Gaussian nature of the COBE data from multipoint correlations

Samuel I. Vainshtein and Andrea Malagoli

Department of Astronomy and Astrophysics, University of Chicago, Illinois 60637

Katepalli R. Sreenivasan

Mason Laboratory, Yale University, New Haven, Connecticut 06520-8286

(Received 5 October 1995)

Important information about the early universe can be obtained from a study of the cosmic background radiation (CBR) in the microwave range. We study high-order correlation and structure functions of temperature fluctuations collected from the first two years of Differential Microwave Radiometer (DMR) observations from the Cosmic Background Explorer (COBE) satellite. The intent is to determine whether the radiation data possess significant deviations from Gaussianity. The most difficult problem in drawing meaningful conclusions is the presence of instrumental noise, which is quite strong even after a two-year averaging has been performed. We have taken into account the noise in various ways. Within the limitations imposed by the noise, our study shows that the fluctuations are quite likely to be Gaussian-like. The results can, therefore, be said to favor the inflation scenario of the universe. [S0556-2821(96)02912-8]

PACS number(s): 98.70.Vc, 98.80.Bp, 98.80.Cq

I. INTRODUCTION

Experimental data on temperature fluctuations in cosmic background radiation (CBR) have been available for a few years now (see, for example, Refs. [1–6]). This radiation in the microwave wavelengths is believed to be the remnant ''afterglow'' of the early universe, corresponding to the recombination time at which baryonic matter and radiation were decoupled. The properties of CBR are, therefore, of great interest in the context of the formation of the early universe. The data, obtained primarily from the Cosmic Background Explorer (COBE) satellite, have been used recently to obtain some insight into this question, in particular, to shed light on cosmology theories. There are two families of competing theories of formation of the universe at the time of the matter-radiation decoupling: one of them involves expansion and structure formation due to gravitational instabilities (see, e.g., [7]), and the other involves symmetry-breaking processes, such as topological defects, resulting in a textured and intermittent universe (see, e.g., [8]). Careful analysis (see, for example, Refs. [4, 5]) shows that the fluctuations in the COBE data are significant beyond the various uncertainties due to noise sources, but the question that remains is whether such fluctuations represent genuine intermittency (or strong deviations from Gaussianity) or are more akin to fluctuations inherent in a Gaussian distribution. We believe that the fundamental importance of these considerations needs no further elaboration here.

A one-point analysis of the background radiation reveals that it is essentially Gaussian; the probability distribution function is also known to deviate little from Gaussian (see Sec. II for details and some references). Unfortunately, this conclusion is not decisive because of the substantial contribution of instrumental noise. Single-point statistics can be considered to represent the limit of multipoint statistics for zero separation. It can be argued (see Sec. II) that, while all correlations at zero separation are influenced by instrumental noise, the prospects for learning something useful are better

if one considers multipoint statistics for nonzero separation. Low-order multipoint statistics have already been studied: The autocorrelation corresponding to the second moment has been obtained in Refs. [9,10] and the third moment, corresponding to the skewness, has been discussed in [11] and obtained in Refs. [12,13]; see also Ref. [14] for a detailed discussion of correlation properties of COBE data. These two low-order moments are also consistent with a Gaussian multivariant distribution.

If the departures from Gaussianity are not pronounced, it is clear that one should examine high-order correlations to arrive at a definitive conclusion. This is the goal of this paper. The type of high-order statistics we consider is analogous to those used in the study of fluid turbulence where the issue of non-Gaussianity looms large. For example, even though temperature fluctuations in heated homogeneous flows are closely Gaussian, the smaller scales display strong deviations from Gaussianity. One of the principal aims of turbulence theory is the characterization of this scale-dependent deviations from Gaussianity, or the so-called ''intermittency'' effect. This effect is best explored by studying the so-called structure functions and correlations of magnitudes of quantities of relevance such as velocity and temperature. Motivated by these considerations, we compute for the COBE Differential Microwave Radiometer (DMR) data these same quantities and compare them with Gaussian counterparts. We are, however, constrained in carrying the analogy with turbulence too far because of the strong effects of instrumental noise in the COBE data. Despite this difficulty, we explore ways of extracting useful information about the temperature distribution on the sky. The inference appears to be that statistically substantiated deviations from Gaussianity do not exist.

The rest of the paper is organized as follows. After a preliminary analysis of data in Sec. II, we explore their correlations and structure functions in Sec. III and study amplitude correlations of high-order statistics in Sec. IV. In Sec. V, we introduce briefly the new topic of ''high points'' in the

sky, and conclude the paper with Sec. VI, where some principal results are reiterated.

II. PRELIMINARY CONSIDERATIONS

Single-point statistics are the simplest measures of the nature of the data. For the COBE data, these are best evaluated as averages out of the galactic plane, i.e., for the galactic latitude b satisfying the (somewhat arbitrary but reasonable) constraint $|b| \geq 15°$. If t is the temperature fluctuation around the mean value of about 2.736 K (to within a few millikelvin), direct calculations [15] result in

$$\langle t \rangle = 2.07 \ \mu K, \quad \langle |t| \rangle = 154.0 \ \mu K, \tag{1a}$$

$$\langle t^2 \rangle = 3.83 \times 10^4 \ \mu K^2, \tag{1b}$$

$$\langle t^4 \rangle = 4.88 \times 10^9 \ \mu K^4, \tag{1c}$$

$$\langle t^6 \rangle = 1.11 \times 10^{15} \ \mu K^6, \tag{1d}$$

$$\langle t^8 \rangle = 3.57 \times 10^{20} \ \mu K^8. \tag{1e}$$

From these numbers, one obtains

$$f = \frac{\langle t^4 \rangle}{\langle t^2 \rangle^2} = 3.33,$$

$$f_6 = \frac{\langle t^6 \rangle}{\langle t^2 \rangle^3} = 19.8,$$

and

$$f_8 = \frac{\langle t^8 \rangle}{\langle t^2 \rangle^4} = 166,$$

not very different from the Gaussian values of 3, 15, and 105, respectively (for more detailed calculations of this nature, see Refs. [13,14]). However, because of instrument noise information about the real distribution cannot be assessed from one-point statistics.

Indeed, let $t = \tilde{t} + n$, where $\tilde{t}$ is the sky temperature, and n is noise. Then, for homogeneous fluctuations,

$$K_2(x - x') = \langle t(x)t(x') \rangle = \langle \tilde{t}(x)\tilde{t}(x') \rangle + \langle n^2 \rangle \delta_{xx'}. \tag{2}$$

Here angular brackets denote an ensemble average (which, in principle, implies an average over all experiments and all possible universes). It was supposed above that the noise is uncorrelated at different points, and statistically independent of the signal [16]. The last term in Eq. (2) cannot be neglected if the noise is comparable to, or higher than, the signal $\tilde{t}$; then, the correlation function at the origin is strongly influenced by noise.

