CORRELATION FUNCTIONS FROM THE PERSEUS-PISCES REDSHIFT SURVEY

S. A. BONOMETTO,[1,2] A. IOVINO,[3] L. GUZZO,[4] R. GIOVANELLI,[5] AND M. HAYNES[5]

ABSTRACT

The three- and four-point correlation functions are evaluated from the Perseus Pisces galaxy redshift sample. Former results on the two-point function ξ are also reviewed. In the range of distances from 1 h^{-1} Mpc to 10 h^{-1} Mpc, three- and four-point functions do not agree with the expressions of the hierarchical clustering model. On the contrary they are steeper and can be fitted by adding terms with a higher power of ξ. Such results from a redshift catalog partially modifies the picture of the large-scale matter distribution obtained from early two-dimensional samples. The relation among these results, *biased* models of galaxy formation, and observed velocity fields is discussed.

Subject headings: galaxies: clustering — galaxies: formation — methods: numerical

1. INTRODUCTION

One of the main aims of cosmology is to attain a quantitative description of the distribution of matter over scales not affected by dissipative phenomena. As part of this effort, during the last 20 years, a large deal of work has been devoted to measuring correlation functions. Much attention was concentrated on the two-point function, which is however not sufficient in general to distinguish among different clustering models. Higher order correlation functions are the natural way to extend such information. Other kinds of statistical measures were also suggested, e.g., the void probability function (White 1979; Fry 1984; Schaeffer 1984; Bouchet & Lachièze-Rey 1986; Balian & Schaeffer 1989) and the multifractal analysis (Mandelbrot 1982; Paladin & Vulpiani 1984, 1987; Pietronero 1987; Martinez et al. 1990; Valdarini, Borgani, & Provenzale 1992, and references therein). However, it must be outlined that correlation functions are more than a standard reference point in this context.

Their statistical significance is related to the distribution $P[\epsilon_x]$ of density fluctuations $\epsilon_x \equiv \epsilon(x) = \rho(x)/\langle\rho\rangle - 1$, where $\rho(x)$ is the density in the point x in three-dimensional space and $\langle\rho\rangle$ is its mean value. If we perform a McLaurin functional expansion of the Fourier transform of $P[\epsilon_x]$,

$$Z[J_x] = \int D\epsilon_x P[\epsilon_x] e^{iJ_x \epsilon_x} \qquad (1.1)$$

(often called *partition functional*), the expansion coefficients are essentially the unreduced n-point functions (see, e.g., Ramond 1981). Here and below we adopt the continuous Einstein convention: $J_x \epsilon_x = \int d^3 x J(x)\epsilon(x)$. Furthermore, if we consider $W = \ln Z$, its McLaurin expansion directly reads

$$W[J_x] = \sum_{n=2}^{\infty} \frac{i^n}{n!} \xi^{(n)}_{x_1,\ldots,x_n} J_{x_1} \ldots J_{x_n}, \qquad (1.2)$$

[1] Dipartimento di Fisica dell'Università di Milano, Via Celoria 16, I-20133 Milano, Italy.

[2] I.N.F.N.—Sezione di Milano.

[3] Osservatorio Astronomico di Brera, Via Brera 28, I-20121 Milano, Italy.

[4] Osservatorio Astronomico di Brera, Sede di Merate, Via Bianchi 46, I-22055 Merate (CO), Italy.

[5] Department of Astronomy and National Astronomy and Ionosphere Center,[6] Space Sciences Building, Cornell University, Ithaca, NY 14853.

[6] The National Astronomy and Ionosphere Center is operated by Cornell University under a cooperative agreement with the National Science Foundation.

where the $\xi^{(n)}_{x_1,\ldots,x_n}$ are the n-point reduced (or connected) functions.

Accordingly, should all connected functions $\xi^{(n)}$ with $n > 2$ vanish, the distribution $P[\epsilon]$ is Gaussian. This can be shown by starting from the inverse of equation (1.1), which reads

$$P[\epsilon_x] = \mathcal{N} \int DJ_x Z[J_x] e^{-iJ_x \epsilon_x}, \qquad (1.3)$$

(here $\mathcal{N}$ is a normalization constant), and by replacing there the expression

$$Z[J_x] = \exp -\tfrac{1}{2}\xi_{x_1, x_2} J_{x_1} J_{x_2}.$$

If we consider $\xi_{x_1 x_2} \equiv \xi^{(2)}_{x_1, x_2}$ as a continuous matrix and take its inverse $K_{x_1, x_2} = [\xi^{-1}]_{x_1, x_2}$, it is then straightforward to obtain that

$$P[\epsilon] = \mathcal{N} \exp -\tfrac{1}{2}\epsilon_{x_1} K_{x_1, x_2} \epsilon_{x_2}, \qquad (1.4)$$

i.e., that $P[\epsilon]$ is Gaussian.

A nonvanishing two-point function provides us with the information on the way how galaxy distribution deviates from Poisson. Nonvanishing higher order functions, instead, tell us how galaxy distribution deviates from Gaussian. When still in its linear stages, matter distribution had to be non-Poissonian, i.e., inhomogeneities had to exist, which could later evolve toward present epoch structures. On the contrary non-Gaussian behavior needs not to be present; in fact, standard inflationary theories give rise to a purely Gaussian spectrum.

Nonlinear terms in gravitational evolution equations directly cause deviations from pure Gaussian behavior. As soon as the evolution itself becomes nonlinear we must therefore expect that $P[\epsilon_x]$ shows a substantial non-Gaussian signature.

Furthermore, if there is a class of objects exhibiting a biased distribution vis-a-vis that of the matter, even if the matter distribution is Gaussian, the distribution of such objects is non-Gaussian.

Finally, it is also quite possible that primeval fluctuations had a nonstandard origin and spectra were non-Gaussian even in their linear stages.

