Fermi National Accelerator Laboratory

Fermilab-Pub-94-043-A

Non-Gaussian statistics of pencil beam surveys

Luca Amendola

March 16, 1994

Osservatorio Astronomico di Roma
Viale del Parco Mellini, 84
00136 Rome - Italy
and
NASA/Fermilab Astrophysics Center [a]
Fermi National Accelerator Laboratory
PO Box 500
Batavia IL 60510 - USA

Abstract

We study the effect of the non-Gaussian clustering of galaxies on the statistics of pencil beam surveys. We find that the higher order moments of the galaxy distribution play a dominant role in the probability distribution for the power spectrum peaks. Taking into account the observed values of the moments of galaxy distribution we derive the probability distribution for the power spectrum modes in non-Gaussian models and show that the probability to obtain the $128h^{-1}$ Mpc periodicity found in pencil beam surveys is raised by more than one order of magnitude, up to 1%. Further data are needed to decide if non-Gaussianity alone is sufficient to explain the $128h^{-1}$ Mpc periodicity, or if extra large scale power is necessary.

Subject headings: cosmology: large-scale structure of the Universe; galaxies: clustering

[a] Address until Oct. 1994; e-mail amendola@fnas06.fnal.gov

Operated by Universities Research Association Inc. under contract with the United States Department of Energy

1 Introduction

The surprising discovery of a $128h^{-1}$ Mpc periodicity in the distribution of galaxies (Broadhurst *et al.* 1989) has raised an intense debate about the statistical significance of the signal detected. The main question is whether the periodicity is consistent with the local observations or is rather to be regarded as a new feature appearing only when very large scales ($\gg 100h^{-1}$ Mpc) are probed. In the original paper by Broadhurst *et al.* (1989; BEKS) the statistical significance of the peak in the one-dimensional power spectrum was assessed making use of an external estimator, i.e. adopting a model for the clustering of galaxies. The clustering was assumed to be described by the usual correlation function $\xi(r) = (r/r_0)^{-\gamma}$ up to the scale of $30h^{-1}$ Mpc, without any correlation beyond this scale, and without any higher order moment. As Szalay *et al.* (1991) pointed out, however, external estimators are very model dependent. Even slightly different assumptions, concerning e.g. selection functions or the parameters r_0, γ, can result in dramatic variations of statistical significances. Indeed, Kaiser & Peacock (1991), investigating essentially the same dataset as BEKS, but including the effects of a realistic survey geometry, found that the mean power (or "noise") level was to be significatively raised. This resulted in a much higher probability to find a peak as large as, or larger, the one at $128h^{-1}$ Mpc, so as to reconcile the standard model of galaxy clustering with the BEKS data. The strong dependence of the results upon the clustering model seems to force one to use an internal estimator of the noise level, i.e. an estimator which depends only on the data and on their expected statistics. This has been done by Szalay *et al.* (1991), who showed then that the probability to find a peak as high as the one in the BEKS data, or higher, is $2.2 \cdot 10^{-4}$, matching the original estimate. Luo & Vishniac (1993) also confirmed the result that, while the rest of the power spectrum is consistent with the hypotheses of clustering and Gaussianity, the single prominent spike at $128h^{-1}$ Mpc is not. They also showed that even a delta-like feature in the tridimensional power spectrum of the galaxy distribution can barely account for the BEKS spike. As we will show below, the probability estimate on which these conclusions are based relies essentially on two hypotheses: *a*) that the spatial bins of the BEKS survey are uncorrelated, i.e. that the clustering beyond $30h^{-1}$ Mpc is negligible, and *b*) that the components of the power spectrum can be assumed, by virtue of the central limit theorem, to be Gaussian distributed. The very fact that the probability estimate based on these two hypotheses is as low as $2.2 \cdot 10^{-4}$ points to the conclusion that one of the two, or both, are false. This implies either that there is some previously unknown, and theoretically unexpected, feature in the tridimensional power spectrum at large scale, or that it is the other hypothesis, the Gaussianity, to be abandoned. The first possibility has been explored for instance in Voronoi simulations (see e.g. Coles 1990, SubbaRao & Szalay 1992), or in truncated HDM models (Weiss & Buchert 1993). Unlike the precedent studies, in this paper we consider in detail the latter way out.