In the COBE satellite data we are using here, the background temperature has been obtained by averaging observations at each pixel over a two-year period. One also has, at each pixel, the root-mean-square (rms) fluctuation around the pixel mean. Since the mean is constant over the period of observations, this is indeed the rms noise, calculated with error propagation analysis; this is all we know about the noise characteristics. The rms noise is a strong function of the position in the sky (i.e., it is inhomogeneous); see

[4,5,17]. The task is to eliminate noise effects, where possible, only from the knowledge of the distribution of its rms value. As can be imagined, this imposes severe constraints on our ability to eliminate noise effects.

Returning now to the variance around the mean value of the second-order correlations, one can write the following more general expression for the correlation function by assigning different weights w_i to different pixels, because, as mentioned above, the noise rms is different at different pixels. We take the standard definition for the weights, namely, $w_i = 1/\langle n_i^2 \rangle$. Then,

$$K_2(r) = \frac{1}{M} \sum_{i,i'(r)} \frac{w_i t_i t_{i'(r)} w_{i'(r)}}{M_i(r)}$$

$$= \frac{1}{M} \sum_{i,i'(r)} \frac{w_i \tilde{t}_i \tilde{t}_{i'(r)} w_{i'(r)}}{M_i(r)} + \Delta K_2,$$

$$\Delta K_2 = \frac{1}{M} \sum_{i,i'(r)} \frac{1}{M_i(r)} w_i \{ \tilde{t}_i n_{i'(r)} + n_i \tilde{t}_{i'(r)} + n_i n_{i'(r)} \} w_{i'(r)}. \tag{3}$$

Here, the index i corresponds to any pixel, $i'(r)$ to pixels such that the angle between them and the ith pixel is r, $M = \Sigma_i w_i$, and $M_i(r) = \Sigma w_{i'}(r)$.

Expression (3) tells us about the mean correlation function, but nothing about the variances around the mean value. This variance is $\langle \Delta K_2^2 \rangle$, which can be written as

$$\sigma^2(r) = \frac{1}{M^2} \sum_{i,i'(r),j,j'(r)} \frac{w_i w_j w_{i'} w_{j'}}{M_i(r) M_j(r)} \{ \langle \tilde{t}_i \tilde{t}_j \rangle \langle n^2 \rangle \delta_{i'j'}$$

$$+ \langle \tilde{t}_i \tilde{t}_{j'} \rangle \langle n^2 \rangle \delta_{ji'} + \langle \tilde{t}_j \tilde{t}_{i'} \rangle \langle n^2 \rangle \delta_{ij'} + \langle \tilde{t}_{i'} \tilde{t}_{j'} \rangle \langle n^2 \rangle \delta_{ij}$$

$$+ \langle n^2 \rangle^2 [\delta_{ii'} \delta_{jj'} + \delta_{ij} \delta_{i'j'} + \delta_{ij'} \delta_{ji'}] \}, \tag{4}$$

where we have already considered the noise to be Gaussian. This expression is equivalent to the diagonal elements of a covariance matrix [3].

If $r \neq 0$, which means that $i \neq i'$ and $j \neq j'$, a rough estimation of (4) (which, for simplicity, also takes all the weights to be comparable) gives

$$\sigma(r) \sim \frac{1}{\sqrt{NN(r)}} \sqrt{\{ K_2(r) \langle n^2 \rangle + \langle n^2 \rangle^2 \}}, \tag{5}$$

where N is the number of pixels, $N(r) = \langle N_i(r) \rangle = \Sigma_i N_i(r)/N$, and $N_i(r)$ is the number of pixels such that the angle between them and the ith pixel is r. At the origin, $r = 0$, $i = i'$, $j = j'$, and $N_i = 1$, so that

$$\sigma(0) = \left(\langle n^2 \rangle^2 \left[1 + \frac{2}{N} \right] + \frac{4 \langle \tilde{t}^2 \rangle \langle n^2 \rangle}{N} \right)^{1/2} \approx \langle n^2 \rangle. \tag{6}$$

Equation (6) demonstrates the following. Not only is the correlation function at $r = 0$ determined strongly by noise [as already seen from Eq. (2)], but so also is the variance around the mean. The correlation function at $r \neq 0$ may have a closer connection to the sky temperature, provided the noise is acceptably low, or the product $NN(r)$ in (5) is sufficiently large, i.e., one has acceptable statistics. Therefore, in order to find measures of the sky temperature, for example,

$$T_2 = \langle \widetilde{t}\,^2 \rangle,$$

it is better to study the behavior of K_2 at $r \neq 0$ and extrapolate it to $r=0$ in some sensible way, rather than study $K_2(0)$ itself.

The same situation holds for high-order moments. For the fourth-order correlation

$$\langle [t(x)^2 - \langle t(x)^2 \rangle][t(x')^2 - \langle t(x')^2 \rangle] \rangle$$
$$= \langle t(x)^2 t(x')^2 \rangle - \langle t(x)^2 \rangle \langle t(x')^2 \rangle$$

we can write, as in Eq. (2), that

$$\langle [t(x)^2 - \langle t(x)^2 \rangle][t(x')^2 - \langle t(x')^2 \rangle] \rangle$$
$$= \langle [\widetilde{t}(x)^2 - \langle \widetilde{t}(x)^2 \rangle][\widetilde{t}(x')^2 - \langle \widetilde{t}(x')^2 \rangle] \rangle$$
$$+ [4\langle \widetilde{t}(x)^2 \widetilde{t}(x')^2 \rangle \langle n^2 \rangle + 2\langle n^2 \rangle^2] \delta_{xx'}. \quad (7)$$

Changing $x' \to x+r$, and averaging expression (7) over all pixels positioned at x, we get

$$K_4(r) = \overline{\langle t(x)^2 t(x')^2 \rangle - \langle t(x)^2 \rangle \langle t(x')^2 \rangle}.$$

We have two kinds of averages here. One is over different pixels (with index i), and another one is over all pixels in one bin, i.e., separated by a distance r, denoted in (3) with index $i'(r)$. The latter corresponds to the $\langle \cdots \rangle$ average in the expression above, and the first to the overbar. For a homogeneous process, we can do only with one kind of average, and this can be written as

$$K_4(r) = \langle t(x+r)^2 t(x)^2 \rangle - \langle t^2 \rangle^2 \quad (8)$$

(no overbar needed). Thus, for this correlation, the errors are not accumulated for $x \neq x'$.

Since the distribution of the signal could be inhomogeneous (this being a conservative statement because the noise appears to be so), we will use an expression equivalent to Eq. (7) and compare it with Eq. (8). In order to get this equivalent expression, consider again the correlation $\langle t(x+r)^2 t(x)^2 \rangle$, which can be written, analogously to Eq. (2), as

$$\langle t(x)^2 t(x')^2 \rangle = \langle \widetilde{t}(x)^2 \widetilde{t}(x')^2 \rangle + \delta n_4(x,x'),$$

$$\delta n_4(x,x') = \langle n(x)^2 \rangle \langle n(x')^2 \rangle + \langle n(x)^2 \rangle \langle \widetilde{t}(x')^2 \rangle$$
$$+ \langle n(x')^2 \rangle \langle \widetilde{t}(x)^2 \rangle + 4\langle \widetilde{t}(x)\widetilde{t}(x') \rangle \langle n(x)^2 \rangle \delta_{xx'}$$
$$+ [\langle n(x)^4 \rangle - \langle n(x)^2 \rangle^2] \delta_{xx'},$$

where

$$\langle \widetilde{t}(x)^2 \rangle = \langle t(x)^2 \rangle - \langle n(x)^2 \rangle,$$

and $\langle \widetilde{t}(x)\widetilde{t}(x') \rangle$ is found from Eq. (2).