The subject of this paper is a measure of the three- and four-point correlation functions applying the moment method to the Perseus Pisces redshift Survey (PPS). We rederive and discuss also in detail the two-point correlation function, given its importance for the subsequent higher order analysis. The

analysis leads to some fairly unexpected conclusions. Previous angular sample analyses led to the introduction of the so-called *hierarchical clustering model*, according to which the reduced *n*-point function $\xi^{(n)}$ can be written in the form $\sum_t Q_t^{(n)}[\xi^{(2)}]^{n-1}$ (Peebles 1980; § 73; *see below*). This model is *not* confirmed by the present three-dimensional data analysis: the three- and four-point functions are found to be steeper than predicted by the above model.

The paper is organized as follows. Section 2 describes the data. In § 3 we discuss the analytical techniques used in the analysis together with the theoretical predictions related to the above described scenarios. In § 4 the results of the analysis are presented, while in § 5 they are discussed and conclusions are drawn.

2. THE DATA

The PPS database was compiled by Giovanelli and Haynes during the last decade (see, e.g., Giovanelli & Haynes 1993, and references therein). Thanks to its completeness, its large volume, and the fact of containing a number of both under-dense and overdense regions (i.e., *voids* and density enhancements like the Perseus-Pisces supercluster), it provides an excellent sample to investigate the properties of galaxy distribution.

The survey consists mainly of highly accurate 21 cm H I line redshifts, partly unpublished, obtained with the NAIC 305 m telescope in Arecibo[7] at declinations between 0° and 38°, and with the NRAO[8] 300 foot telescope formerly in Green Bank (Giovanelli & Haynes 1985; Giovanelli, Haynes, & Chincarini 1986; Haynes et al. 1988; Giovanelli & Haynes 1989). The radio data are complemented with optical observations of early-type galaxies carried out at the 2.4 m telescope of the McGraw-Hill Observatory[9] (Wegner, Haynes, & Giovanelli 1994), plus a number of redshifts kindly provided by J. Huchra and other smaller sources in the public domain.

All through this work we shall refer to PPS as the region bound by [R.A. = 22^h–04^h], [Decl. = 0°–45°], which is somewhat more restrictive than the region surveyed by Giovanelli, Haynes, and coworkers (in order to reject regions of sky where the incompleteness is too high). The catalog used comprises those redshifts obtained before 1991 December, for a total of 5183 galaxies. Among them, 3854 have Zwicky magnitudes (Zwicky et al. 1961–1968, hereafter CGCG) of 15.7 or brighter.

For the analysis presented here, the following corrections have been applied to the observed, heliocentric, radial velocities. First, a velocity correction $\Delta v = 300 \sin l \cos b$ km s^{-1} was applied to take into account galactic rotation and the motion of the Galaxy with respect to the bulk of the Local Group. A simplified model was applied for Virgocentric flow. The projection of the Galaxy's motion with respect to the Virgo Cluster onto the line of sight to each galaxy of the sample, was added to the radial velocity. Furthermore, it was assumed that galaxies at distance R from M87 fall toward Virgo according to an inverse-distance law for $R < R_o$, where

R_o is the Local Group's distance from M87, while those between R_o and $2R_o$ fall with a velocity that decreases linearly between 250 km s^{-1} at R_o and 0 at $2R_o$. The latter correction affects a minority of galaxies in the sample, which is located in the anti-Virgo octant of the sky, and where most galaxies are at velocities larger than 1500 km s^{-1}. Finally, velocity dispersions in rich clusters were compressed to eliminate virial *fingers*. Further details on these corrections and on the survey in general can be found in Haynes & Giovanelli (1988) and Iovino et al. (1993).

Because the survey region's northern parts are at relatively low galactic latitude, extinction is a serious concern. To this end, Zwicky magnitudes are corrected for extinction using the absorption maps produced by D. Burstein & C. E. Heiles (1988–1992, private communication). The extinction map of the survey region is shown in Giovanelli et al. (1986). Excluding galaxies fainter than $m_{Zwicky} = 15.5$ from the catalog, after correcting extinction, eliminates incompleteness on the far side of the survey wherever absorption is less or equal to 0.2 mag. This was the strategy used in Guzzo et al. (1991, hereafter G91). Here we decided to be more conservative, with the aim to check possible effects of extinction on the results obtained by G91 on the shape of $\xi(r)$ (which, nevertheless, were in very good agreement with the results of the same analysis performed on the CfA$_{14.5}$). Therefore, we clipped out those regions which might be more seriously affected by extinction, restricting the sample within the following ranges: [R.A. = 22^h–03^h10^m], [decl. = 0°–42°30′]. We were left with 3525 galaxies with Zwicky magnitude brighter than 15.5 (after correction for extinction), and velocity in the range 0–13,000 km s^{-1}.

Among them, we selected a volume limited sample (VLS in the following), by requiring that the absolute magnitude of the galaxy be $M \le -19 + 5 \log h$ and its distance $d \le 79 \ h^{-1}$ Mpc (here h is the value of the Hubble parameter in units of 100 km s^{-1} Mpc^{-1}). This sample contains 948 galaxies and the main results presented here are based on it. Part of the analysis was repeated also on the whole apparent-magnitude limited sample (mLS in the following). The use of VLS has two main advantages. First, it is homogeneous in the galaxy properties, including only objects within a well defined luminosity range. Secondly, it represents the best possible option in terms of equally weighing different redshift ranges of the sample. In our case, owing to the form of the selection function, the Perseus-Pisces supercluster, at a distance of ~50 h^{-1} Mpc, lies in the distance range wherefrom the top contribution to mLS comes. If not treated with suitable care, such relevant structure can cause an overestimate of the mean galaxy density and a consequent reduction of the correlation length. Using VLS, instead, minimizes such effect, as low-density regions behind and beyond the supercluster are included with the same weight of the supercluster volume.

3. METHODS OF ANALYSIS

3.1. *The Two-Point Function*

It is well established that the two-point function is fitted by the expression $\xi(r) = (r_o/r)^\gamma$ up to ~10 h^{-1} Mpc. The present analysis of three- and four-point functions is based on such expression for the two-point function.

