (compare $c f$. equation 3.3) as expected.
The ccdf for the lognormal is simply
$Q(u)=Q_{1}(\log u)$.
(vi) Rectangular

If $X_{1}$ and $X_{2}$ are independent standard Gaussian processes then
$Z=\exp \left[-\frac{1}{2}\left(X_{1}^{2}+X_{2}^{2}\right)\right]$
has uniform density on the interval [ 0,1 ]. Defining a Rayleigh process
$R=\left(X_{1}^{2}+X_{2}^{2}\right)^{1 / 2}$
we have
$Z=\exp \left(-R^{2} / 2\right)$
and it is straightforward to evaluate the Jacobian of the transformation $\left(R, R^{\prime}\right) \rightarrow\left(Z, Z^{\prime}\right)$ to obtain the joint probability distribution
$\Pi\left(Z=u ; Z^{\prime}=v\right)=\frac{1}{\sqrt{2 \pi \sigma_{1} u}} \exp \left(-\frac{v^{2}}{2 u^{2} \sigma_{1}^{2}}\right)$.
Hence, integrating, we obtain the upcrossing rate from equation (2.10) to be
$N_{u}=\frac{\sigma_{1} u}{\sqrt{2 \pi}}$.
The ccdf is discontinuous:

$$
\begin{align*}
Q(u) & =1-u, & & u \leqslant 1 \\
& =1, & & u>1 \tag{3.41}
\end{align*}
$$

(vii) Gumbel type I

The Gumbel type I distribution has pdf
$p(x)=\exp [-x-\exp (-x)], \quad-\infty<x<\infty$
and is well-known in the study of order statistics. ${ }^{\star}$
The associated Gumbel type II and type III distributions can be obtained via a transformation $x \rightarrow \log x$ in (3.42) and differ in the domain of dependence. If $X_{1}$ and $X_{2}$ are Gaussians with unit variance and zero mean then the process
$\Pi\left(X_{1}, X_{2}\right)=-\log \left(\frac{X_{1}^{2}+X_{2}^{2}}{2}\right)$
has the Gumbel-I density because $\frac{1}{2}\left(X_{1}^{2}+X_{2}^{2}\right)$ possesses an exponential distribution. The joint probability of the Gumbel-I process and its derivative is thus
$p(u, v)=\frac{1}{2 \sigma_{1} \sqrt{ } \pi} \exp \left[-\frac{3 u}{2}-\exp (-u)+\frac{v^{2}}{2 \sigma_{1}^{2}} \exp (-u)\right]$
and the upcrossing rate is
$N_{u}=\frac{\sigma_{1}}{\sqrt{ } \pi} \exp \left[-\frac{u}{2}-\exp (-u)\right]$.
The ccdf is given by
$Q(u)=1-\exp [-\exp (-u)]$.
In all the examples calculated above the abundance and size of low-intensity regions can be obtained by noting that the requisite downcrossing rate of the field below some level is equal to the upcrossing rate in each case whilst the complementary cumulative function $Q(u)$ should be replaced by the cumulative function $P(u)=1-Q(u)$ in the expressions (2.13)-(2.17). We should point out that the non-Gaussian processes (ii)-(vii) admit exact solutions for the mean frequency and size of excursions because they are all at some level related to Gaussian processes. However, the form of the resulting processes is sufficiently varied and comprehensive to provide a useful guide as to the robustness of the conclusions drawn about the statistics of excursions using solely Gaussian processes. Some of our examples (ii)-(iv) behave in a very similar fashion to the Gaussian process whilst the others do not. To keep the algebra simple we have used underlying Gaussian processes with unit variance. Different distributions are obtained if the underlying process has variance $\sigma^{2}$. For reference, the means and variances of the distributions (i)-(viii) obtained in the general case are displayed in Table 1 and distribution functions for the case $\sigma^{2}=1$ are shown in Fig. 2.
The transformations which we have applied to the Gaussian process in order to obtain analytic expressions for the marginal distributions are by no means unique. For example, in the case of the rectangular distribution (vi), instead of choosing $Z$ as in (3.36) we could have taken
$Z=\left(\frac{1}{2 \pi}\right) \arctan \left(X_{1} / X_{2}\right)$
*The Gumbel distribution arises in the study of first-ranked members of independent subsamples in sampling theory (Gumbel 1966). It has been used in a study of first-ranked galaxies in rich clusters and loose groups by Bhavsar \& Barrow (1985) to determine whether their first-ranked member galaxies have properties consistent with having been selected from a single underlying population. If independent subsamples are taken from an underlying probability distribution possessing an exponential tail then the probability distribution of the first-ranked members of the independent subsamples will approach the universal Gumbel form in the limit that the population becomes large.