The scheme of this paper is as follows. First, we derive the probability distribution of the components of a one-dimensional power spectrum in presence of higher order moments of the spatial distribution, making use of the Edgeworth expansion. Second, we ask ourselves which is the probability to find a spike as high as, or higher than, the one in the BEKS data in such non-Gaussian galaxy distribution. Finally, adopting the actual higher order moments found in local ($\leq 100h^{-1}$ Mpc) observations, we will show that the formal probability for the

BEKS periodicity increases by more than one order of magnitude, approaching 1%. Even a slight variation in the $128h^{-1}$ Mpc peak significance can bring this value up to several percent or down outside the three sigma threshold. As the data stand, our conclusion is that we cannot exclude that non-Gaussianity alone is sufficient to explain the data to a confidence level higher than 99% or so. The more interesting result is however that the non-Gaussianity of the power spectrum modes is decisive in estimating the likelihood of rare events.

Let us note that we will not question in any way the reliability of the BEKS data or of their noise estimate. Rather, we derive our conclusion only taking into due account the *already known* level of non-Gaussianity in the galaxy distribution.

2 Non-Gaussian pencil beam statistics

The BEKS data consist in a set of counts along a survey geometry that approximates a long, thin cylinder directed towards the galactic poles. The galaxy positions are binned in N small cylinders of radius $R = 3h^{-1}$ Mpc and radial length $30h^{-1}$ Mpc, out to $L/2 \sim 1000h^{-1}$ Mpc in both directions. The details of the survey are given in the original paper (BEKS) and in Szalay *et al.* (1991). Let us denote the cell counts as n_i, with $i = 1,..N \approx 70$. The discrete Fourier transform of the dataset is

$$f_k = \frac{1}{P}\sum_{j=1}^{N} n_j \exp(i2\pi k r_j/L)\,, \tag{1}$$

where $r_j = 30jh^{-1}$ Mpc is the radial distance to the $j-th$ bin, and $P = \sum n_j$ is the total number of galaxies (396 in BEKS). The counts n_j have mean $\hat{n} = P/N$ and variance $\sigma^2 =< (n_j - \hat{n})^2 >$ as well as higher order irreducible moments (or cumulants, or disconnected moments) k_n. Let us recall that $k_2 = \sigma^2\,,\quad k_3 =< (n_j - \hat{n})^3 >\,,\quad k_4 =< (n_j - \hat{n})^4 > -3k_2^2$, etc., and that for a Gaussian distribution $k_n = 0$ for $n > 2$. The power spectrum of the dataset n_j is defined as (for $k = 1, 2, ..., N/2$)

$$A_k = |f_k|^2\,. \tag{2}$$

Let us define the quantity

$$a_k = (\sigma^2 N/2)^{-1/2} \sum_j n_j \cos(2\pi k r_j/L)\,. \tag{3}$$

Squaring a_k we obtain $a_k^2 = (\,\mathrm{Re} f_k)^2/[\sigma^2(N/2P^2)]$. Likewise, we can define the quantity $b_k = \sum_j n_j \sin(2\pi k r_j/L)/\sigma\sqrt{N/2}$ and form the modulus

$$z \equiv a_k^2 + b_k^2 = \frac{A_k}{\sigma^2(N/2P^2)}\,. \tag{4}$$

The problem is now to find out the probability distribution density (PDD) of A_k when we know the one for n_j. First, however, we have to derive the PDD of a_k and b_k. They are constructed as a linear sum of N independent variables (as long as the various n_j are uncorrelated), so by the central limit theorem a_k, b_k should tend to be Gaussian distributed.

However, since $N \sim 70$ is not really very large, one should check if the higher order terms are significant. This is indeed what will be shown to happen. We make use of the so-called Edgeworth expansion (see e.g. Cramer 1966, Abramovicz & Stegun 1972, whose notation we will follow), according to which the variable

$$x = \frac{\sum_i (Y_i - m_i)}{(\sum_i \sigma_i^2)^{1/2}} \tag{5}$$

(the sums run over N terms) where the Y_i are independent random variables with mean m_i, variance σ_i and n-th order cumulants $k_{n,i}$, is distributed like a function $f(x)$ that can be expanded in powers of $N^{-1/2}$ around a normal distribution. To the order $1/N$ the Edgeworth expansion is

$$f(x) \sim G(x)\left[1 + \frac{\gamma_1}{6N^{1/2}}\mathrm{He}_3(x) + \frac{\gamma_2}{24N}\mathrm{He}_4(x) + \frac{\gamma_1^2}{72N}\mathrm{He}_6(x) + O(N^{-3/2})\right]. \tag{6}$$

Here, $G(x)$ is the normal distribution, $\mathrm{He}_n(x) \equiv (-1)^n (d^n G(x)/dx^n)/G(x)$ is the Hermite polynomial of order n, and the expansion coefficients are given by the general formula (Abramovitz & Stegun 1972)

$$\gamma_{r-2} = \frac{(\sum_i k_{r,i}/N)}{(\sum_i \sigma_i^2/N)^{r/2}}. \tag{7}$$