In this paragraph we want to describe how such expression can be compared with another expression of the two-point function, initially obtained from the PPS and CfA$_{14.5}$ surveys

(G91; see also Dekel & Aarseth 1984), and subsequently confirmed also on angular samples (Calzetti, Giavalisco, & Meiksin 1992). According to G91, a detailed fit to the data over a more extended distance range of separations is obtained using a two-power-law expression:

$$\xi(r) = \begin{cases} A_1 r^{\gamma_1} - 1 & \text{for} \quad r < r_b \\ A_2 r^{\gamma_2} - 1 & \text{for} \quad r_b < r < r_c . \end{cases}$$

Continuity requires that $A_1 = A_2 r_b^{\gamma_2 - \gamma_1}$, while $A_2 r_c^{\gamma_2} = 1$, so this expression contains four parameters. For PPS data, G91 obtain $\gamma_1 \simeq -1.8$, $\gamma_2 \simeq 0.8$, $r_b \simeq 3.5 \ h^{-1}$ Mpc and $r_c \simeq 25 \ h^{-1}$ Mpc. This expression follows naturally the excess correlation (*shoulder*), which is observed at $r \sim r_o$ in virtually all optical galaxy samples, both in spatial and angular analyses (e.g., Shanks et al. 1989; Maddox et al. 1990), and provides a good fit to the data in the whole range 1–25 h^{-1} Mpc [i.e., where $\xi(r) > 0$]. Given the fact that the data used here differ slightly from those used by G91, due to the elimination of the higher absorption outer region, we shall reestimate the two-point correlation function. We apply the standard estimator $\xi(r) = fGG/GR - 1$, where GG is the number of galaxy pairs with separations in the range $r \pm dr$. This is normalized by the number of cross pairs GR between the data and a set of points distributed randomly within the survey boundaries, with density f times that of the galaxies. Errors on $\xi(r)$ are evaluated through bootstrap resampling (a more advanced procedure will be adopted for all other measures in this work; see § 3.4). We anticipate that, at least for $r \leq 10 \ h^{-1}$ Mpc, the usual fit to $\xi(r)$ with a single power law is still widely acceptable, and since $10 \ h^{-1}$ Mpc is the scale where we stop the high-order analysis presented here, we can safely use the simple $\xi(r) = (r_o/r)^\gamma$ model.

This will be seen in more detail by considering the expected number of galaxies in a sphere of radius R centered on a galaxy, which reads

$$\langle N_R \rangle = n V_R [1 + K_1 \xi(R)] \tag{3.1}$$

[here n is the galaxy number density within the sample, $V_R = (4\pi/3)R^3$, $K_1 = 3/(3 - \gamma)$]. Taking then a galaxy set Poisson distributed, V_R is expected to contain $N_{e,R} = n V_R$ galaxies. Accordingly, the value of the correlation length r_o can be obtained from the integral counts:

$$r_o = r \left\{ \left[\frac{\langle N_r \rangle}{\langle N_{e,r} \rangle} - 1 \right] \frac{1}{K_1} \right\}^{1/\gamma} . \tag{3.2}$$

To evaluate $N_{e,r}$ we created a large number ($\mathcal{N} \sim 10^3$) of artificial samples, whose points are distributed at random in the same volume occupied by the actual sample; $N_{e,r}$ is estimated by averaging among the numbers $\tilde{N}_{e,r}^{(i)}$ ($i = 1, \dots, \mathcal{N}$) of points in each V_R. In this way also boundary effects (arising when the radius R of a sphere of volume V_R exceeds the distance of galaxy from the boundary of the sample volume) are taken into account. In the next paragraphs we shall need also $(N_{e,r})^n$. To take into account boundary effects this is obtained directly from the averages

$$\frac{\sum_{i=1}^{\mathcal{N}} (\tilde{N}_{e,r}^{(i)})^n}{\mathcal{N}} .$$

If $\xi(r)$ has a perfect power law behavior, the value of r_o obtained from equation (3.2) is independent of r. On the contrary, a two-power-law behavior leads to expecting slight variations of r_o even in the 1–10 h^{-1} Mpc scale range. In § 4 we

shall give the values of r_o and their expected errors for different r's. Then we shall compare such values and errors with the expectation that r_o is constant (usual model) and with the r_o scale dependence predicted by the two-power-law expression.

The conclusions drawn from this comparison will both confirm the robustness of the two-power-law expression, and show that it is safe to keep to the well established expression $\xi(r) = (r_o/r)^\gamma$, for $r < 10 \ h^{-1}$ Mpc.

3.2. n-Point Function Expressions

The so-called *hierarchical model* (HM) expressions for the three- and four-point connected functions read

$$\xi_{012}^{(3)} = Q[\xi_{01}\xi_{12} + \xi_{12}\xi_{20} + \xi_{20}\xi_{01}] , \tag{3.3}$$

$$\xi_{0123}^{(4)} = R_a[\xi_{01}\xi_{12}\xi_{23} + \text{sym (12 terms)}]$$
$$+ R_b[\xi_{01}\xi_{02}\xi_{03} + \text{sym (4 terms)}] . \tag{3.4}$$

In the case of the four-point function we meet the simplest example of different topologies within a function of a given order. The following graphs are given to facilitate the understanding of more complicated cases we shall have below:

$$\xi_{0123}^{(4)} = R_a \times [\sqcap + \sqsubset + \sqcup + \sqsupset + \sldiag + \diagbox + \diagbox$$
$$+ \diagbox + \bowtie + \bowtie + \bowtie + \bowtie] + R_b \times [\diagbox + \diagbox + \diagbox + \diagbox] .$$

Let us also consider here the so-called *Kirkwood superposition* (KW) expressions. They are obtained from the assumption that

$$1 + \xi_{dis, x_1, \dots, x_n}^{(n)} = \prod_{i < j = 1}^{n} (1 + \xi_{ij}) . \tag{3.5}$$

In the case of the three-point function this yields a simple form for the connected function:

$$\xi_{012}^{(3)} = \xi_{01}\xi_{12} + \xi_{12}\xi_{20} + \xi_{20}\xi_{01} + \xi_{01}\xi_{12}\xi_{20} . \tag{3.6}$$

We shall indicate as *generalized Kirkwood* (GK) expression the one obtained from KW by setting two coefficients in front of the terms $\propto \xi^2$ and $\propto \xi^3$:

$$\xi_{012}^{(3)} = Q[\xi_{01}\xi_{12} + \text{sym}] + Q'\xi_{01}\xi_{12}\xi_{20} . \tag{3.7}$$

This gives back the HM expression for $Q' = 0$.