Therefore the sky temperature correlation can be written for nonzero r as

$$R_4(r) = \langle \widetilde{t}(x+r)^2 \widetilde{t}(x)^2 \rangle$$
$$= \langle t(x+r)^2 t(x)^2 \rangle - \delta n_4(r)$$
$$= \langle [t(x+r)^2 - \langle n(x+r)^2 \rangle][t(x)^2 - \langle n(x)^2 \rangle] \rangle, \quad (9)$$

where $\delta n_4(r) = \overline{\delta n_4(x+r,x)}$ is the average value of $\delta n_4(x+r,x)$ over all x. This formula expresses the fourth-order correlation through observed second- and fourth-order correlations and rms noise at each point.

Similarly, we have (for $r \neq 0$)

$$R_8(r) = \langle \widetilde{t}(x+r)^4 \widetilde{t}(x)^4 \rangle$$
$$= \langle t(x+r)^4 t(x)^4 \rangle - \delta n_8(r)$$
$$= \langle [t(x+r)^4 - 6\widetilde{t}(x+r)^2 n(x+r)^2 - \langle n(x+r)^4 \rangle]$$
$$\times [t(x)^4 - 6\widetilde{t}(x)^2 n(x)^2 - \langle n(x)^4 \rangle] \rangle$$
$$= \langle [t(x+r)^4 - 6t(x+r)^2 n(x+r)^2 + 5\langle n(x+r)^4 \rangle]$$
$$\times [\langle t(x)^4 - 6t(x)^2 n(x)^2 + 5\langle n(x)^4 \rangle] \rangle. \quad (10)$$

Basically, these two formulas, (9) and (10), are subtracting the noise contribution $\langle n(x+r)^2 n(x)^2 \rangle$ from $\langle t(x+r)^2 t(x)^2 \rangle$ and $\langle n(x+r)^4 n(x)^4 \rangle$ from $\langle t(x+r)^4 t(x)^4 \rangle$; the other terms in these equations are of higher order and small.

Let us return to the variances. For the fourth-order correlation, expressions like (4)–(6) can be constructed by changing variables to $\widetilde{t}\,^2 + m$ from $\widetilde{t} + n$, where m is a new noise variable. Thus, again, $\langle t^4 \rangle$ would represent the statistics of the noise, and one has to study $R_4(r)$ in order to get the fourth-order moment for the temperature.

In principle, complete information about the errors gives the covariance matrix, including its off-diagonal elements. This has been described in Ref. [9], where the matrix has been written down for the correlation function $K_2(r)$. An attempt to construct a covariance matrix for the fourth-order correlation leads to cumbersome expressions; at the present stage it is not clear how to handle them when processing the data. Therefore we shall take account of only the diagonal elements when calculating the errors. While this procedure does not guarantee complete estimation of noise effects, the χ^2 analysis of the deviations of the calculated correlations from these noise models shows that the error propagation, performed this way, correspond reasonably well to these deviations—at least in the most important cases. Some cases do show that the error model gives slightly greater errors than the standard deviations, thus overestimating the instrument errors slightly. These cases are of less importance for the main results. Good consistency of error propagation due to instrumental noise with the observed deviations from the fitted lines, at least for the most important cases of high-order correlations, suggests that the diagonal elements of the covariance matrix provide a sensible approximation for the error estimation.

A potential problem with studying the high-order moments is that the data analysis in making the DMR map has been manipulated with data cuts [17]. In particular, this means that the measured data in excess of three standard deviations are excluded. This ''bad data'' flagging might have increasing influence on high-order statistics. Because of the dominance of the noise (see Sec. II A), we may hope that these limits exceed several more standard deviations of the sky temperature fluctuation, and therefore do not influence the conclusions excessively. However, this artifact should be kept in mind while interpreting the results.
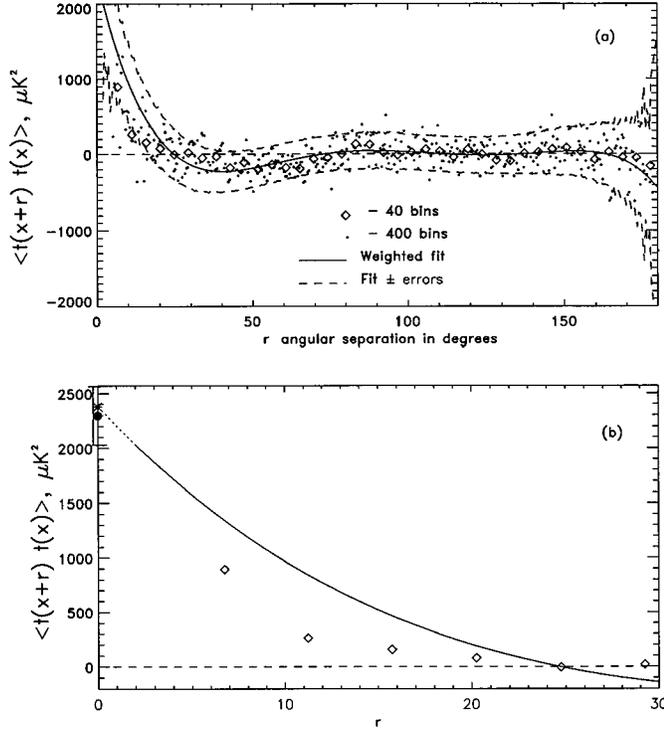
FIG. 1. Second-order correlation function $K_2(r)$, defined by Eq. (3), plotted against the angular separation r in degrees: (a) represents the whole COBE map while (b) shows data near the origin. In these and all other figures to follow, diamonds represent averages over 40 bins and dots averages over 400 bins. In (a), the solid line is the least-square polynomial weighted fit to the 400-bin data. The dashed lines shown on either side of the full line represent one standard deviation on either side of the fit. The same notation will be used in all other figures as well. The asterisk in (b) represents the value of the fit function extrapolated to zero lag, and the filled circle is obtained independently by Eq. (11). They represent the best estimate for the ''real'' temperature variance over the sky.

III. CORRELATION AND STRUCTURE FUNCTIONS

A. Correlations

It is useful to start with the correlation function $K_2(r)$, although it has been studied more than once for the COBE data [4,5,9,10]. We first note that the present calculations of the correlation function yield results identical with those obtained before. Second, as is well known, the correlation function is almost zero except for small separation angles; see Fig. 1(a). Hereafter, we will find a polynomial fit for all correlations. For example, the best fit (least-square polynomial weighted fit) for $K_2(r)$ is depicted in Fig. 1(a) (solid line). This function, denoted as $f_2(r)$, deviates to some degree from zero, especially for small angles r (though the errors are bigger in this region). There is, of course, a variability to each data point in the figure, as determined through error propagation analysis.