In general, to the order $1/N^m$ we will need the $(2m+2)$-th order cumulant of Y_i. Notice that properly speaking the Edgeworth expansion is not a probability distribution, since it is not positive definite. However, what is relevant here is that the error one makes to the order N^{-m} is indeed of order $N^{-m-1/2}$ (Cramer 1966). The Edgeworth expansion has been used recently in astrophysics by several authors to quantify slight deviations from Gaussianity (Juszkiewicz *et al.* 1993, and references therein). In the form (6), the Edgeworth expansion can be considered as the mathematical expression of the central limit theorem for finite values of N. Now we can notice that, as long as the counts n_j are uncorrelated, the variables a_k and b_k are indeed in the form (5), where $Y_i = (2/N)^{1/2} n_i \theta_i$ and $\theta_i = \cos(2\pi k r_i/L)$ or $\sin(2\pi k r_i/L)$, and where $\sum m_i = 0$. This allows us to apply the Edgeworth expansion. To evaluate the expansion coefficients we need the sums

$$\sum_i k_{r,i} = (2/N)^{r/2} k_r \sum_i \theta_i^r, \tag{8}$$

where k_r is the r-th cumulant of the counts n_i, that is the observable quantity. For $r = 2$ the variance squared replaces k_2 in Eq. (8). The sum over the sines and cosines can be approximated as $\sum \theta_i^r = N < \cos^r(x) >$, where the average indicates the mean value of the function over a period. Then, the sum over the odd moments of the variables Y_i vanishes, so that $\gamma_{r-2} = 0$ for r odd. For the even moments we have that $\sum_i \sigma_i^2 = \sigma^2$, while the non vanishing coefficients γ_{r-2} are

$$\gamma_{r-2} = 2^{r/2} < \cos^r(x) > (k_r/\sigma^r). \tag{9}$$

Due to the symmetry of the Fourier modes, the expansion in powers of $N^{-1/2}$ becomes an expansion in powers of $1/N$.

Since we wish to include terms to the order $1/N^2$ in the Edgeworth expansion, we need the coefficients γ_i for $i = 2, 4$. As we have shown, this requires an estimate of the 4-th and 6-th order cumulants of the counts n_i. The higher order moments of the galaxy counts have been calculated for different surveys (Saunders *et al.* 1991; Bouchet, Davis & Strauss 1992; Gaztañaga 1992; Loveday *et al.* 1992). Gaztañaga (1993) gives the cumulants up to k_9 for the APM angular catalog, by converting the angular moments to the corresponding spatial ones. The general result is that, for counts on scales ranging from some megaparsecs to more than 50 h^{-1} Mpc, i.e. in the linear or mildly non-linear regime, the dimensionless cumulants $\mu_m = k_m/\hat{n}^m$ obey the hierarchical scaling relation

$$\mu_m = S_m \mu_2^{m-1}\,, \tag{10}$$

where S_m are the scaling constants (we have checked that the shot-noise correction is negligible in our case). Let us note that the count n_i on a scale $R \gg r_c$, if r_c is the correlation length of the fluctuation field, is a sum over $N \sim (R/r_c)^3$ independent random variables. Then, by applying the Edgeworth expansion to n_i, one can easily see that, to the lowest order in the variance, the scaling relation is a purely *mathematical consequence* of the Edgeworth expansion. On these scales, the physics of the clustering process is contained in the scaling constants S_m, and has been widely investigated in several works, from the book of Peebles (1980) to recent generalizations as in Bernardeau (1993).

Coming back to the Edgeworth expansion coefficients, we see that, from (9) and (10) ,