In the case of the four-point function we have some more complications. The KW expression now contains terms $\propto \xi^3$, $\propto \xi^4$, $\propto \xi^5$, and $\propto \xi^6$. Terms $\propto \xi^3$ and $\propto \xi^4$ can be still handled with standard techniques. Higher powers give rise to statistical problems that will be discussed elsewhere. Lt us rather give the GK expression that we shall need in this analysis:

$$\xi_{0123}^{(4)} = R_a \times [\sqcap + \cdots + \sldiag + \cdots + \bowtie + \cdots]$$
$$+ R_a \times [\diagbox + \cdots]$$
$$+ R_c \times [\diagbox + \diagbox + \diagbox + \diagbox + \diagbox + \diagbox + \diagbox$$
$$+ \diagbox + \bowtie + \bowtie + \bowtie + \bowtie]$$
$$+ R_d \times [\square + \bowtie + \bowtie] .$$

The HM expression was widely tested in the analysis of two-dimensional samples. There, however, a generalized Limber equation is needed to pass from two to three dimensions (Peebles & Groth 1977; Sharp, Bonometto, & Lucchin 1984). Schaeffer and his coworkers (see, e.g., Balian & Schaeffer 1989, and references therein) performed a great deal of work on HM expressions, analyzing its relation with scaling and

showing that the observed void distribution is compatible with $Q_n = (2^{n-1}/n)^a$ (here the constants appearing in all order connected functions are indicated by Q), with $a \sim 1 \pm 1$.

KW expressions were used to fit data in different physical contexts, as gas dynamics. However, KW expressions for n-point functions can be expected to hold if galaxies are related to very high peaks of an underlying distribution (Matarrese, Lucchin, & Bonometto 1986; Szalay 1988; Borgani & Bonometto 1990). This is exactly true for underlying Gaussian distributions, pure ϑ thresholds and in the very high peak limit. If such conditions are not verified quite cumbersome expressions hold which are difficult to handle. GK expressions can be considered to be reasonable approximations to such expressions.

It is also important to outline that, in the case of the three-point function Tóth, Hollósi, & Szalay (1989) have shown that, if ξ has the standard power-law form $\propto r^{-\gamma}$, the cubic terms should project into a term of the form

$$w_{012} = Q' L_\gamma \frac{w_{01} w_{12} w_{20}}{\theta_{01} \theta_{02} \theta_{20}} ,$$

with

$$L_\gamma = \pi^2 \left[\int_{-\infty}^{\infty} dx (1 + x^2)^{-\gamma/2} \right]^{-3}$$

and this should be detectable at small angles θ.

3.3. Moment Analysis

In equation (3.2) we showed how $C_1 = \langle N_r \rangle / \langle N_{e,r} \rangle$ is related to r_o. Here we shall consider the moments

$$C_n = \frac{\langle N_r^n \rangle}{\langle N_{e,r}^n \rangle} \tag{3.8}$$

and the mixed-exponent quantities

$$C_{n,m} = \frac{\langle N_r^n \rangle}{\langle N_{e,r}^m \rangle} , \tag{3.9}$$

which are necessary to subtract the shot-noise due to the point structure of the sample. We shall also use the quantity

$$\tilde{C}_1 = C_1 - 1 .$$

We shall report here the relations connecting C_n and $C_{n,m}$ to the constants Q and R. For the HM case, such relations can be obtained from Peebles (1980; § 59) and from Sharp et al. (1984). A generalization to the GK case, for the three-point function, is provided by Pons, Iovino, & Bonometto (1991). Here we shall give also a further extension including the GK case for the four-point function.

In the case of the three-point function we have that

$$Q = \frac{C_2 - C_{1,2} - c_1 \tilde{C}_1 - 1}{c_2 \tilde{C}_1^2 + q c_3 \tilde{C}_1^3} . \tag{3.10}$$

Here $q = Q'/Q$ is the ratio between the coefficients of the cubic and quadratic terms. Furthermore

$$c_1 = \frac{J_2}{K_1} + 2 , \quad c_2 = 1 + 2 \frac{K_2}{K_1^2} , \quad c_3 = \frac{H}{K_1^3} \tag{3.11}$$

with

$$J_2 = \left(\frac{3}{4\pi}\right)^2 \int_0^1 d^3x_1 d^3x_2 x_{12}^{-\gamma} = \frac{9 \times 2^{\gamma_3}}{\gamma_3 \gamma_4 \gamma_6} \tag{3.12}$$

(x_{ij} is the distance between the integration points x_i and x_j, $\gamma_s = s - \gamma$), K_1 is the same quantity appearing in equation (3.1),

$$K_2 = \left(\frac{3}{4\pi}\right)^2 \int_0^1 d^3x_1 d^3x_2 (x_1 x_{12})^{-\gamma} , \tag{3.13}$$

while

$$H = \left(\frac{3}{4\pi}\right)^2 \int_0^1 d^3x_1 d^3x_2 (x_1 x_2 x_{12})^{-\gamma}$$

$$= \frac{27}{8\pi\gamma_2} \int_0^1 dx_1 x_1^{\gamma_1} \int_0^1 dx_2 x_2^{\gamma_1} (|x_1 + x_2|^{\gamma_2} - |x_1 - x_2|^{\gamma_2})$$

$$\tag{3.14}$$

(for $\gamma = 1.51$, it is $H = 1.96$).