That the correlation function possesses a weak, yet statistically significant, structure can be demonstrated as follows. Suppose we fit the function $\{K_2(r)-f_2(r)\}$ to a straight line, $y=a+br$. If f_2 is a proper representation of the correlation function, then y should be close to zero, within errors; that is to say, $a=b=0$. Using χ^2 analysis, we obtain

$$a=-6.12\times10^{-9}, \quad b=6.32\times10^{-11},$$

$$\chi^2=291.3 \quad (n_f=393), \quad Q=0.999,$$

where χ^2 represents the deviation from the straight line, n_f represents the number of degree of freedom, and Q is the goodness of fit. Note that the latter is the probability that a value of χ^2 as poor as

$$\chi^2(a,b)=\sum_i^N \left(\frac{y_i-a-bx_i}{\sigma_i}\right)^2$$

should occur by chance.

On the other hand, a white noise with no correlation would give a simple zero line, $y=0$. Therefore we may try to fit $K_2(r)$ itself to a straight line, to see the significance of the structures. Then,

$$a=1.49\times10^{-5}, \quad b=-2.54\times10^{-7},$$

$$\chi^2=463.0 \quad (n_f=393), \quad Q=0.0064.$$

Thus $\{K_2(r)-f_2(r)\}$ definitely fits the zero line better than $K_2(r)$.

Figure 1(b) also depicts the extrapolation to zero separation, or lag, of the function $f_2(r)$. The zero-separation value can be calculated independently from Eq. (2) as

$$\langle\widetilde{t}^{\,2}\rangle=\langle t^2\rangle-\langle n^2\rangle, \tag{11}$$

which yields

$$K_2(r=0)=\langle\widetilde{t}^{\,2}\rangle=2297\pm269 \ \mu K^2. \tag{12}$$

This quantity is presented with its error bars in Fig. 1(b) (filled circle at zero lag). Remarkably, it almost coincides with the extrapolated value of the correlation $f_2(0)$ (asterisk on the figure), increasing our confidence in the estimate.

We note that both these values, given by the extrapolated fit and by Eq. (12), are still small compared with (1b), suggesting a large effect of the noise. Indeed, the ratio at the origin of $[\widetilde{t}^{\,2}/(t^2-\widetilde{t}^{\,2})]^{1/2}$ is a measure of the sky-temperature-to-noise ratio. From the present estimates, this number is of the order of 0.25—which clearly emphasizes the need to examine the four-year average COBE data, for which this ratio can be expected to improve by a factor of about 1.4.

We now consider the higher order correlation $K_{13}(r)=\langle t(x+r)t(x)^3\rangle$, shown in Fig. 2. It is clear that there is no correlation at large distances. Figure 2(b) depicts this correlation at the origin and, for comparison, the Gaussian counterpart $[=3K_2(r)^2]$. The correlation does not really deviate from Gaussian.

B. Structure functions

Structure functions were introduced by Kolmogorov [18] for the study of the ''universal'' aspects of high-Reynolds-number turbulence. The qth-order structure function is defined as

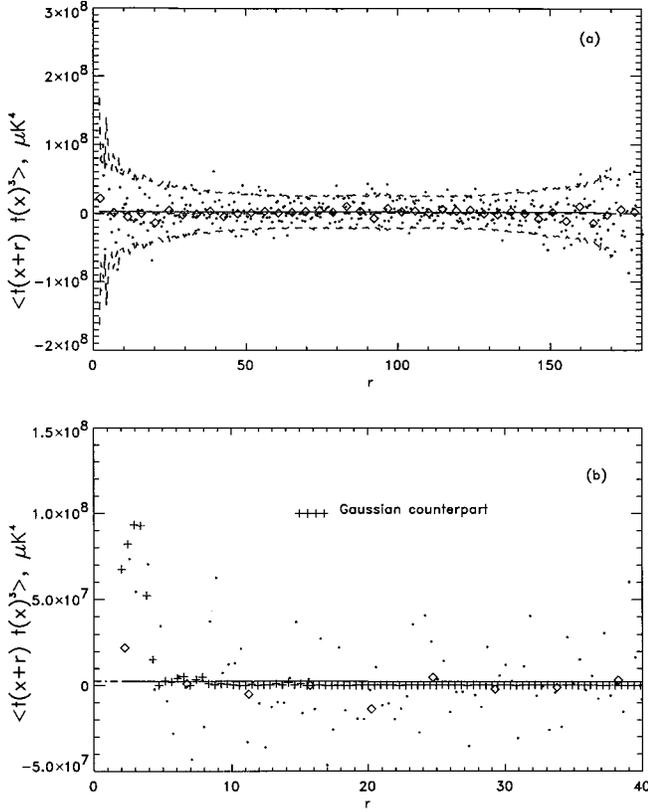$$S_q(r)=\langle[t(x+r)-t(x)]^q\rangle,$$

FIG. 2. The correlation $\langle t(x+r)t(x)^3 \rangle$. Solid line corresponds to the best polynomial fit, $--$ is the best fit $\pm$ errors. (a) and (b) have the same meaning as for Fig. 1. It can be seen that the best fit is practically a zero line, which varies even less than the Gaussian counterpart. This means that this correlation essentially vanishes at nonzero lag. This behavior would be expected for a process with small departure from Gaussianity.

for any positive integer q. The principal point of studying structure functions is that they focus on scales of size r. The structure functions can be reduced to a combination of the more familiar correlation functions. In particular, if $q=2$,

$$S_2(r) = 2[\langle t^2 \rangle - K_2(r)]. \tag{13}$$

Thus, increasing correlation for some r corresponds to decreasing $S_2(r)$, and vice versa. It follows from the definition of $S_q(r)$ that this feature is also true for structure functions of arbitrary order (i.e., any q).

We have to construct, however, "noiseless" structure functions. Analogous to Eqs. (9) and (10), we get, for the second order:

$$\widetilde{S}_2(r) = \langle [\widetilde{t}(x+r) - \widetilde{t}(x)]^2 \rangle = \langle [t(x+r) - t(x)]^2 \rangle$$
$$- \langle n(x+r)^2 \rangle - \langle n(x)^2 \rangle. \tag{14}$$