$$\gamma_{r-2} \approx A_r S_r \mu_2^{(r-2)/2}\,, \tag{11}$$

where the numerical factor $A_r = 2^{r/2} < \cos^r(x) >$ equals 3/2, 5/2 for $r = 4, 6$, respectively, and vanishes for r odd. The value of $\mu_2 = \sigma^2/\hat{n}^2$ can be expressed as a function of the correlation function (e.g. Peebles 1980), $\mu_2 = (\xi_0 + 1/\hat{n})$, where $\xi_0 = V^{-2} \int d^3r_1 d^3r_2 W(r_1)W(r_2)\xi(\mathbf{r}_1 - \mathbf{r}_2)$ and where W is the window function corresponding to the BEKS cylindrical cells of volume V. Since $\hat{n} \approx 6$ and ξ_0 is of order unity, we can approximate μ_2 with ξ_0, so that $\gamma_{r-2} \approx A_r S_r \xi_0^{(r-2)/2}$. The values of γ_2, γ_4 will result to be crucial. Several uncertainties, however, prevent their exact estimate. For ξ_0 we must rely on very local observations of the correlation function parameters γ and r_0. For $\gamma = 1.8$ and $r_0 = 5h^{-1}$ Mpc one gets $\xi_0 \approx 1$ (see e.g. Szalay *et al.* 1991), but this value depends on r_0 as r_0^γ. We can consider the range $\xi_0 \in (0.8 - 1.3)$ as an acceptable one. For S_4, S_6, one problem is that we need the scaling constants for quite elongated cylindrical cells, while the observations have been carried out mostly for large spherical or cubic cells. Since the scaling coefficients are indeed expected to be constant only for $r \gg r_0$ (e.g. Lahav *et al.* 1993, Colombi, Bouchet & Schaeffer 1994), we cannot simply extrapolate their large-scale values down to our cell size. For S_4, Gaztañaga (1993) reports from the APM catalog values ranging from 20 to 40, for spherical cells of radius $R = 10h^{-1}$ Mpc to $R = 1h^{-1}$ Mpc. For S_6, the range of values is even larger, from 10^3 to $9 \cdot 10^3$. In the mildly non-linear regime, these values are in agreement with the results of higher-order perturbation theory (Bernardeau 1993). One can naively assume for our cylindrical cells of volume $V \approx 850(h^{-1}\text{ Mpc})^3$ the value corresponding to a spherical cell of radius $R = (3V/4\pi)^{1/3} \approx 6h^{-1}$ Mpc. Then we obtain $S_4 \approx 25$ and $S_6 \approx 2000$. The reference values for the expansion coefficients can be then $\gamma_2 \approx 40$ and $\gamma_4 \approx 5000$. However, values as large as $\gamma_2 = 60$ and $\gamma_4 = 10^4$ can still be perfectly consistent with the observations. We will explore numerically all this range of values.

Let us apply now the Edgeworth expansion (6) to a_k or b_k. Due to the cancellation of the odd orders in $N^{-1/2}$, only the even Hermite polynomials are left in Eq. (6), so that $f(a_k)$ is an even function of a_k. The PDD for $y = a_k^2$ is then $P(y) = 2f(a_k)|da_k/dy| = f(y^{1/2})/y^{1/2}$, that is, to the order $(1/N^2)$,

$$
\begin{aligned}
P(y &= a_k^2) \sim g_1 P_1 + \frac{\gamma_2}{24N}\left[g_5 P_5 - 6 g_3 P_3 + 3 g_1 P_1\right] \\
&+ \frac{\gamma_4}{720N^2}\left[g_7 P_7 - 15 g_5 P_5 + 45 g_3 P_3 - 15 g_1 P_1\right] \\
&+ \frac{\gamma_2^2}{1152N^2}\left[g_9 P_9 - 28 g_7 P_7 + 210 g_5 P_5 - 420 g_3 P_3 + 105 g_1 P_1\right] \equiv \sum_j c_j P_j(y), \qquad (12)
\end{aligned}
$$

where $g_n \equiv 2^{n/2}\Gamma(n/2)$ and $P_n(y) = g_n^{-1} y^{n/2-1} e^{-y/2}$ is the χ^2 PDD with n degrees of freedom. Now that we have the PDD for a_k^2, b_k^2 we must find the distribution for $z = a_k^2 + b_k^2$. Let us denote with $\phi(t)$ the characteristic function (CF) of a generic probability distribution $P(x)$, where $\phi(t) = \int e^{itx} P(x) dx$. The general theorems about probability distributions say that the CF of the sum of two independent variables is the product of the CF of the variables. Furthermore, by linearity, we see that the CF of $P = P_1 + P_2$ is $\phi(P_1) + \phi(P_2)$. We are to use these two properties to derive the general distribution $P(A_k)$. First, we calculate the CF $\phi(P)$ for $P(y)$ given by (12), $P(y) = \sum_j c_j P_j$. Denoting the CF for the χ^2 distribution P_n as $\psi_n \equiv (1-2it)^{-n/2}$, we have

$$
\phi[z] = \phi[a_k^2]\phi[b_k^2] = \phi^2[y] = (\sum_j c_j \psi_j)^2\,, \qquad (13)
$$

where the sum runs over all the χ^2 PDD in the expansion (12), with the same c_j's. Now, since $\psi_n \psi_m = \psi_{n+m}$, we can see that the CF for the unknown distribution $P(z)$ is a sum of χ^2 CFs, so that the final result $P(z)$ is again a sum of χ^2 PDDs. Before writing down the result, we note that $z = A_k/[\sigma^2(N/2P^2)] = 2A_k/A_0$, where A_0 is the noise level in the notation of BEKS. It follows $A_0 = \sigma^2 N/P^2 = (\xi_0/N + 1/P)$, which gives an external estimate of the noise level. However, as already remarked, the estimate of A_0 by Szalay *et al.* (1991) is internal in that is not based on a *a priori* model for $\xi(r)$, but rather on fitting the observational distribution function for A_k *at small amplitudes* with the exponential $P(A_k) = (1/A_0)\exp(-A_k/A_0)$, as it should be in the purely Gaussian case (or for $N \to \infty$). The same internal estimate applies here, since as we will see the Gaussian and non-Gaussian PDD are equivalent at low amplitudes.