In the case of the four-point function the situation is slightly more complicated. What is evaluated is the combination

$$R_4 = R_a + \rho R_b$$

$$= \frac{C_3 + 2C_{1,3} - 3\tilde{C}_3 - 3c_4 \tilde{C}_2 - 3(c_1 - 2)\tilde{C}_1 - 1}{c_a \tilde{C}_1^3 + 3\sigma \tilde{C}_1^4} \tag{3.15}$$

of the *hierarchical* coefficients R_a and R_b. They can be corrected for the contribution of a quartic term (GK) by the second term at the denominator, where

$$\sigma = c_c \frac{R_c}{R_4} + c_d \frac{R_d}{R_4} . \tag{3.16}$$

In equation (3.15), $\tilde{C}_2 = C_2 - C_{1,2} - c_1 \tilde{C}_1 - 1$, $\tilde{C}_3 = C_{2,3} + C_2 - C_{1,2} - C_1$, and

$$c_4 = \frac{J_3}{2K_2 + K_1^2} . \tag{3.17}$$

Here

$$J_3 = \left(\frac{3}{4\pi}\right)^3 \int_0^1 d^3x_1 d^3x_2 d^3x_3 (x_{12} x_{23})^{-\gamma} . \tag{3.18}$$

Furthermore

$$c_a = 6\left(\frac{K_{3a}}{K_1^3} + \frac{K_2}{K_1^2}\right) \tag{3.19}$$

enters in the definition of the coefficients

$$\rho = \frac{1}{c_a}\left(\frac{K_{3b}}{K_1^3} + 1\right) \tag{3.20}$$

$$c_c = \frac{1}{c_a}\left(H + \frac{H_{c1}}{K_1} + 2\frac{H_{c2}}{K_1}\right) , \quad c_d = \frac{H_d}{c_a K_1} .$$

Such coefficients contain the new integrals

$$H_{c1} = \left(\frac{3}{4\pi}\right)^3 \int_0^1 d^3x_1 d^3x_2 d^3x_3 (x_1 x_{12} x_{13} x_{23})^{-\gamma}, \tag{3.21}$$

$$H_{c2} = \left(\frac{3}{4\pi}\right)^3 \int_0^1 d^3x_1 d^3x_2 d^3x_3 (x_1 x_3 x_{13} x_{23})^{-\gamma}, \tag{3.22}$$

$$H_d = \left(\frac{3}{4\pi}\right)^3 \int_0^1 d^3x_1 d^3x_2 d^3x_3 (x_1 x_{12} x_3 x_{23})^{-\gamma}. \tag{3.23}$$

Suitable geometrical choices can reduce the integration from nine to six variables. At that stage we perform a Monte Carlo
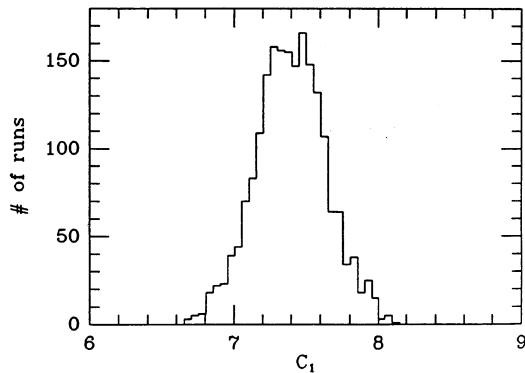
FIG. 1.—Example of a result of the Monte Carlo analysis used to estimate the errors on the parameters. Here the nearly Gaussian distribution of C_1 is shown.
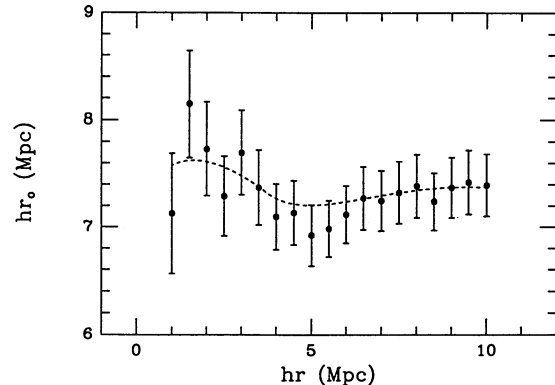

FIG. 2.—Clustering length deduced averaging over spheres of radius r and approximating $\xi(r)$ with a single power law with slope $\gamma = 1.51$. The dashed line is the expected behavior if the two–power-law model holds.

integration using $\sim 2,000,000$ random points. For $\gamma = 1.51$ we obtain

$$H_{c1} = 14.6 \,, \quad H_{c2} = 2.35 \,, \quad H_d = 4.19 \,. \qquad (3.24)$$

3.4. Error Analysis

Besides estimating C's and derived quantities it is essential to work out the expected errors on such estimates. To this aim we followed a pattern which is a generalization of the bootstrap procedure.

We selected at random subsamples made of half galaxies of the sample as centers of spheres, while all galaxies were kept as possible *companions*. C's were then estimated using a large number (2000) of independent subsamples. For the sake of example, in Figure 1 we report an histogram representing the distributions of the values obtained for C_1 with $rh = 3$ Mpc. All histograms have a similar bell shape. The standard deviations ΔC evaluated from such distributions are therefore an *overestimate* of the error for each C.

Starting from the C averages, we worked out r_o, Q, and R_4, according to equations (3.2), (3.10), and (3.15). Error propagation was studied numerically, by taking random points in the intervals $C \pm \Delta C$. All errors given below and in the figures correspond to 3 σ, estimated in the above fashion.

4. RESULTS

4.1. Two-Point Correlation Function

In Figure 2 we give the behavior of the correlation length r_o estimated from the integral counts with $\gamma = 1.51$, using (3.2). This value of γ is obtained from a direct fit to the two-point function, discussed here below. As already outlined, the standard single power-law behavior must be consistent with equal r_o-values for all r's.

An inspection of Figure 2 then shows that the value of r_o is fairly stable in the range 1–10 h^{-1} Mpc, so that the chosen approximation is justified. However, there is no way to fit a single horizontal straight line through all error bars, which represent 3 standard deviations. (Let us also remind that points and error bars at different r-values are not statistically independent.)

The dashed line, instead, is the result of integrating the two power-law model (with the observed slopes and amplitudes), to obtain N_r, and then r_o from equation (3.2). The global trend of the points and, in particular, the minimum around $4 \, h^{-1}$ Mpc, is thus coherent with a two–power-law behavior. What pre-

vious inspection implicitly suggested to consider as an excess noise is now the expected behavior.