Table 1. Means and variances of the probability distributions studied in Section 3.

| Distribution | Mean, $\mu$ | Variance, $\Sigma^{2}$ |
| :--- | :--- | :--- |
| Gaussian | 0 | $\sigma^{2}$ |
| Rayleigh | $\sigma \sqrt{\pi / 2}$ | $(2-\pi / 2) \sigma^{2}$ |
| Maxwell | $\sigma \sqrt{8 / \pi}$ | $(3-8 / \pi) \sigma^{2}$ |
| Chi-squared $\left(\chi_{n}^{2}\right)$ | $n \sigma^{2}$ | $2 n \sigma^{4}$ |
| Lognormal | $\exp \left(\frac{1}{2} \sigma^{2}\right)$ | $\exp \left(\sigma^{2}\right) \times\left[\exp \left(\sigma^{2}\right)-1\right]$ |
| Rectangular | $\left(1+\sigma^{2}\right)^{-1}$ | $\sigma^{4}\left(1+2 \sigma^{2}\right)^{-1}\left(1+\sigma^{2}\right)^{-2}$ |
| Gumbel type I | $\gamma^{\star}-\log \left(\sigma^{2}\right)$ | $\pi^{2} / 6$ |

${ }^{\star} \gamma=0.577 \ldots$ is Euler's constant.
which also has a rectangular marginal distribution on $[0,1]$. However, it is a different process and leads to intractable integrals.

It is probable that our method could be extended to several other more complicated nonGaussian processes like Weibull, Student- $t$, Fisher- $Z$ and Snedecor- $F$.


Figure 2. Probability densities for (i) Rayleigh and Maxwell; (ii) Chi-squared with $n=2,5$ and 30 ; (iii) lognormal, and (iv) Gumbel-I distributions.

We have concentrated attention upon only two statistical properties of the extrema of the stochastic fields (i)-(vii) above because these are both mathematically tractable and physically relevant in astronomical applications. However, in the case of Gaussian fields it is possible to go further and calculate the probability distribution function of the sizes of excursion regions above a threshold $u$ in the limit of large $u$ as (Belyaev \& Nosko 1969)
$\operatorname{Lt}_{u \rightarrow \infty} P\left[u \theta_{u} \geqslant x\right]=\exp \left[-\frac{\sigma_{1}^{2} x^{2}}{8}\right]$.
If we specialize results of Longuet-Higgins (1957) to the case of isotropic Gaussian fields then the mean length of contour per unit area of the field reduces to
$\langle l\rangle=\frac{\sigma_{1}}{2 \sigma} \exp \left[-\frac{u^{2}}{2 \sigma^{2}}\right]$
where $\sigma^{2}$ is the variance and $\sigma_{1}$ is defined by equation (4.1). It does not appear to be possible to produce results analogous to (3.49) for the non-Gaussian fields we have studied.

## 4 A comparison of Gaussian and non-Gaussian behaviour

Before we can compare meaningfully the statistics of Gaussian and non-Gaussian excursions we need to standardize the distributions (i)-(vii). This we do in two ways:
(a) We determine number densities and areas of two-dimensional fields above a standard level $u$ defined to be a fixed number of rms units from the mean. Alternatively, (and more realistically if we wish to use observations to distinguish between Gaussian and non-Gaussian underlying statistics using the high-level excursion behaviour), lacking a good estimate of the rms fluctuation, we could compare the behaviour at different percentile levels.
(b) We standardize the derivative part of the bivariate distribution. A measure of the 'noisiness' of a Gaussian process with covariance function $\xi(\theta)$ is provided by
$\sigma_{1}^{2}=-\xi^{\prime \prime}(0)=($ mean square derivative $)$.
Hence, we standardize by introducing a relative noisiness parameter, $\beta$, defined in general» by
$\beta=\frac{1}{2 \pi}\left(\frac{\text { mean square derivative }}{\text { variance }}\right)^{1 / 2}$
where the variance is given by $\xi(0)$.
The mean square derivatives and variances together with the form of $N_{u}$ in terms of $\beta$ for the examples given in Section 3 are listed in Table 2.

In Table 2 the noisiness of the Gumbel-I distribution and the consequent form of $N_{u}(\beta)$ are based upon a different definition of $\beta$. We note that, for a Gaussian process,
$\int_{-\infty}^{\infty} N_{u} d u=\frac{\sigma_{1}}{\sqrt{2 \pi}} \equiv N$
so by reference to this case we define a dimensionless noise parameter for the Gumbel-I as
$\beta_{*}=\frac{N}{\sqrt{2 \pi}} \frac{1}{(\text { variance })^{1 / 2}}$.