Figure 3 depicts $S_2(r)$. It clearly reflects the correlation on Fig. 1(a), as expected from Eq. (13). Substituting into Eq. (13) of the correlation function obtained independently without "noise subtraction," one obtains the structure function which practically coincides with that depicted on Fig. 3. (Note that there is no need to do the noise subtraction when constructing the correlation function.) This coincidence sim-
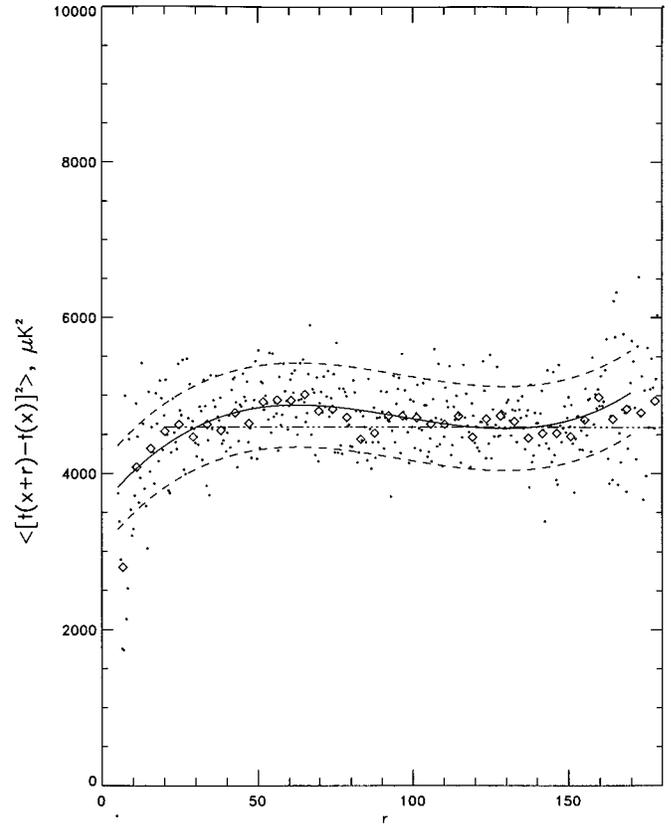

FIG. 3. The second-order structure function, defined by Eq. (14). The solid line corresponds to the best fit and the $--$ line to fit $\pm$ errors. The $\cdots$ line corresponds to the asymptotic value. Note that the fit corresponds to the correlation function depicted in Fig. 1 and that the asymptotic value is rather close the mean-square sky temperature $2\langle \widetilde{t}^2 \rangle$. This means that the noise subtraction is self-consistent.

ply means that the calculations with noise subtraction are self-consistent. Another consistency check is provided by the asymptotic value of the structure function: at infinity the correlation vanishes, and $S_2(r) \rightarrow 2\langle \widetilde{t}^2 \rangle$. Indeed, the structure function quickly approaches this value, defined by Eq. (12) (the dash-dot line in Fig. 3).

Fitting $\{S_2(r) - f_{2S}(r)\}$, where $f_{2S}(r)$ is a polynomial fit, to a straight line results in $\chi^2 = 305.4$ ($n_f = 367$) and $Q = 0.989$. For a white noise, the structure function would be zero at $r=0$ and constant (coinciding with the dash-dotted line on Fig. 3) elsewhere. Fitting the structure function itself to a straight line, i.e., treating it as a noise, results in $\chi^2 = 382.7$ ($n_f = 367$), and $Q = 0.252$. Thus the goodness of fit is statistically poorer if we treat the structure function as white noise.

It is clear from Fig. 3 that $S_2(r \rightarrow \infty) \rightarrow 2\langle \widetilde{t}^2 \rangle$ has been achieved over many bins. It may at first be thought that this asymptotic value provides a more reliable estimate of $\langle \widetilde{t}^2 \rangle$ than that given in Fig. 1(b). But this is deceptive because the correlation function is averaged over N^2 points while the structure functions are averaged simply over the number N itself. In any case, the estimate $\langle \widetilde{t}^2 \rangle$ from structure functions is close to that from Fig. 1(b).

In summary, we may conclude that the second-order structure function shows consistency with the noise subtrac-
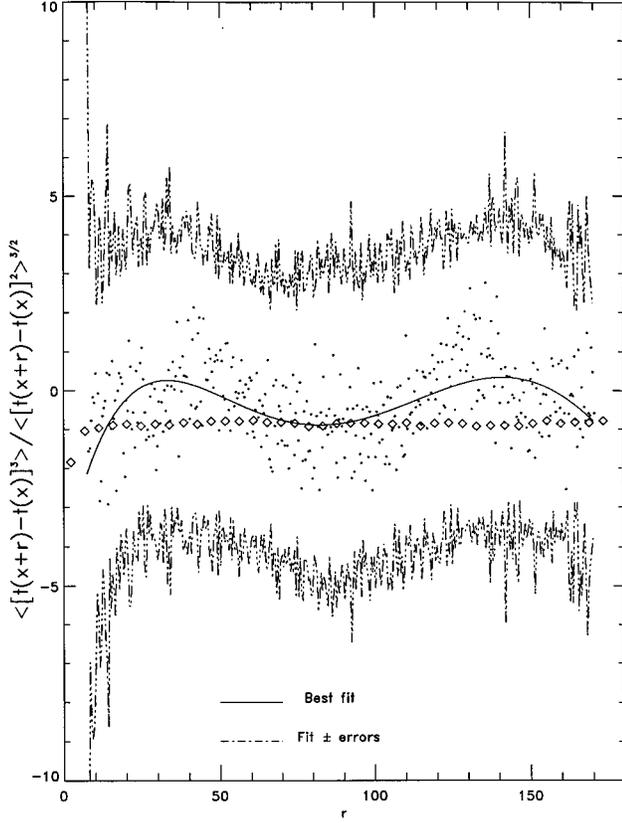
FIG. 4. The skewness defined in Sec. III B. All deviations from zero are clearly within errors.

tion method, i.e., both $S_2(r)$ and its asymptotic value correspond to the correlation function given by (13), and to the independent estimate given by Eq. (12).

The first indication of the deviation from Gaussianity would be a nonvanishing value of the skewness [11,12]. Consider the structure function of the third order,

$$\langle [t(x+r)-t(x)]^3 \rangle,$$

and define skewness, in analogy with turbulence theory, as

$$S = \frac{\langle [t(x+r)-t(x)]^3 \rangle}{\langle [t(x+r)-t(x)]^2 \rangle^{3/2}}.$$

This quantity, which in general may depend on r, is depicted in Fig. 4 for the COBE data. The skewness is essentially zero. More precisely, it is within the Gaussian variance. Our conclusion, based on the structure function, is in agreement with that of Ref. [12] based on three-point correlation studies. We again fit the function by a polynomial and denote it by $f_3(r)$. Now, χ^2 for $\{S(r)-f_3(r)\}$ is 21.63, $n_f=362$, even less than for the case when we fit the skewness itself to the zero line [19]. In that case $\chi^2=26.19$. That is to say, the zero line fits the skewness values marginally better. This implies that the multipoint distribution function is actually symmetric.