In Figure 3 we show directly the plot of the two-point correlation function estimated on the VLS. This plot confirms the findings of G91, although obtained without clipping out the area of higher extinction. In particular, the function $1 + \xi(r)$ is well modeled with two power laws. Here a formal least-square fit in the two ranges gives logarithmic slopes, respectively, $\gamma_1 = 1.44 \pm 0.03$ and $\gamma_2 = 0.76 \pm 0.06$. With respect to G91 the small-scale slope has decreased by ~ 0.3: this is most probably due to the exclusion from the new sample of the very rich Perseus cluster at RA $\sim 3^h 2$.

It is also clear that fitting $\xi(r)$ with a single power law is still a sufficient approximation for $r < 10 \, h^{-1}$ Mpc. From this fit, in the range 1–10 h^{-1} Mpc, we have the slope $\gamma = 1.51 \pm 0.04$; the correlation length is $r_o = 7.4 \pm 0.7 \, h^{-1}$ Mpc.

Altogether, while the two-power law model seems to be in better agreement with data, still a single power-law fit to $\xi(r)$, for $r < 10 \, h^{-1}$ Mpc, can be considered to be valid.

4.2. Three- and Four-Point Functions

In Figure 4 we plot the values of Q obtained for $Q' = 0$ (*open circles*) and for $Q'/Q = 1$ (*filled circles*). The latter case provides
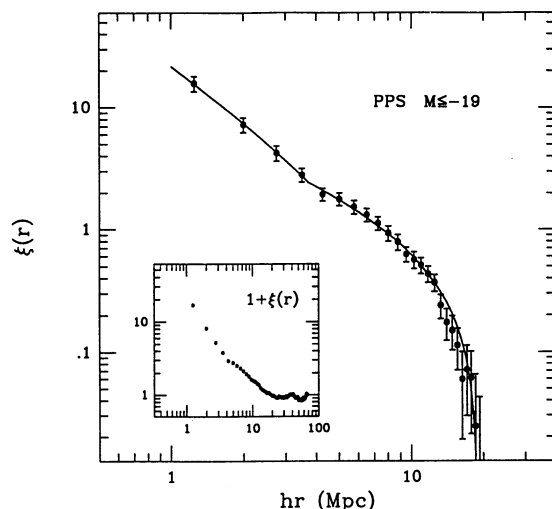

FIG. 3.—Two-point correlation function estimated from the $M \leq 19$, $d \leq 79 \, h^{-1}$ Mpc volume limited sample (VLS). The inset shows $1 + \xi(r)$.
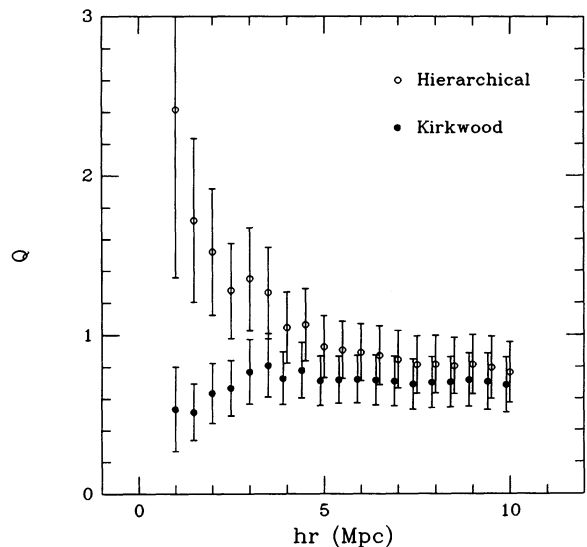
FIG. 4.—Coefficient Q for the three-point function of the VLS, as obtained from the hierarchical model expression (*open circles*), and assuming also a cubic term still with coefficient Q (*filled circles*).

FIG. 6.—Same as Fig. 4, but for the mLS

results which depend less on the radius r of the averaging sphere and can be fitted by a single straight horizontal line. The value of Q obtained for $Q'/Q = 1$ is, however, different from unity. Therefore the distribution is not KW.

Assuming HM expression and considering spheres of 3–5 h^{-1} Mpc, values $Q \sim 1$, with an error $\sim 25\%$, are recovered.

The situation is even more remarkable in the case of the four-point function (see Fig. 5). Here we plot the outputs which correspond to $R_c/R = R_d/R = 0$ and to $R_c/R = R_d/R = 1$. The latter case gives results independent from r. In the former case, instead, the dependence on r is quite strong. Here again, in the interval 3–5 h^{-1} Mpc, assuming an HM expression, we have $R_4 \sim 1.3$ with an error of $\sim 30\%$.

In order to illustrate the advantages of using VLS, in Figure 6 we provide the outputs concerning the three-point function
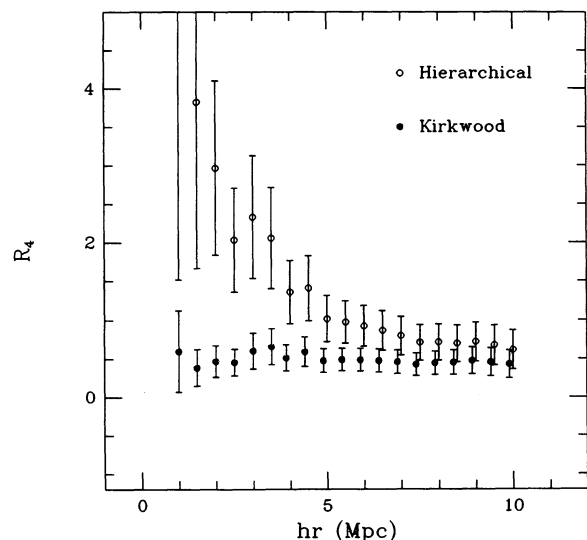
also for mLS. The random sample, in this case, was suitably modulated using the observed selection function. In spite of the larger number of galaxies (~ 3000 vs. ~ 900), error bars are not significantly reduced. There is however a point which can be outlined here: in the small scale range, the value of Q might be consistent with pure KW.

Similar conclusions can be drawn in the analysis of the four-point function.