[^1]Table 2. Mean square derivatives, variances and computations of the expected number of excursions, $N_{u}$, above the level $u$ for the distributions listed in Table 1.

| Distribution | Mean square <br> derivative | Variance | Upcrossing rate <br> $N_{u}(\beta)$ |
| :--- | :--- | :--- | :--- |
| Gaussian | $\sigma_{1}^{2}$ | 1 | $\beta \exp \left(-u^{2} / 2\right)$ |
| Rayleigh | $\sigma_{1}^{2}$ | $2-\pi / 2$ | $[2 \pi(2-\pi / 2)]^{1 / 2} \beta u \exp \left(-u^{2} / 2\right)$ |
| Maxwell | $\sigma_{1}^{2}$ | $3-8 / \pi$ | $2(3-8 / \pi)^{1 / 2} \beta u^{2} \exp \left(-u^{2} / 2\right)$ |
| $\chi_{n}^{2}$ | $4 n \sigma_{1}^{2}$ | $2 n$ | $\beta \sqrt{2 / \pi}(u / 2)^{(n-1) / 2}[\exp (-u / 2)] /(\Gamma(n / 2)$ |
| Lognormal | $\mathrm{e} \sigma_{1}^{2}$ | $\mathrm{e}(\mathrm{e}-1)$ | $\beta(\mathrm{e}-1)^{1 / 2} \exp \left[-(\log u)^{2} / 2\right]$ |
| Rectangular | $\sigma_{1}^{2} / 3$ | $1 / 12$ | $\beta u(\pi / 2)^{1 / 2}$ |
| Gumbel type $\mathrm{I}^{\star}$ | - | $\pi^{2} / 6$ | $\left(\beta_{*} \pi / \sqrt{3}\right) \exp [-(u / 2)-\exp (-u)]$ |

$\star \beta_{*}$ is defined by equation (4.4) and given by equation (4.5) for the Gumbel-I distribution.

For Gumbel-I, $N=\sigma_{1}$ and so
$\beta_{*}($ Gumbel $)=\frac{\sigma_{1} \sqrt{3}}{\pi^{3 / 2}}$.
Hence, for Gumbel-I, equation (3.40) becomes

$$
\begin{equation*}
N_{u}=\beta_{*} \frac{\pi}{\sqrt{3}} \exp \left[-\frac{u}{2}-\exp (-u)\right] \tag{4.6}
\end{equation*}
$$

## 5 Application to the cosmic microwave background

In this section we shall apply the results obtained in Section 4 to the temperature distribution of the microwave background radiation over the celestial sphere. This problem has been studied, under the assumption that the angular fluctuations, $\Delta T(\theta)$, in the measured temperature of the radiation constitute a Gaussian random process, by Sazhin (1985a, b), Zabotin \& Nasel'skii (1985) and Bond \& Efstathiou (1987). Our computations generalize these earlier results to the case where the temperature fluctuations are described by a stochastic field with non-Gaussian statistics.

As usual we shall assume that the temperature fluctuations are described by a two-dimensional stochastic process on a plane of area $4 \pi$ rather than on a sphere. In order to compute the variance, $\Sigma^{2}$, and noise parameter, $\beta$, of the process we would need to know the power spectrum of the density inhomogeneities in the underlying cosmological model.

The results obtained are displayed in Tables 3, 4, 5 and 6 and in Figs 3 and 4. The case of a Gaussian field is included for comparison. In order to obtain an equivalence with the results of Sazhin (1985a, b) one requires a noisiness parameter $\beta$ of roughly 12 , which, in their notation produces for the Gaussian case $\sigma_{\mathrm{v}}^{2} / \sigma^{2}=4 \pi^{2} \beta^{2} \simeq 6000$. This gives a $\beta$ which is measured in units of $\mathbf{s r}^{-1}$. Our threshold level, $u$, is labelled by $v=-5,-4, \ldots, 4,5$ where
$u=\mu+v \Sigma$
and $\mu$ and $\Sigma^{2}$ are the mean and variance of the distribution in question (see Table 1). Our $v$ corresponds to that defined by Bond \& Efstathiou (1987) in the Gaussian case. Only in the Gaussian case are the results for the number and size distributions of the excursion symmetric under change of sign of $v$. In Figs 3 and 4 and Tables 3-6 there is a symmetry between the statistics

Table 3. Expected sizes of excursion regions above level $u$ or below level - $u$ (defined in terms of rms from the mean by equation 5.1) in units of $10^{-4} \mathrm{sr}$. The noisiness (see equations 4.2 and 4.5 ) has been taken to have the value $\beta=12$ in the examples shown.