Finally, the normalized distribution function for A_k to the order $1/N^2$ is

$$
\begin{aligned}
P(z = 2A_k/A_0) &= P_2 + a_1(P_6 - 2P_4 + P_2) + a_2(P_{10} - 4P_8 + 6P_6 - 4P_4 + P_2) \\
&+ a_3(P_8 - 3P_6 + 3P_4 - P_2)\,, \qquad (14)
\end{aligned}
$$

where $a_1 = \gamma_2/4N$, $a_2 = (19/96)\gamma_2^2/N^2$ and $a_3 = \gamma_4/24N^2$. Eq. (14) gives the general PDD for the power spectrum amplitudes relative to a set of pencil beam counts with scaling coefficients S_4, S_6. When $S_4 = S_6 = 0$ we return to the exponential distribution $P(z) = P_2$

on which the calculation of BEKS and of all the other works on the subject were based. As a digression, let us note that Eq. (14) can also be written in a very compact symbolic form, $P(z) = P_2[1 + \sum_{j=1}^{3} a_j(P_2 - 1)^{j+1}]$, in which the "products" of χ^2-functions P_n are to be defined such that $P_n \cdot P_m \equiv P_{n+m}$.

We can see from $P(z)$ why the higher order terms are important. Since the peak-to-noise ratio $X \equiv A_k/A_0$ found by BEKS is very large, $X_{BEKS} = 11.8$, the terms containing higher order χ^2 functions will dominate over the P_2 term when integrated to give the cumulative probability, even if the constants a_j are small, i.e. even if N is large. Actually, for any given N there is a value z_c such as the higher order terms dominate over the lower orders in the integral $\int_{z_c}^{+\infty} P(z)dz$. This is a consequence of the fact that, while the convergence of any distribution $f(x)$, where x is as in (5), to the normal one for $N \to \infty$ is ensured by the central limit theorem, the convergence itself needs not be uniform. The fractional difference between the cumulative distribution of $f(x)$ and the one relative to a normal distribution can be arbitrarily large for large deviations from the mean.

We can now directly compare the PDD (14) with the power spectrum coefficients found by BEKS. We use the tabulated values provided by Luo & Vishniac (1993), binned in peak-to-noise intervals of 0.5. We plot in Fig. 1 the cumulative function of the BEKS coefficients versus peak-to-noise ratio (a point at abscissa X represents the fraction of values of A_k in the BEKS data with peak-to-noise ratio larger than X) and compare this with our theoretical cumulative function

$$F(X) = \int_{2X}^{+\infty} P(z)dz\,, \tag{15}$$

where $X = A_k/A_0 = z/2$ is the peak-to-noise ratio. The functions plotted are for the Gaussian case ($\gamma_2 = \gamma_4 = 0$), and for three possible values of the constant γ_2: from bottom to top, $\gamma_2 = 20, 40, 60$, fixing $\gamma_4 = 5000$. It is clear that as γ_2 increases, the observed distribution becomes more and more consistent with the non-Gaussian behavior, except for the last point, the $128h^{-1}$ Mpc spike, which appears still far away from its expected frequency value. However, we can estimate now the probability to have $X_{BEKS} = 11.8$ or higher in one of the ~ 30 k-bins to which BEKS assigned the data (see Szalay *et al.* 1991 for a detailed exposition) and compare with the very unlikely value $2.2 \cdot 10^{-4}$ originally found for $\gamma_2 = \gamma_4 = 0$. The non-Gaussian result is

$$P(> 11.8) \approx 30F(11.8) = 0.003 - 0.01\,, \tag{16}$$

for the range $\gamma_2 = 20 - 60$. We obtain a record value $P(> 11.8) = 0.015$ assuming the extremal values $\gamma_2 = 60$ and $\gamma_4 = 10^4$. The inclusion of non-Gaussianity pushed the probability to obtain the BEKS spike by more than one order of magnitude, *without any need to invoke non standard features in the galaxy distribution.* The result (16) states that the BEKS spike should occur roughly in 0.3-1.0% of the cases if the very large scale galaxy distribution has to be consistent with the local observations of variance and kurtosis. The result can be interpreted in two complementary ways. If further data do not reduce the peak significativity, our result indicates that is difficult, though not impossible, for the non-Gaussianity alone to explain the observations. On the other hand, if the observed peak-to-noise ratio is somewhat reduced, the deep scales probed by the BEKS pencil beam can become consistent, at a two sigma level, with the local picture of galaxy clustering. In any case, it appears that the effect of non-Gaussianity cannot be neglected when one estimates the probability of rare events.