5. DISCUSSION

The main results of the present work are the presence of a cubic term in the three-point function and of a quartic term in the four-point function. Results on three- and four-point functions were first obtained from angular catalogs, like the Zwicky and Lick compilations (Groth & Peebles 1977; Fry & Peebles 1978; Sharp et al. 1984). Recently Szapudi, Szalay, & Boschan (1992) used the moment method on the two-dimensional Lick sample and Meiksin, Szapudi, & Szalay (1992) extended the same method to the *IRAS* galaxy sample. The range of angles considered in analyses of Zwicky and Lick counts corresponds to distances r in the range $0.5 \lesssim hr/\mathrm{Mpc} \lesssim 8$. Points above 4 h^{-1} Mpc are however given with very large errors and only in the interval $1.5 < hr/\mathrm{Mpc} < 4$ there is some contradiction between the estimate of this work and the results obtained from Zwicky and Lick counts.

In the work of Meiksin et al. (1992), instead, in the case of the three-point function, it is shown that observations are compatible with $Q^3 - Q' = 0.9 \pm 0.4$ (for $\gamma = 2$) or with $Q^3 - Q' = 1.7 \pm 1$. (for $\gamma = 1.6$). Such results do not disagree with the outputs of our analysis.

A comparison with the results from other three-dimensional samples is not easily made. Huchra, Davis, & Geller (1979, unpublished) are reported by Peebles (1980) to have estimated $Q = 0.80 \pm 0.07$ on the CfA 14.5 sample. On the basis of an unpublished sample selected by Rood, Peebles (1981) found agreement with HM expressions for the three-point function and gave $Q = 0.68 \pm 0.05$. However, the triplets from which such measure is taken are situated at the vertices of triangles whose shortest side has a projected length never exceeding 1.3 h^{-1} Mpc. (The longest projected triangle sides do not exceed

FIG. 5.—Coefficient R_4 (linear combination of R_a and R_b, see text) for the VLS, as obtained for the hierarchical model expression (*open circles*) and assuming quartic terms with coefficients equal to $R_a = R_b$ (*filled circles*).
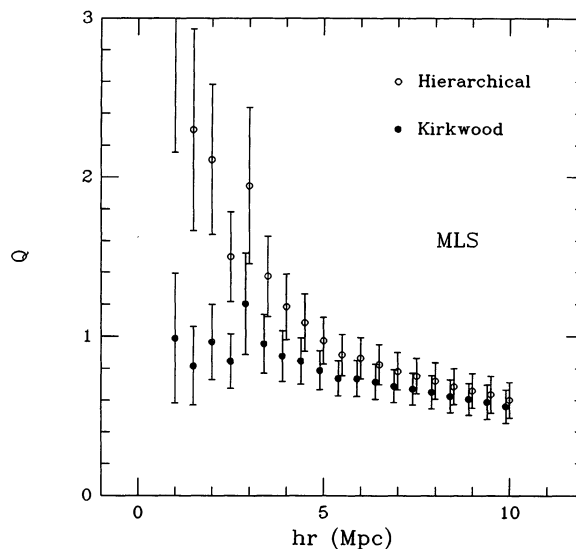
4.8 h^{-1} Mpc.) It is therefore hard to say whether a real contradiction exists with the present result.

More recently Gaztañaga (1992), performing counts in random cells of diameters ranging from 2 to 22 h^{-1} Mpc, based on volume limited samples worked out from the CfA$_{14.5}$ catalog and on the Southern Sky Redshift Survey (da Costa et al. 1991), obtained significant results on three- and four-point functions. The paper shows a reasonable agreement between data and a hierarchical clustering model. However, as is often done, errors are estimated using formal least-square fits and their quantitative significance can be hardly evaluated. In spite of that, formal errors can be a fair indication of the noise present in the data: a ∼20% error is present for the output on Q, while such error approaches ∼90% for the output on R_4. (Although not clearly stated, these errors seem to account for a 2 σ formal error level.) Assuming hierarchical clustering greatly lowers the above error level. No quantitative comparison is however provided with other clustering models. Altogether it is hard to say whether such results really disagree with those of this paper, although this impression can hardly be avoided.

The range of distances considered here shows substantial overlaps only with the one taken in Gaztañaga's paper. The significance of the distance range we took can be better focused if we consider that a critical galaxy separation $\Delta R_c \simeq 1$–$2 \, h^{-1}$ Mpc, related to pairwise velocities, exists, which bears a triple meaning: (i) for $\Delta R \gtrsim \Delta R_c$ the distance component along the z direction begins to be reliably appreciable from z measures. (ii) For $\Delta R \gtrsim \Delta R_c$ the relative displacement of galaxies in a Hubble time is not so heavy to destroy clustering features related to matter distribution at the time of galaxy formation. (iii) Accordingly, for scales $R \lesssim \Delta R_c$, we are in a fully nonlinear regime (both from the Eulerian and the Lagrangian points of view).

The scale ΔR_c is obtained considering the component along the z direction of the average galaxy pairwise velocity. Pairwise velocities are very mildly dependent on galaxy separation and their average component along z is estimated to be $\sigma_z \sim$ 150–200 km s^{-1}. Actual velocities are smaller for field galaxies; statistical corrections can be applied however to galaxies in clusters; to our aims corrections are equivalent to reduce σ_z by a factor ∼2. Accordingly

$$\frac{\Delta R_c}{h^{-1} \text{ Mpc}} \simeq \frac{\sigma_z}{100 \text{ km s}^{-1}} \qquad (5.1)$$

lays between 1 and 2.

Owing to point (i), it would have scarce significance to use a redshift sample to estimate clustering properties at distances below ∼1 h^{-1} Mpc, while such estimates are fairly reliable above ∼2 h^{-1} Mpc. This corresponds to the range of distances above which our analysis is effective.

The point (ii) is related to precise physical arguments. As was discussed in § 1, biased thories of galaxy origin predict KW expressions. Such theories are however intrinsically *geometrical*: they do not contemplate dynamical effects over scales exceeding the interior of matter density peaks. Accordingly, bias predictions can only be tested if the displacement from the place where galaxies formed can be neglected. Once more, this begins to be allowed above ΔR_c.