|  | -4 | -3 | -2 | -1 | +1 | +2 | +3 | +4 | +5 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Gaussian | 0.7 | 1.0 | 2.0 | 4.8 | 4.8 | 2.0 | 1.0 | 0.7 | - |
| Rayleigh | - | - | - | 2.8 | 7.1 | 3.9 | 2.5 | 2.0 | - |
| Maxwell | - | - | 0.2 | 3.3 | 6.5 | 3.4 | 2.2 | 1.8 | - |
| Lognormal | - | - | - | - | 2.1 | 1.4 | 1.0 | 0.9 | 0.9 |
| Rectangular | - | - | - | 44.2 | 11.1 | - | - | - | - |
| Gumbel-I | - | - | 3.1 | 10.4 | 3.8 | 1.0 | 0.3 | 0.07 | 0.02 |
| $X_{n}^{2}, \mathrm{n}=1$ | - | - | - | - | 5.6 | 4.1 | 3.2 | 2.6 | 2.3 |
| 2 | - | - | - | - | 5.5 | 3.7 | 2.8 | 2.2 | 1.9 |
| 3 | - | - | - | 1.7 | 5.5 | 3.5 | 2.5 | 2.0 | 1.6 |
| 4 | - | - | - | 2.4 | 5.4 | 3.3 | 2.4 | 1.9 | 1.6 |
| 5 | - | - | - | 2.9 | 5.4 | 3.2 | 2.3 | 1.7 | 1.5 |
| 10 | - | - | 0.2 | 3.7 | 5.3 | 2.9 | 2.0 | 1.5 | 0.9 |
| 20 | - | - | 0.4 | 4.1 | 5.1 | 2.7 | 1.7 | 1.2 | 0.6 |
| 30 | - | - | 1.1 | 4.2 | 5.1 | 2.6 | 1.6 | 1.1 | 0.4 |

Table 4. Expected numbers of hot (cold) spots on the microwave sky above (below) the levels $u$ $(-u)$ in units of the rms from the mean defined by equation (5.1). The noisiness (see equations 4.2 and 4.5) has been taken to have the value $\beta=12$ in the examples shown.

|  | $v$ |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | -4 | -3 | -2 | -1 | +1 | +2 | +3 | +4 | +5 |
| Gaussian | 6 | 165 | 1457 | 4196 | 4196 | 1457 | 165 | 6 | - |
| Rayleigh | - | - | - | 7454 | 2877 | 1200 | 285 | 40 | - |
| Maxwell | - | - | 2961 | 6236 | 3133 | 1248 | 261 | 28 | - |
| Lognormal | - | - | - | - | 5741 | 2802 | 2131 | 1368 | 926 |
| Rectangular | - | - | - | 601 | 8367 | - | - | - | - |
| Gumbel-I | - | - | 28 | 1591 | 4708 | 5573 | 5848 | 5923 | 5945 |
| $x_{n}^{2}, n=1$ | - | - | - | - | 2692 | 1561 | 870 | 470 | 251 |
| 2 | - | - | - | - | 3079 | 1698 | 833 | 386 | 167 |
| 3 | - | - | - | 6852 | 3261 | 1742 | 785 | 321 | 125 |
| 4 | - | - | - | 6043 | 3374 | 1757 | 741 | 278 | 94 |
| 5 | - | - | - | 5664 | 3451 | 1763 | 707 | 248 | 77 |
| 10 | - | - | 111 | 5030 | 3651 | 1751 | 593 | 161 | 44 |
| 20 | - | - | 665 | 4727 | 3797 | 1713 | 488 | 102 | 22 |
| 30 | - | - | 869 | 4610 | 3865 | 1690 | 433 | 78 | 7 |
| 50 | - | - | 1050 | 4505 | 3937 | 1654 | 377 | 62 | - |

Table 5. Expected sizes of hotspots at different percentile levels in units of $10^{-4} \mathrm{sr}$. The noisiness (see equations 4.2 and 4.5) has been taken to have the value $B=12$ in the examples shown.