In Fig. 2 we display the behavior of $P(>11.8)$ *vs.* γ_2 for several values of γ_4. Assuming $\gamma_4 = 5000$, the BEKS spike is inside the three sigma levels for $\gamma_2 > 12$, and inside the 1% level for $\gamma_2 > 54$. If γ_4 is larger, the range is widened. A value of X_{BEKS} smaller by even a ten percent would result in quite higher values of $P(>11.8)$, as shown by the dot-dashed curve in Fig. 2. A probability of 3% can be then reached assuming γ_2, γ_4 in the acceptable range. Notice that for a purely Gaussian distribution the probability would still be outside the three sigma level.

3 Conclusions

We have shown that the higher order moments of the galaxy clustering play a dominant role in assessing the significance level of peaks in one-dimensional power spectra. The scaling constants S_4, S_6 and the correlation average ξ_0 on the spatial bin in a pencil beam survey are combined in the crucial parameters γ_2, γ_4. Assuming that the spatial bins are uncorrelated, the question raised by the remarkable periodicity discovered by Broadhurst *et al.* (1989) in the very large scale galaxy clustering can then be expressed in the following way: are the values of S_4, S_6 and of ξ_0 determined by local observations compatible with the clustering of galaxies at the very deep scales probed by the pencil beam? To give an answer, we have to determine the probability distribution for the power spectrum amplitudes of a non-Gaussian field sampled in spatial bins. We find by means of the Edgeworth expansion, truncated for sake of simplicity to the order N^{-2}, that the BEKS most prominent spike around $128h^{-1}$ Mpc has an occurrence probability of roughly 10^{-2} for acceptable values of γ_2, to be compared with the value $2.2 \cdot 10^{-4}$ obtained by Szalay *et al.* (1991) neglecting the higher order correction. If the BEKS periodicity is not smeared out by further observations, we should probably conclude that the spatial bins cannot be assumed uncorrelated. Indeed, this is what the large coherent structures reported in deep surveys seem to require. In any case, the non-Gaussian terms have to be taken into account when estimating the likelihood of rare events. It is also possible that further orders in the Edgeworth expansion make a non-negligible contribution to the expected probability: an estimate of this, however, would require a much more precise knowledge of cumulants to the eighth order and beyond than is available at present.

We also compared the peak occurrences of the full spectrum of BEKS and found it in a good agreement with our non-Gaussian probability distribution, especially if large values for γ_2 are allowed. This raises the possibility that further pencil beam data can measure the parameter γ_2, i.e. the product $S_4\xi_0$, down to very deep distances (the distribution being less sensible to the other parameter, γ_4). However, because of the central limit theorem, the power spectrum statistics is certainly not the best place to look for non-Gaussianity. This is even more true for the three-dimensional power spectrum $P_3(k)$: being it an average over k-directions, we expect that non-Gaussian deviations are negligible, except perhaps for very small wavenumbers.

Acknowledgments

I thank Stéphane Colombi for useful discussions. This work has been supported by DOE and NASA under grant NAGW-2381. I also acknowledge CNR (Italy) for financial support.

References

Bernardeau F. 1993, preprint

Bouchet F., Davis M., & Strauss M. 1992, in Proc. of the 2nd DAEC Meeting on the Distribution of Matter in the Universe, eds. G.A. Mamon, & D. Gerbal, 287

Broadhurst T.J., Ellis R.S., Koo D.C. & Szalay A. 1990 Nature 343, 726

Coles P. 1990 Nature 346, 446

Colombi S., Bouchet F.R., & Schaeffer R.,1994, Astron. Astrophys. 281, 301

Cramer H. 1966, Mathematical Methods of Statistics (Princeton: Princeton Univ. Press)

Gaztañaga E. 1992, Ap. J. , 398, L17

Gaztañaga E. 1993, Mon. Not. R. Ast. Soc. 000,000

Juszkiewicz R., & Bouchet F.R. 1992, in Proc. of the 2nd DAEC Meeting on the Distribution of Matter in the Universe, eds. G.A. Mamon, & D. Gerbal, 301

Juszkiewicz R., Weinberg D. H., Amsterdamski P., Chodorowski M. & Bouchet F. 1993, preprint (IANSS-AST 93/50)

Kaiser N. & Peacock J.A. 1991, Ap. J. 379 482

Lahav O., Itoh M, Inagaki S. & Suto Y. 1993, Ap. J. 402, 387

Loveday J., Efstathiou G., Peterson B. A. & Maddox S.J. 1992, Ap. J. , 400, L43

Luo S. & Vishniac E.T. 1993, Ap. J. 415, 450.