Analogous to Eq. (14), we can construct a ''noiseless'' fourth-order structure function:
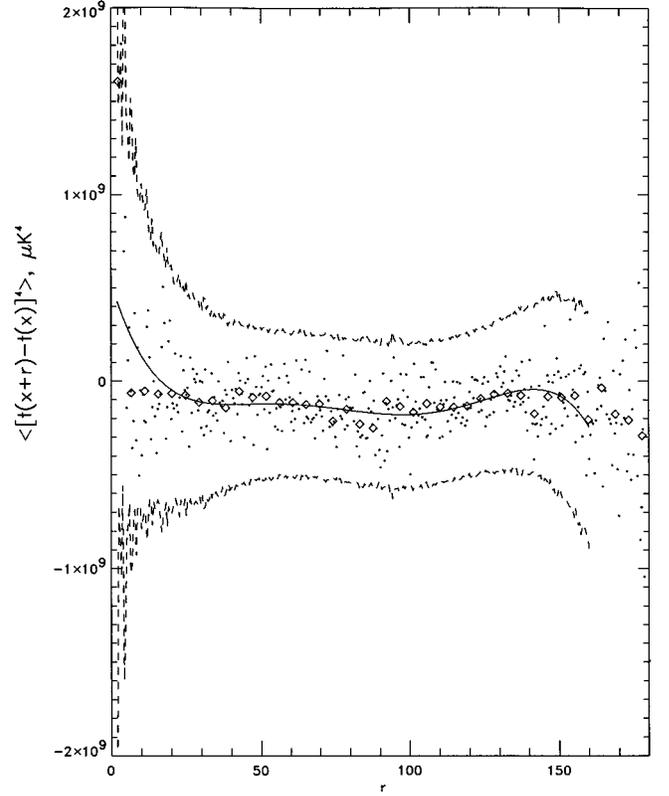

FIG. 5. The fourth-order structure function, $\langle [\tilde{t}(x+r)-\tilde{t}(x)]^4 \rangle$, defined by Eq. (15).

$$\begin{aligned}
\widetilde{S}_4(r) &= \langle [\tilde{t}(x+x)-\tilde{t}(x)]^4 \rangle \\
&= \langle [t(x+x)-t(x)]^4 \rangle - \langle n(x+r)^4 \rangle - \langle n(x)^4 \rangle \\
&\quad + 6[\langle n(x+r)^2 \rangle^2 + \langle n(x)^2 \rangle^2 + \langle n(x+r)^2 \rangle \langle n(x)^2 \rangle] \\
&\quad - 6\langle [t(x+x)-t(x)]^2 \rangle [\langle n(x+r)^2 \rangle + \langle n(x)^2 \rangle].
\end{aligned}$$
(15)

Figure 5 shows that this structure function is embedded within errors and that noise effects are too strong. Four-year observations of COBE could presumably give a more substantial result, and it would be interesting to study them from this point of view.

IV. CORRELATIONS OF MAGNITUDES

In this section, we study the correlations of absolute values of temperature fluctuations. We consider even moments (and so do not explicitly write the absolute value signs) and turn to the fourth-order correlation $R_4(r)$ defined in Eq. (9). The results are shown in Fig. 6. One glance at the figure suffices to show that the statistics are far from perfect. In fact, $\chi^2=167.1$ ($n_f=374$). This small value of χ^2 means that the errors are somewhat overestimated, which might be related to the neglect of off-diagonal elements in the covariance matrix; see the end of Sec. II.

In spite of this limitation, the (weighted) average of $R_4(r)$ (which essentially corresponds to the asymptotic value of the fourth-order correlation) gives for $R_4(r\to\infty)=\langle \tilde{t}^2 \rangle^2$ the value
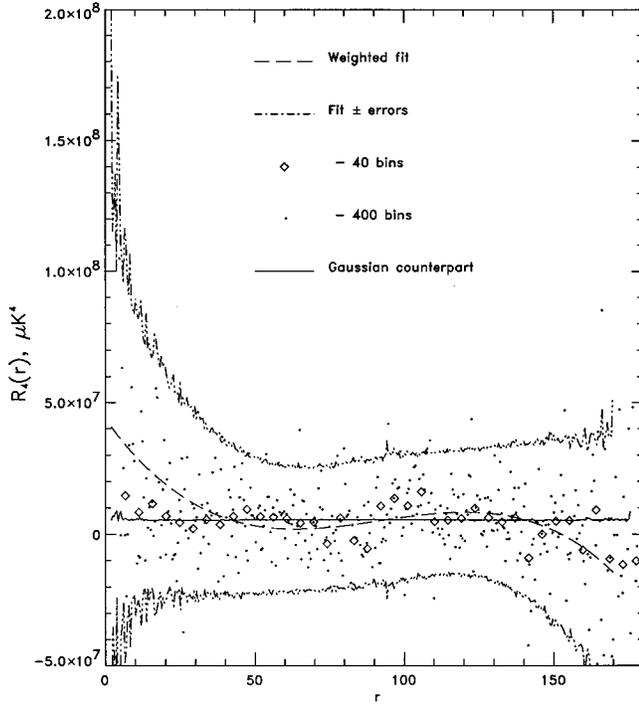
FIG. 6. The fourth-order correlation of magnitudes, $R_4(r) = \langle \tilde{t}(x+r)^2 \tilde{t}(x)^2 \rangle$, defined according to Eq. (9).

$$\langle R_4 \rangle = (5.51 \times 1.45) \times 10^6 \ \mu\text{K}^4.$$

The ratio of this value to the square of $\langle \tilde{t}^2 \rangle$ given by Eq. (12) is 1.04 which, being close to unity, is another check on self-consistency.

The second realistic feature of this correlation is that the "smoothed value," i.e., the 40-bin presentation of $R_4(r)$, does not really deviate from its Gaussian counterpart $[=\langle \tilde{t}^2 \rangle^2 + 2K_2(r)^2]$. Indeed, if we prescribe the value of $\langle \tilde{t}^4 \rangle$ to be the first diamond in Fig. 6, the flatness factor will be

$$f = \frac{\langle \tilde{t}^4 \rangle}{\langle \tilde{t}^2 \rangle^2} = 2.76.$$

If, on the other hand, we trust the fit $f_4(r)$ and prescribe $\langle \tilde{t}^4 \rangle = f_4(0)$, one obtains

$$f = 8.30.$$

However, because of the uncertainties inherent in this extrapolation as well as due to the large noise, we cannot claim that this represents a reliable deviation from the Gaussian value ($f = 3$).

Finally, Fig. 7 depicts the eighth-order correlation $R_8(r)$ defined by expression (10). It is clearly within the error. It is hard to say that the correlation deviates significantly from its Gaussian counterpart $[=72\langle \tilde{t}^2 \rangle^2 K_2(r)^2 + 24K_2(r)^4 + 9\langle \tilde{t}^2 \rangle^4]$.

V. BRIGHT OBJECTS IN THE SKY

We now introduce the quantity
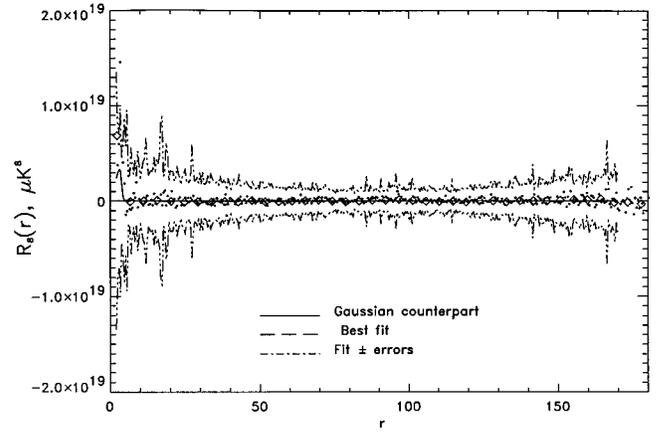
$$\text{dif}(x) = t(x)^2 - n(x)^2, \tag{16}$$

FIG. 7. The eighth-order correlation of magnitudes, $R_8(r) = \langle \tilde{t}(x+r)^4 \tilde{t}(x)^4 \rangle$, defined by Eq. (10).

which, at any position in the sky, corresponds to the square of the temperature above the background noise. It is clear that Eq. (16), when averaged over the whole sky, yields $\langle \tilde{t}^2 \rangle$. However, Eq. (16) has the advantage that it shows positions in the sky where the *local* temperature exceeds the *local* noise. If we prescribe, for instance, that the local temperature exceed three standard deviations of the noise locally, $t(x)^2 \geq 9n(x)^2$ so that $\text{dif}(x) \geq 8n(x)^2$, we will be looking at exceedingly large temperature fluctuations, corresponding to either very low or high density variations. While one hopes that these "outstanding objects" in the sky are real, there is no guarantee that they do not arise from big fluctuations of the instrumental noise itself. One may hope that at least some of them would survive in the four-year observations, in which case they would deserve more trust. The situation is the same as with the structures on the conventional COBE map. The latter is given in Fig. 8(a), with all the bright objects drawn on it. These objects are listed in Table I.