|  | Percentile |  |  |  |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $75 \%$ | $90 \%$ | $95 \%$ | $97.5 \%$ | $99 \%$ | $99.5 \%$ | $99.9 \%$ | $99.95 \%$ | $99.99 \%$ |
| Gaussian | 6.8 | 3.6 | 2.6 | 2.0 | 1.6 | 1.3 | 1.0 | 0.9 | 0.7 |
| Rayleigh | 9.2 | 5.6 | 4.3 | 3.5 | 2.8 | 2.4 | 1.9 | 1.7 | 1.4 |
| Maxwell | 8.6 | 5.1 | 3.9 | 3.1 | 2.5 | 2.2 | 1.7 | 1.5 | 1.4 |
| Lognormal | 4.0 | 2.1 | 1.5 | 1.2 | 0.9 | 0.7 | 0.5 | 0.5 | 0.3 |
| Rectangular | 4.9 | 0.5 | 0.1 | 0.03 | 0.005 | 0.001 | $4 \times 10^{-5}$ | $1 \times 10^{-5}$ | $4 \times 10^{-7}$ |
| Gumbel I | 8.2 | 2.5 | 1.1 | 0.5 | 0.2 | 0.1 | 0.02 | 0.01 | 0.002 |
| $X_{n}^{2}, \mathrm{n}=1$ | 8.2 | 5.2 | 4.0 | 3.3 | 2.6 | 2.3 | 1.7 | 1.6 | 1.3 |
| 2 | 8.0 | 4.8 | 3.7 | 3.0 | 2.4 | 2.1 | 1.6 | 1.5 | 1.2 |
| 3 | 7.8 | 4.6 | 3.5 | 2.9 | 2.3 | 2.0 | 1.5 | 1.3 | 1.1 |
| 4 | 7.7 | 4.5 | 3.4 | 2.8 | 2.2 | 1.9 | 1.5 | 1.3 | 1.1 |
| 5 | 7.6 | 4.4 | 3.3 | 2.7 | 2.2 | 1.9 | 1.4 | 1.2 | 1.1 |
| 10 | 7.5 | 4.2 | 3.1 | 2.5 | 2.0 | 1.7 | 1.3 | 1.1 | 1.0 |
| 20 | 7.3 | 4.0 | 3.0 | 2.4 | 1.9 | 1.6 | 1.2 | 1.1 | 0.9 |
| 30 | 7.2 | 4.0 | 2.9 | 2.3 | 1.8 | 1.6 | 1.2 | 1.1 | 0.9 |
| 50 | 7.1 | 3.9 | 2.9 | 2.3 | 1.8 | 1.5 | 1.2 | 1.0 | 0.9 |
| 100 | 7.0 | 3.8 | 2.8 | 2.2 | 1.7 | 1.5 | 1.1 | 1.0 | 0.8 |

of upcrossings above a level $u$ and downcrossings below a level $-u$ only in the Gaussian case. However, in all the non-Gaussian examples it is possible to carry out a transformation of $u \rightarrow-u$ to interchange the results for the high and low level crossings.

Looking first at the results for Rayleigh, Maxwell and $\chi^{2}$, it is clear that although the numerical values for the numbers and areas of hot spots are different in these cases, the form of the curves are rather similar. As expected, higher order $\chi^{2}$ processes give results close to the Gaussian [this is expected because the form of equation (3.17) means, that with higher $n$, processes will tend to the Gaussian by virtue of the Central Limit Theorem]. Indeed, asymptotically these processes all have an expected hotspot area which goes as $1 / u^{2}$. At this point one might conjecture that this similarity in high level excursion behaviour is related to the fact that these distributions are all of exponential type in the sense of Gnedenko (1941) (see also Gumbel 1966).夫 However, the Gumbel distribution is itself of exponential type but has an expected hotspot area that falls exponentially as $u \rightarrow \infty$. However, we conjecture that bivariate distributions, both of whose marginal distributions are of exponential type, will have the $1 / u^{2}$ high-level behaviour.

The lognormal distribution is extremely positively skew and the hotspot behaviour therefore exhibits a long positive tail. Note that, apart from a constant factor the behaviour of this process at different percentile levels is exactly the same as that of the Gaussian process. This would be true for any process which is a unique function of a single Gaussian process.
*A distribution is defined to be of exponential type if its probability density function $F^{\prime}(x)$ satisfies
$\operatorname{Lt}_{x \rightarrow \infty} \frac{d}{d x}\left(\frac{1-F(x)}{F^{\prime}(x)}\right)=0$
so all statistical moments must exist and the distribution must be unbounded either to the left or to the right.






Figure 4. Plots (solid lines) of the expected number of excursions above the level $u=\mu+v \Sigma$ over a plane microwave 'sky' of area $4 \pi$ steradians for the following distributions: (i) Rayleigh; (ii) Maxwell; (iii) lognormal; (iv) Gumbel-I; (v) Chi-squared $n=2$; (vi) Chi-squared $n=5$; (vii) Chi-squared $n=30$; (viii) Rectangular. In each case the result for


Table 6. Expected numbers of hotspots occurring at different percentile levels. The noisiness (see equations 4.2 and 4.5) has been taken to have the value $\beta=12$ in the examples shown.