Peebles P.J.E. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton Univ. Press)

Saunders W., Frenk C., Rowan-Robinson M., Efstathiou G., Lawrence A., Kaiser N., Ellis R., Crawford J., Xia X.-Y., Parry I. 1991, Nature , 349, 32

SubbaRao M. & Szalay A. 1992, Ap. J. 391, 483

Szalay A.S., Ellis R.S., Koo D.C. & Broadhurst T.J. 1991, in Proc. After the First Three Minutes, ed. S. Holt, C. Bennett & V. Trimble (New York: AIP)

Weiss A. & Buchert T. 1993, Astron. Astrophys. 274,1

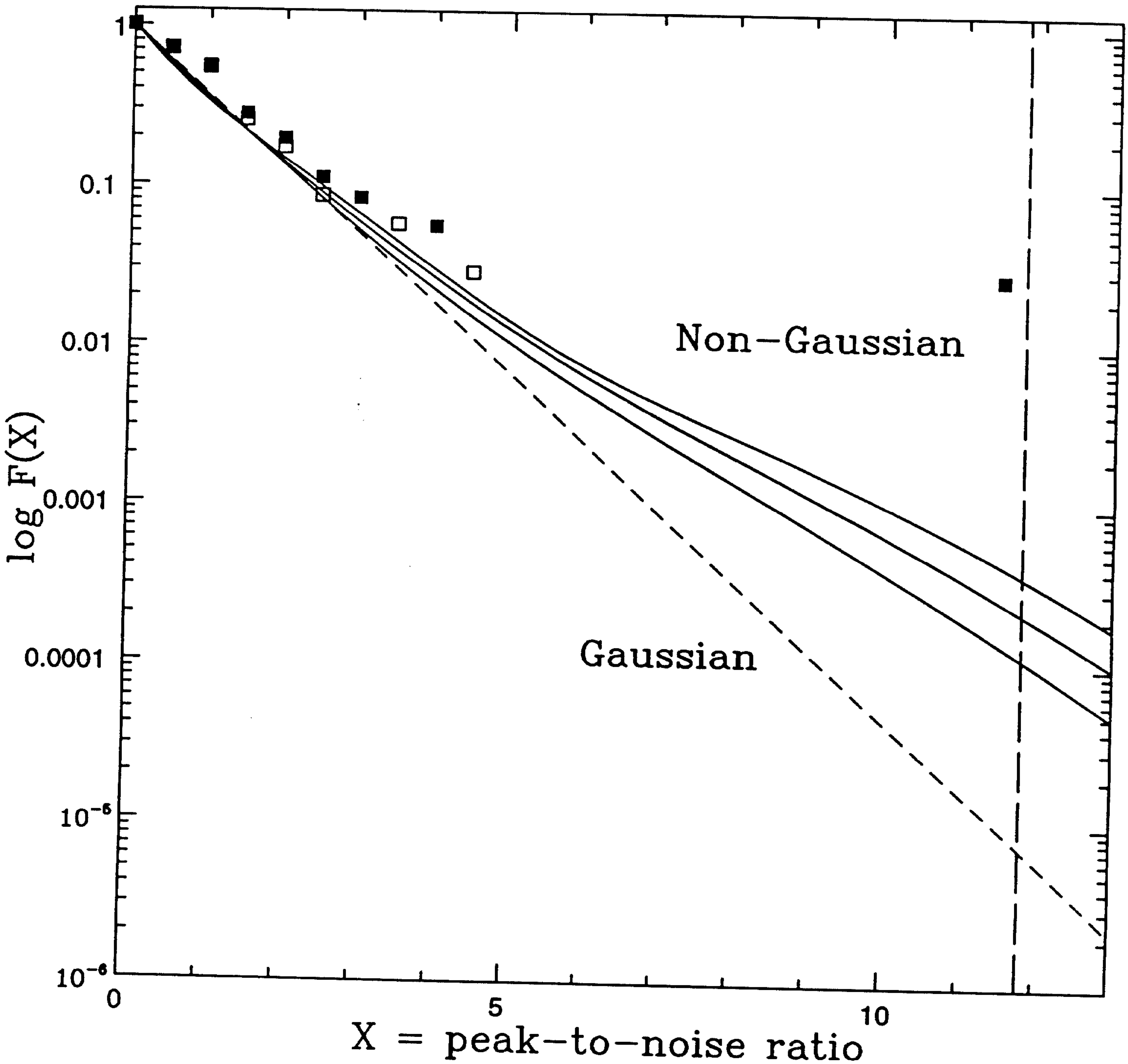
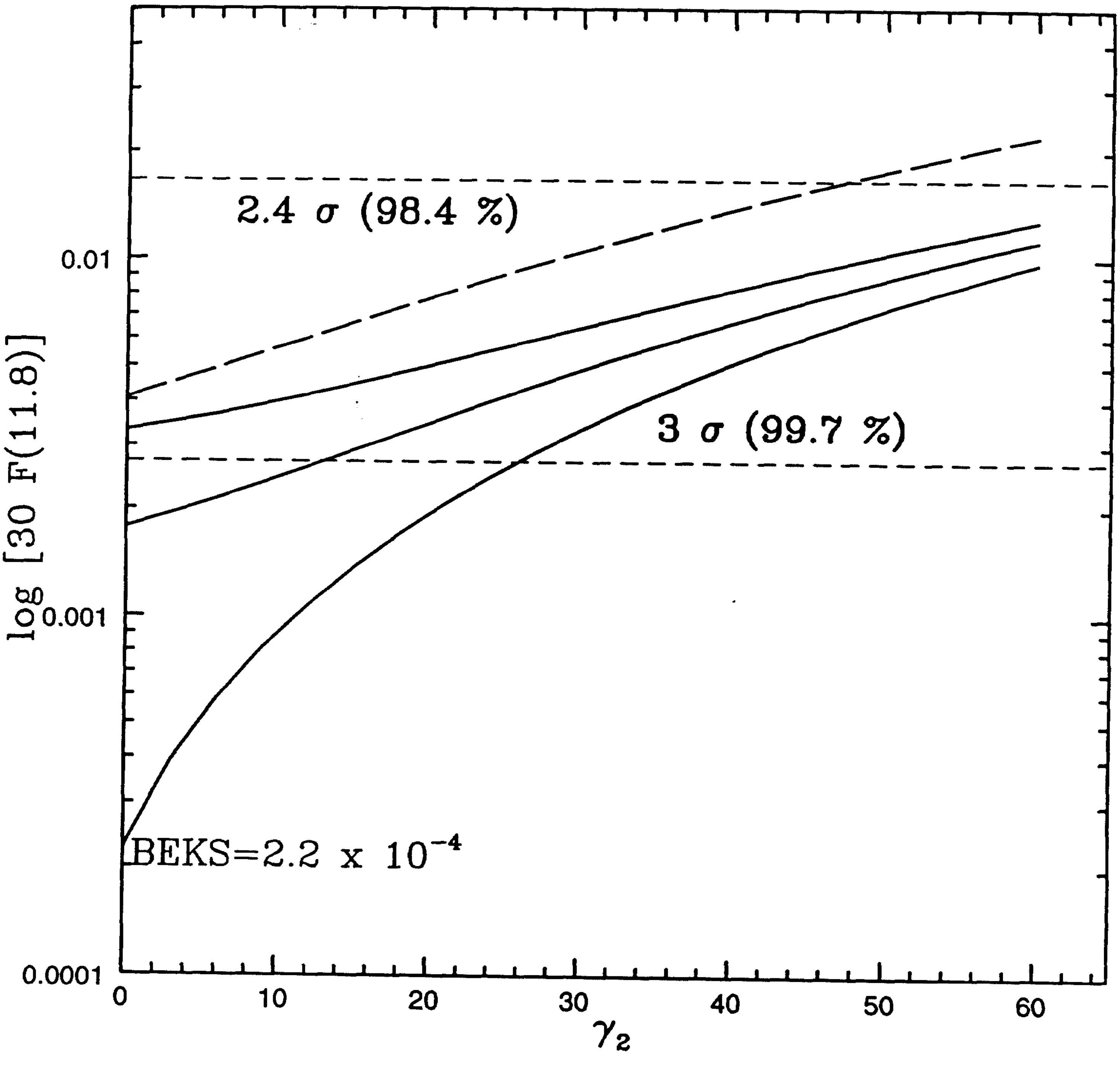

log F(X)
1
0.1
0.01
0.001
0.0001
10^{-5}
10^{-6}
0
5
10
X = peak-to-noise ratio
Non-Gaussian
Gaussian

FIG. 1



2.4 σ (98.4 %)
3 σ (99.7 %)
BEKS=2.2 x 10^-4
0.01
0.001
0.0001
log [30 F(11.8)]
0
10
20
30
40
50
60
γ2

FIG 2