In turn this relates to the onset of a nonlinear dynamical situation. The presence of large velocities over scales of a few Mpc witnesses that, at the present time, nonlinearity has extended from the distance range characterizing the interior of density peaks, up to some Mpc. It is therefore fair to consider three distance ranges: at $R < \Delta R_c$ there is a fully nonlinear regime; for

$$\Delta R_c < R < \left(\frac{2}{b^2}\right)^{1/\gamma} r_o \qquad (5.2)$$

(b is the linear bias factor) from the Eulerian point of view, we are in nonlinear regime, i.e., density fluctuations $\delta\rho/\rho$ exceed unity; however the dynamics can still be quasi-linear from the Lagrangian point of view and can be treated with perturbative Lagrangian techniques (e.g., Zel'dovich approximation and its improvements, see Buchert 1992). Above $r_{nl} = (2/b^2)^{1/\gamma} r_o$ we attain a fully linear regime. For $b \simeq 2.4$ it turns out that $r_{nl} \simeq r_b$, which is the point where there appears to be a change of slope in the two-point function, according to the two–power-law model.

In order that KW expression significantly differs from HM expression, scales less than r_o ought to be considered. Between ΔR_c and r_o we can therefore expect to detect features related to biased galaxy formation. Peculiar velocities erase such features at $R < \Delta R_c$.

According to these considerations our findings on three- and four-point functions become essentially consistent both with previous measures of higher order functions and with G91 findings on the shape of the two-point function. In our opinion it is reasonable to conclude that our outputs show the gradual passage from a linear to a nonlinear regime and support a biased theory of galaxy formation with a bias factor for bright galaxies ∼2–2.5.

REFERENCES

Balian, R., & Shaeffer, R. 1989, A&A, 226, 373
Borgani, S., & Bonometto, S. A. 1990, ApJ, 348, 398
Bouchet, F. R., & Lachiéze-Rey, M. 1986, ApJ, 302, L37
Buchert, T. 1992, MNRAS, 254, 729
Calzetti, D., Giavalisco, M., & Meiksin, A. 1992, ApJ, 398, 429
da Costa, L. N., Pellegrini, P., Davis, M., Meiksin, A., Sargent, W., & Tonry, J. 1991, ApJS, 75, 935
Davis, P., & Peebles, P. J. E. 1983, ApJ, 267, 46
Dekel, A., & Aarseth, S. J. 1984, ApJ, 340, 625
Fry, J. 1984, ApJ, 277, L5
Fry, J., & Peebles, P. J. E. 1978, ApJ, 221, 19
Gaztañaga, E. 1992, ApJ, 398, L17
Giovanelli, R., & Haynes, M. P. 1985, AJ, 90, 2445
———. 1988, in IAU Symp. 130, Large-Scale Structure in the Universe, ed. J. Adouze, M. C. Pelletan, & A. Szalay (Dordrecht: Reidel), 113
———. 1989, AJ, 97, 633
———. 1993, ApJ, in press

Giovanelli, R., Haynes, M. P., & Chincarini, G. 1986, ApJ, 300, 77
Groth, E. J., & Peebles, P. J. E. 1977, ApJ, 217, 385
Guzzo, G., Iovino, A., Chincarini, G., Giovanelli, R., & Haynes, M. 1991, ApJ, 382, L5 (G91)
Haynes, M. P., & Giovanelli, R. 1988, in Large-Scale Motions in the Universe, ed. V. C. Rubin & G. V. Coyne, S.J. (Princeton: Princeton Univ. Press), 31
Haynes, M. P., Giovanelli, R., Starosta, B. M., & Magri, C. 1988, AJ, 95, 607
Iovino, A., Giovanelli, R., Haynes, M. P., Chincarini, G., & Guzzo, G. 1993, MNRAS, in press
Maddox, S. J., Efstathiou, G., Sutherland, W. J., & Loveday, J. 1990, MNRAS, 242, 43P
Mandelbrot, P. P. 1982, The Fractal Geometry of Nature (San Francisco: Freeman)
Martinez, V. J., Jones, B. J. T., Dominguez, R., & van de Weygaert, R. 1990, ApJ, 357, 50
Matarrese, S., Lucchin, F., & Bonometto, S. A. 1986, ApJ, 310, L21
Meiksin, A., Szapudi, I., & Szalay, A. S. 1992, ApJ, 394, 87

Paladin, G., & Vulpiani, A. 1984, Nuov. Cimento Lett., 41, 82
————. 1987, Phys. Rep., 156, 147
Peebles, P. J. E. 1980, The Large-Scale Structure of the Universe (Princeton: Princeton Univ. Press)
————. 1981, Proc. 10th Texas Symp., ed. R. Ramaty & F. C. Jones (New York: NY Acad. Sci.), 287
Peebles, P. J. E., & Groth, E. J. 1977, ApJ, 217, 385
Pietronero, L. 1987, Physica, A144, 257
Pons, M. J., Iovino, A., & Bonometto, S. A. 1991, AJ, 102, 495
Ramond, P. 1981, Field Theory: A Modern Primer (Reading: Benjamin Commings)
Schaeffer, R. 1984, A&A, 134, L15

Shanks, T., Hale-Sutton, D., Fong, R., & Metcalfe, N. 1989, MNRAS, 237, 589
Sharp, N., Bonometto, S. A., & Lucchin, F. 1984, A&A, 130, 7
Szalay, A. S. 1988, ApJ, 333, 21
Szapudi, I., Szalay, A. S., & Boschán, P. 1992, ApJ, 390, 350
Tóth, G., Hollósi, J., & Szalay, A. S. 1989, ApJ, 344, 75
Valdarnini, R., Borgani, S., & Provenzale, A. 1992, ApJ, 394, 422
Wegner, G. A., Haynes, M. P., & Giovanelli, R. 1994, in preparation
White, S. D. M. 1979, MNRAS, 186, 145
Zwicky, F., Herzog, E., Karpowicz, M., Kowal, C. T., & Wild, P. 1961–1968, Catalogue of Galaxies and of Clusters of Galaxies, Vols. 1–6 (Pasadena: Caltech) (CGCG)