|  |  |  |  | Per | entile |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 75\% | 90\% | 95\% | 97.5\% | 99\% | 99.5\% | 99.9\% | 99.95\% | 99.99\% |
| Gaussian | 4592 | 3502 | 2418 | 1553 | 807 | 474 | 129 | 72 | 18 |
| Rayleigh | 4948 | 2247 | 1462 | 890 | 449 | 258 | 67 | 37 | 9 |
| Maxwell | 3649 | 2464 | 1627 | 1008 | 503 | 288 | 74 | 42 | 9 |
| Lognormal | 7890 | 6018 | 4155 | 2668 | 1387 | 815 | 222 | 124 | 31. |
| Rectangular | 6396 | $2.3 \times 10^{4}$ | $5.13 \times 1.0^{4}$ | $1.08 \times 10^{5}$ | $2.79 \times 10^{5}$ | $5.63 \times 10^{5}$ | $2.84 \times 10^{6}$ | $5.68 \times 10^{6}$ | $2.84 \times 10^{7}$ |
| Gumbel-I | 3853 | 5081 | 5512 | 5731 | 5863 | 5909 | 5944 | 5949 | 5952 |
| $x_{n}^{2}, n=1$ | 3854 | 2419 | 1554 | 952 | 476 | 274 | 72 | 40 | 10 |
| 2 | 3940 | 2619 | 1703 | 1049 | 524 | 300 | 79 | 43 | 10 |
| 3 | 4015 | 2727 | 1785 | 1102 | 551 | 317 | 83 | 45 | 11 |
| 4 | 4069 | 2799 | 1840 | 1138 | 570 | 328 | 85 | 47 | 11 |
| 5 | 4110 | 2852 | 1880 | 1166 | 584 | 337 | 88 | 48 | 12 |
| 10 | 4181 | 2999 | 1995 | 1244 | 628 | 364 | 95 | 52 | 13 |
| 20 | 4315 | 3120 | 2093 | 1313 | 665 | 385 | 101 | 56 | 14 |
| 30 | 4360 | 3178 | 2142 | 1349 | 684 | 397 | 105 | 58 | 14 |
| 50 | 4407 | 3242 | 2194 | 1386 | 707 | 411 | 1.09 | 60 | 15 |
| 100 | 4458 | 3310 | 2253 | 1427 | 732 | 427 | 114 | 63 | 15 |

In the case of the rectangular distribution, although the distribution of the fluctuations is symmetric about the mean, the upcrossing rate is not symmetric and the characteristics of the hotspots are therefore different to those of the coldspots. Note also that all of the hotspots and coldspots are confined within a narrow range of the mean because the rectangular distribution is bounded to the left and to the right. (The previous cases have probability distributions which are bounded only to the left.)

The Gumbel-I example demonstrates a very peculiar behaviour which is unlike the other cases. In all previous examples, as the level $v$ increases so both the number of hotspots and their mean areas decrease. In the Gumbel-I case the number of hotspots above the level $v$ tends to a constant as $v \rightarrow \infty$ although the mean area of the regions becomes very small so that the total flux received from these regions goes to zero. Although this example is not physically very realistic it does demonstrate that one can use our method to treat fields which differ markedly from the Gaussian case.

In all cases it can be seen that greater differences in behaviour are displayed at levels $u$ specified in terms of standard deviations from the mean rather than at levels specified by percentiles. This suggests that good estimates of $\mu$ and $\Sigma$ would improve the prospects for using these properties as discriminators between Gaussian and non-Gaussian fluctuations.

Our results are not immediately comparable with those given by Bond \& Efstathiou (1987) because their calculations refer to local maxima and not just regions above some threshold level. One region above a level $u$ might well include several local maxima above the level. To apply our method to the form of fluctuations studied by Bond \& Efstathiou (adiabatic and isocurvature modes of cold dark matter) one would have to choose a different value of $\beta$ to that used above.

Their notation ( $\gamma$ and $\theta_{*}$ ) is related to our $\beta$ by
$4 \pi^{2} \beta^{2}=\gamma^{2} / \theta_{*}^{2}$
and $\beta^{2}$ is then in units of $\operatorname{arcmin}^{-2}$. This corresponds to values which are somewhat larger than studied by Sazhin. ( $\beta$ goes from 40.8 to 49.4 for the adiabatic modes and $\beta=10.2$ for isocurvature modes when measured in the units used by Sazhin.) Our results for the Gaussian case are in agreement with those of Bond \& Efstathiou (1987) and Vittorio \& Juszkiewicz (1987) but differ by a factor of 2 from those in Sazhin (1985a, b) due to a numerical error in the latter papers.

## 6 Discussion

We have developed a simplified method for calculating the expected sizes and numbers of spots of high and low temperature in the cosmic microwave background when its temperature variations are non-Gaussian. The non-Gaussian fields we have studied range from those which just depart slightly from the Gaussian case (e.g. $\chi_{n}^{2}$ with large $n$ ) to those where the behaviour is markedly different (e.g. rectangular and Gumbel-I). This is significant of the fact that our method can be used to treat a very wide class of stochastic fields derivable from the Gaussian.