Another interpretation of the quantity $\text{dif}(x)$ is that, when it is averaged over a few neighboring pixels, say of the order 10, it may reflect cosmic variations of temperature fluctuations better than the noise-affected temperature itself. This quantity, denoted by $\langle \text{dif}(x) \rangle$ has been computed. This is presented, for the ten-pixel average, as a map in Fig. 8(b). The contour lines are set to the zero value of $\langle \text{dif}(x) \rangle$, so that the measured temperature in the positive regions exceeds one standard deviation (for the locally averaged quantity). We again hope that this map of temperature excess, or variation of fluctuations, would reemerge more or less in the same way in maps from four-year averages. Only then could it be trusted. The main point, however, is that there appear to be some small number of regions in the sky where very large temperature variations exist—despite our major conclusion that the temperature fluctuations are by and large Gaussian.

Finally, in analogy with Eq. (16), it would be tempting to calculate the fourth-order moment with subtracted noise,

$$\langle \tilde{t}^4 \rangle = \langle t^4 \rangle - \langle n^4 \rangle - 6\langle n^2 \rangle (\langle t^2 \rangle - \langle n^2 \rangle).$$

Unfortunately, the errors are large for present data. Indeed, we find that the error is mostly defined by the variance of the term $4\tilde{t}n^3$. For the present signal-to-noise ratio, the error
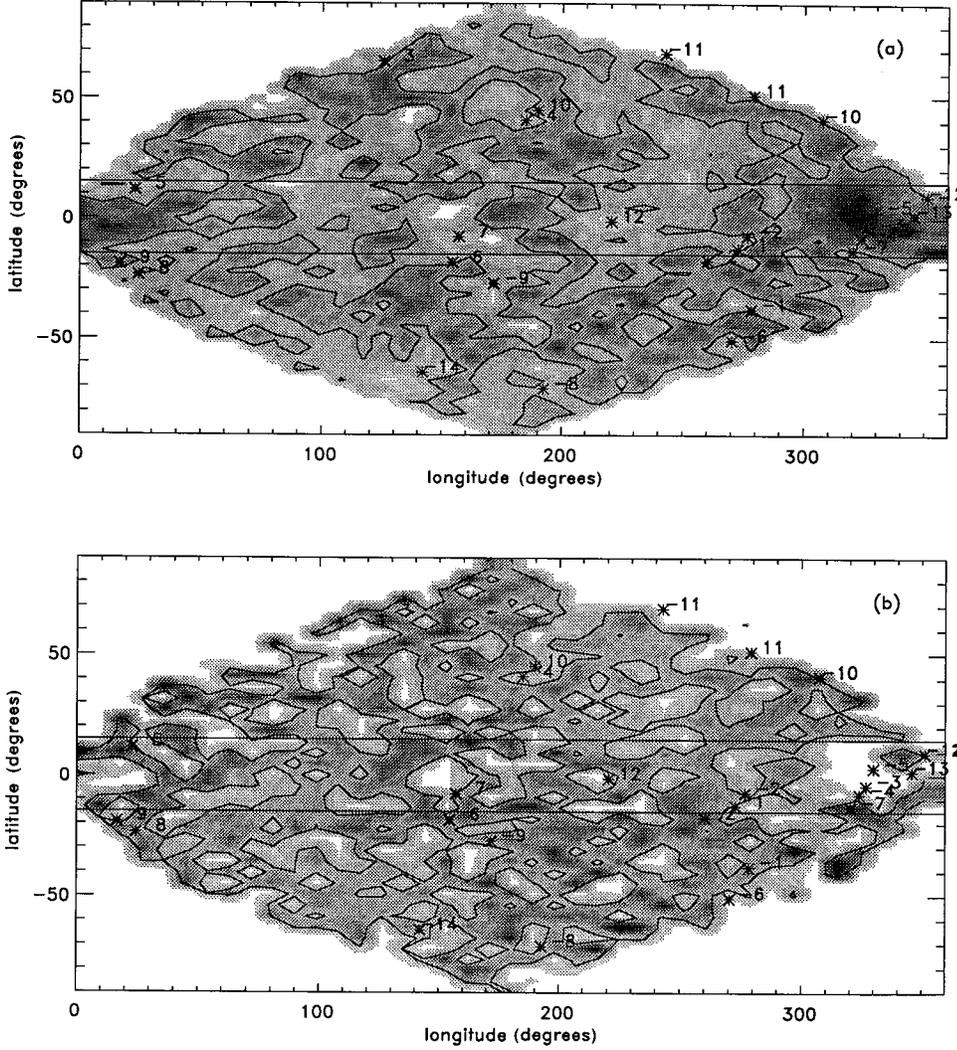
FIG. 8. (a) Conventional COBE map, except that the bright ''objects'' of Table I are added. Contour line corresponds to zero temperature fluctuation. The highest number corresponds to the highest temperature (fluctuation), down to the negative numbers corresponding to the negative temperature (fluctuation). The galactic plane, $b \lesssim 15°$, is separated by two horizontal lines. (b) The map of the locally averaged temperature excess $\langle \text{dif}(x) \rangle$ according to the definition (18). Contour lines are set to zero value, and the numbers have the same meaning as in (a).

would exceed the quantity of $\langle \tilde{t}^4 \rangle$, rendering these calculations meaningless. Four-year observations may present a better opportunity for these studies, but nothing of consequence can be said with the two-year data.

VI. SUMMARY AND DISCUSSION

We have constructed correlations and attempted to exclude instrumental noise. Consideration of noise is crucial for these kinds of data because instrumental noise is a large part of the signal. Another feature of processing these data is that we have taken explicit account of the inhomogeneity of the noise. Indeed, the noise is itself correlated over the sky as a bipole structure and the correlation function is nonzero for nonzero lag.

The main conclusion is that the calculations do not support, within the serious limitations imposed by the noise, any substantial deviation from Gaussianity. The various correlation functions computed are close to their Gaussian counterparts. Zero-lag data are especially vulnerable to instrumental noise, and so no statistical significance can be attributed to the deviation seen on Fig. 6.