Our results indicate that small departures from the Gaussian can lead to processes where the high-level behaviour is only slightly different from the Gaussian case. This has been known experimentally for some time in the case of clustering of density perturbations in three dimensions where the linear theory (which assumes Gaussian behaviour) is known to be fairly accurate even when the rms density fluctuation is of order the mean density, so indicating that the distribution of fluctuations is asymmetric about the mean and therefore non-Gaussian.

We consider the statistical properties of the high-level and low-level regions to be of very great importance when analysing any observed temperature fluctuations in the future. The first fluctuations to be seen will be from those areas where the microwave background fluctuations exceeds some level imposed by system noise in the antenna and receiver of the experimental system being used to measure the sky temperature. Good independent measurements of the mean level and rms deviation would improve the chances of being able to discriminate between Gaussian and non-Gaussian fluctuations using these high and low-level properties of the microwave background on small angular scales. But, until it is possible to reduce the system noise level very much below the rms fluctuations intrinsic to the sky, the temperature map obtained by the measuring device will be censored to some extent. One would be unable to test the complete distribution of fluctuations for goodness of fit to a Gaussian and so be unable to base any goodness of fit test on properties of the high or low level regions.

Although our results for the expected numbers and expected areas of hotspots and coldspots* serve as a useful guide to the behaviour of the processes under consideration, our method cannot be used to place confidence limits on these quantities as it does not enable us to calculate further details of the sampling distributions of these quantities. If we wish to do this using the high or low level behaviour we must simulate Gaussian and non-Gaussian temperature patterns and investigate by means of numerical experiments the power of the properties of high and low-level regions as discriminators between Gaussian and non-Gaussian fluctuations.
Notwithstanding these practical problems we can pose the question as to whether we could hope to test the hypothesis that hotspots above some temperature level possess statistics generated by a Gaussian field with some level of noisiness $\beta$ by using the mean spot size as a test statistic. Let us, for illustration, suppose that some set of $n$ hotspots generated by underlying

[^2]Gaussian fluctuations is observed and that this set constitutes a random sample from a population with a distribution of sizes given by equation (3.48). The size of the sample (which is determined theoretically by the noisiness level $\beta$, and, in practice, by the fraction of the sky covered by observations) and the distribution (3.48) together determine a confidence interval for the mean spot size and this can be used as the basis for a test of the above hypothesis against some alternative. The family of $\chi_{n}^{2}$ distributions provides a useful set of distributions to use as alternatives as the distributions range from very close to the Gaussian (when $n$ is large) to very different (when $n$ is small). We have studied the case when $\beta=12$ and found that it is not possible to reject the Gaussian hypotheses with 99 per cent confidence using hotspots above levels $v>3$ if the hotspots are actually generated by $\chi^{2}$ fluctuations of order $n>5$, even if a complete map of the microwave sky is available. In practice, we would clearly not have a map of the whole sky (due to point source subtraction, obscuration due to galactic emission etc.) and further limitations would be introduced by uncertainties in the estimates of the mean and rms level of the fluctuations so that even the example given above is extremely optimistic. Therefore, the mean spot size above a given temperature level will not be sensitive enough to discriminate between underlying Gaussian statistics and a whole set of other distributions which produce hotspots with similar characteristics for the same value of the noisiness parameter. Prospects are improved if $\beta$ is larger (as would be the case in fluctuations produced in the cold dark matter model which has lots of power on small scales and leads to $\beta \sim 40-50$ ) since the number of hotspots increases as $\beta^{2}$ and hence the associated confidence level for discrimination varies as $\beta^{-1}$. Even so, it is still unlikely that the mean size would be a useful way to discriminate between Gaussian and non-Gaussian fluctuations in the cold dark matter model even if large microwave sky maps were available. However, this does not exclude the possibility that there may exist some other statistical feature beside mean frequency and area of hotspots that can be used to discriminate different underlying statistics significantly.

We conclude that studies of more detailed properties of high level regions of non-Gaussian fields are needed in order to obtain a useful discriminator between Gaussian and non-Gaussian fluctuations. Such studies might also help to alleviate the problem of separating true temperature fluctuations from confusion produced by integrated point sources because such confusion would undoubtedly manifest non-Gaussian statistics.

## Acknowledgments

We thank Roman Juszkiewicz and John Peacock for suggesting improvements to the manuscript. P. Coles receives a SERC postgraduate studentship. J. D. Barrow acknowledges a Nuffield Foundation Fellowship.