It is worth exploring this last statement in more detail. Let us return to the estimation of the errors by Eq. (12). The

exact expression for the second moment can be written as

$$\langle \tilde{t}^2 \rangle = \langle t^2 \rangle - \langle n^2 \rangle - 2\langle \tilde{t}n \rangle. \tag{17}$$

The last term in Eq. (17) vanishes because the sky temperature is statistically independent of the noise, and $\langle \tilde{t} \rangle = 0$, $\langle n \rangle = 0$. This happens, however, only on the average. That is to say, it can be neglected only if there is a sufficient statistical ensemble. Its deviation from zero can be estimated by standard error propagation procedure which gives the estimate

$$2\langle \tilde{t}n \rangle \sim 2 \frac{\langle \tilde{t}^2 \rangle^{1/2} \langle n^2 \rangle^{1/2}}{\sqrt{N}}. \tag{18}$$

This error should be compared with the value of $\langle \tilde{t}^2 \rangle$. These error bars are, in fact, depicted in Figs. 1(a) and 1(b). Fortunately, the two-year observations of COBE data do provide decent error bars. Recall that the zero-lag value has been obtained by two independent methods. One of them consists of constructing a polynomial fit for the correlation function at $r \neq 0$, and extrapolating it to $r = 0$. This fit has information about all the points in different bins. An independent way of obtaining this value is to use Eqs. (11) and (12). It is inter-

TABLE I. Bright objects on the conventional COBE map.

	Temperature	Errors	Latitude	Longitude
12	0.864 377	0.189 747	220.530	−1.277 66[a]
11	0.766 569	0.222 113	338.825	51.298 5
10	0.756 953	0.222 592	193.933	45.117 4
9	0.745 256	0.218 814	7.129 73	−19.119 9
8	0.728 039	0.220 787	10.142 9	−23.715 1
7	0.690 435	0.207 147	156.761	−7.656 89[a]
6	0.654 277	0.209 410	153.055	−18.545 4
5	0.653 849	0.208 033	19.4990	11.706 5[a]
4	0.646 905	0.210 369	186.424	40.997 5
3	0.641 961	0.191 624	48.1382	65.625 2
2	0.552 024	0.150 714	264.266	−17.665 0
1	0.482 502	0.149 424	275.101	−12.691 8[a]
−1	−0.413 945	0.123 104	304.916	−37.997 2
−2	−0.482 978	0.137 430	277.595	−7.252 15[a]
−3	−0.561 857	0.184 247	327.184	−4.244 02[a]
−4	−0.585 290	0.180 302	325.354	−7.804 55[a]
−5	−0.590 034	0.193 764	330.136	3.107 37[a]
−6	−0.590 207	0.171 709	323.491	−50.972 4
−7	−0.604 261	0.176 383	324.126	−13.224 1[a]
−8	−0.616 182	0.187 684	218.101	−70.803 8
−9	−0.636 579	0.208 835	170.747	−27.279 1
−10	−0.671 884	0.221 692	348.736	40.900 0
−11	−0.704 625	0.211 928	354.781	68.895 0
−12	−0.772 940	0.218 140	353.577	9.243 54[a]
−13	−0.791 669	0.208 439	346.031	1.529 15[a]
−14	−0.796 089	0.257 777	93.3616	−64.065 2

[a]Objects in the Galactic plane.

esting to note that this value is higher (by 2.5 times) than the first-bin value of the correlation function in the 40-bin case. This suggests, as can be clearly seen from Fig. 1(a), that the correlation length is somewhat smaller than the resolution of

a pixel, which is about 7°. The correlation length estimated from Fig. 1(b) is about 6.5°, suggesting that the scale of cosmic variations is of that order. It should, however, be recalled that the noise might influence this assessment quite seriously, and that a better idea of cosmic variations can be had from Fig. 8(b) where dif(x) has been plotted.

Note added. We have become aware of two unpublished reports by Coulson *et al.* [20] and Graham *et al.* [21]. These papers address the issue of potential non-Gaussianity of the COBE data and its possible interpretation. However, their approaches are quite different from ours. Graham *et al.* define a quantity which is the average of neighboring q data points in the COBE map, and especially examine the case of the third moment. These quantities are different from the structure functions used in this paper. Graham *et al.* suggest that the observed non-Gaussianity is too strong to be attributed entirely to instrumentation noise. Coulson *et al.* examine the possible texture of the background radiation by means of cosmic string monopole theory. This interesting work bears only weakly on the structures described in Sec. V, and highlighted in Table I of this paper. The two papers are complementary to the present discussion, although different in spirit and detail.

[1] G. F. Smoot *et al.*, Astrophys. J. **396**, L1 (1992).

[2] C. L. Bennet *et al.*, Astrophys. J. **396**, L7 (1992).

[3] E. L. Wright *et al.*, Astrophys. J. **396**, L13 (1992).

[4] G. F. Smoot *et al.*, Astrophys. J. **396**, L1 (1992).

[5] A. Kogut *et al.*, Astrophys. J. **401**, 1 (1992).

[6] A. Kogut *et al.*, Astrophys J. **419**, 1 (1993).

[7] J. Bardeen, P. Steinhardt, and M. Turner, Phys. Rev. D **28**, 679 (1983).

[8] U.-L. Pen, D. N. Spergel, and N. Turok, Phys. Rev. D **49**, 692 (1994).

[9] K. Ganga, E. Cheng, S. Meyer, and L. Page, Astrophys. J. **410**, L57 (1993).

[10] C. L. Bennett *et al.*, Astrophys. J. **436**, 423 (1994).

[11] K. Yamamoto and M. Sasaki, Astrophys. J. **435**, L83 (1994).

[12] G. Hinshaw, A. J. Banday, C. L. Bennett, K. M. Górski, and A. Kogut, Astrophys. J. **446**. L67 (1995).

[13] X. Luo and D. N. Schramm, Phys. Rev. Lett. **71**, 1124 (1993).

[14] X. Luo, Phys. Rev. D **49**, 3810 (1994).

[15] Strictly speaking, the average of the temperature fluctuation should be identically zero. The nonzero value obtained here

simply reflects the differences in the manner in which the average has been removed in the data and the way the fluctuations are processed now. In any case, the average is small compared to $\langle |t| \rangle$, for instance, and is ignored in future calculations.

[16] The nonzero correlation that appears to be prevalent in the noise suggests different considerations, as explored later in the text.

[17] P. B. Keegstra *et al.*, in *Astronomical Data Analysis Software and Systems I*, edited by D. M. Worrall, C. Biemesderfer, and J. Barnes, ASP Conf. Ser. No. 25 (Astronomical Society of the Pacific, San Francisco, 1992), p. 530.

[18] A. N. Kolmogorov, C. R. Acad. Sci. USSR **30**, 299 (1941).

[19] Because χ^2 is small compared to n_f, one obtains a Q value that is almost identically unity. This, however, is misleading and so is not given here or at other places subsequently.

[20] D. Coulson, P. Ferreira, P. Graham, and N. Turok, Report No. PUP-TH-93//1393, 1993 (unpublished).

[21] P. Graham, N. Turok, P. M. Lubin, and J. A. Schuster, Report No. PUP-TH-1408, 1993 (unpublished).