## References

Abramowitz, M. \& Stegun, I. A., eds, 1965. Handbook of Mathematical Functions, Dover, New York. Adlér, R. J., 1981. The Geometry of Random Fields, John Wiley, New York. Bahcall, N. \& Soneira, R., 1983. Astrophys. J., 270, 20.
Bardeen, J. M., Bond, J. R., Kaiser, N. \& Szalay, A., 1986. Astrophys. J., 304, 15.
Barrow, J. D., Juszkiewicz, R. \& Sonoda, D. H., 1983. Nature, 305, 397.
Barrow, J. D., Juszkiewicz, R. \& Sonoda, D. H., 1985. Mon. Not. R. astr. Soc., 213, 917.
Belyaev, Yu. K. \& Nosko, V. P., 1969. Theory Prob. Appl., 14, 296.
Bhavsar, S. P. \& Barrow, J. D., 1985. Mon. Not. R. astr. Soc., 213, 857.
Bond, J. R. \& Efstathiou, G., 1987. Mon. Not. R. astr. Soc., 226, 655.
Coles, P., 1986. Mon. Not. R. astr. Soc., 222, 9p.
Couchman, H., 1986. Ph.D. thesis, University of Cambridge.
Couchman, H., 1987. Mon. Not. R. astr. Soc., 225, 777.
Couchman, H. \& Rees, M., 1986. Mon. Not. R. astr. Soc., 221, 53.

Davis, M., Efstathiou, G., Frenk, C. S. \& White, S. D. M., 1985. Astrophys. J., 292, 371.
Dekel, A. \& Silk, J., 1986. Astrophys. J., 303, 39.
Doroshkevich, A. G., 1970. Astrophysics, 6, 581.
Gnedenko, B. V., 1941. Ann. Math., 44, 423.
Gumbel, E. J., 1966. Statistics of Extremes, Columbia University Press, New York.
Hasofer, A. M., 1974. In: Studies in Probability and Statistics, pp. 153-9, ed. Williams, E. J., Jerusalem Academic Press, Jerusalem.
Jensen, L. G. \& Szalay, A. S., 1986. Astrophys. J., 305, L5.
Kaiser, N., 1984. Astrophys. J., 284, L9.
Klypin, A. A. \& Kopylov, A. I., 1983. Soviet Astr. Lett., 9, 75.
Ling, E. N., Frenk, C. S. \& Barrow, J. D., 1986. Mon. Not. R. astr. Soc., 223, 21p.
Longuet-Higgins, M. S., 1957. Phil. Trans. R. Soc. Lond. A., 249, 321.
Marochnik, L. S., Nasel'skii, P. D. \& Zabotin, N. A., 1980. AN SSSR preprint No. 564.
Novikov, I. D., 1968. Soviet Astr., 12, 427.
Otto, S., Politzer, H. D. \& Wise, M. B., 1986. Phys. Rev. Lett., 56, 1878.
Peacock, J. A. \& Heavens, A. F., 1985. Mon. Not. R. astr. Soc., 217, 805.
Politzer, H. D. \& Wise, M. B., 1984. Astrophys. J., 285, L1.
Rees, M. J., 1985. Mon. Not. R. astr. Soc., 213, 75p.
Sazhin, M. V., 1985a. Mon. Not. R. astr. Soc., 216, 25p.
Sazhin, M. V., 1985b. Soviet Astr. Lett., 11, 239.
Sharpe, K., 1978. Adv. Appl. Prob., 10, 373.
Silk, J., 1985. Astrophys. J., 297, 1.
Sunyaev, R., 1972. Astr. Astrophys., 12, 190.
Vanmarcke, E. H., 1983. Random Fields: Analysis and Synthesis, MIT Press, Cambridge, Massachusetts.
Vittorio, N. \& Juszkiewicz, R., 1987. Astrophys. J., 314, L29.
Wax, N., ed., 1954. Selected Papers on Nosie and Stochastic Processes, Dover, New York.
Zabotin, N. A. \& Nasel'skii, P. D., 1985. Soviet Astr., 29, 614.


[^0]:    * Our definition of $\sigma_{1}^{2}$ differs from that of Bond \& Efstathiou (1987) and Bardeen et al. (1986) and equals one-half and one-third of their values in the two- and three-dimensional cases respectively.

[^1]:    *The Gumbel-I distribution is exceptional and cannot be treated in this manner because the mean square derivative does not exist. A different normalization is proposed for this distribution below.

[^2]:    *The hotspots referred to here are not related to the hotspots of large angular scale which arise in open (negative spatial curvature) anisotropic universes as a result of geodesic focussing effects (Novikov 1968; Barrow, Juszkiewicz \& Sonoda 1983, 1985).

