Clustering properties from finite galaxy samples

A. Provenzale, 1 L. Guzzo 2 and G. Murante 1

1 Istituto di Cosmogeofisica del CNR, Corso Fiume 4, I-10133 Torino, Italy
2 Osservatorio Astronomico di Brera, Sezione di Merate, Via Bianchi 46, I-22055 Merate, Italy

Accepted 1993 August 3. Received 1993 June 10; in original form 1993 March 4

ABSTRACT
The shape of the density correlation function \( g(r) = 1 + \xi(r) \) provides useful information on the properties of galaxy clustering. In particular, a power-law form of \( g(r) \propto r^{-\gamma} \) indicates the presence of a scaling regime with a correlation dimension \( D_2 = 3 - \gamma \), while a flattening of \( g(r) \) is a signature of homogeneity. A relevant issue, then, is whether the standard treatment of boundary effects may affect the behaviour of \( g(r) \) at large separations, for example by inducing a spurious flattening and/or a spurious scaling regime. To answer this question, we consider different types of fractal universe generated by the cosmological \( \beta \)-model and treat them in the same way as the observational data. The results of the analysis show that for a pure fractal universe, with no homogeneity at large separations, the function \( g(r) \) does not display any large-scale flattening, even when boundary effects are corrected by the standard normalization to an ensemble of random points distributed within the sample volume. Analogously, no spurious scaling regime is generated. This demonstrates that the flattening of \( g(r) \) detected in the analysis of the CfA and PP surveys does really indicate a transition to homogeneity; consequently, a pure fractal universe is not consistent with the available observational data. We then use a multi-scale \( \beta \)-model to reproduce the shape of \( g(r) \) observed for the PP survey; an intermediate scaling regime with \( D_2 = 2 \) in the approximate range 4–25 \( h^{-1} \) Mpc is shown to be necessary to reproduce the observational results. As confirmed by recent \( N \)-body experiments, the small-scale regime has to be interpreted in terms of a truly fractal nature of the galaxy distribution, being associated with the outcome of ‘fully developed’ non-linear gravitational clustering. Conversely, we suggest that the intermediate regime has a topological origin, being possibly produced by nearly flat structures resulting from the weakly non-linear gravitational evolution of significant large-scale fluctuations in the initial spectrum.

Key words: methods: statistical – galaxies: clustering – large-scale structure of Universe.

1 INTRODUCTION
The homogeneity of the matter distribution on large scales is a fundamental postulate known as the ‘Cosmological Principle’ (Milne 1935; Peebles 1980). The main observational evidence supporting large-scale homogeneity is (a) the extreme uniformity of the cosmic microwave background (see e.g. Smoot et al. 1992; Wright et al. 1992), (b) the convergence of the optical dipole in, for example, the cluster distribution (Plionis & Valdarnini 1991; Sfaramella, Vettolani & Zamorani 1991), and (c) the rapid decay of the spatial two-point galaxy–galaxy and cluster–cluster correlation functions \( \xi_{gal}(r) \) and \( \xi_{clus}(r) \) at increasing separations, and the scaling with depth of the angular correlation function \( w(\theta) \) (Peebles 1980; Maddox et al. 1990). Analogously, the analysis of the clustering of rich clusters shows evidence of a space-filling distribution on sufficiently large scales (Borgani, Plionis & Valdarnini 1993). We also recall that homogeneity on large scales is a basic assumption implicit in the same definition of the correlation function \( \xi(r) \) (Peebles 1980; Pietronero 1987), which requires the average matter density in the Universe to be a well-defined quantity.
In the last 10 years, however, almost every new galaxy and cluster sample has revealed the existence of 'structures' (e.g. voids, walls, filaments and superclusters) at increasingly large scales, well beyond the standard 'galaxy correlation length' \( r_c = 5 \, h^{-1} \) Mpc. The observational data of a non-homogeneous distribution at these large separations stimulated several discussions on the appropriate size of a 'fair sample' of the Universe (see Einasto, Klypin & Saar 1986). In particular, the existence of large-scale structures has been related to a possible pure fractal nature of the galaxy distribution on all scales by Pietronero (1987) and Coleman, Pietronero & Sanders (1988), who found indications that the density correlation function \( g(r) = 1 + \xi(r) \) does not flatten up to at least \( 20 \, h^{-1} \) Mpc. The idea of a pure fractal universe is, however, in contrast with the evidence for an homogeneous galaxy distribution on large scales advanced, for example, by Martinez & Jones (1990) and Lemson & Sanders (1991) in the analysis of the CfA142 redshift sample, and by Guzzo et al. (1991, hereafter G91) in the analysis of the Perseus-Pisces Redshift Survey (hereafter PP; see Haynes & Giovanelli 1988). The two latter studies have quantitatively indicated a value of \( r_c = 30 \, h^{-1} \) Mpc for the homogeneity scale of the galaxy distribution, based on the flattening of \( g(r) \) beyond \( r_c \). The analysis by G91 has also indicated that \( g(r) \) seems to display two scaling (power-law) regimes characterized respectively by correlation dimensions \( D_2 = 1.2 \) at small separations (\(< 4 \, h^{-1} \) Mpc) and \( D_2 = 2 \) on intermediate scales (\(4 - 30 \, h^{-1} \) Mpc).

In all the above studies, the treatment of boundary effects plays a crucial role. In fact, this problem is particularly relevant in the analysis of redshift samples due to the peculiar shape of the survey volumes. The standard treatment of boundary effects is based on normalizing the galaxy correlation function by the correlation function of a set of random points distributed within the sample volume. A basic question, then, is whether this correction could systematically affect the form of the galaxy correlation function on scales comparable to the sample size.

A possible way to answer the above question is to analyse synthetic distributions with known scaling and statistical properties, and treat them in the same way as the observational data. This approach is applied in Lemson & Sanders (1991), through the use of both Levy flights and Voronoi tessellation models. In this work we want to address some of the issues discussed in that paper as well as more general questions, by using the more flexible (and hopefully realistic) cosmological \( \beta \)-model (Castagnoli & Provenzale 1990, 1991; Provenzale 1991), in the spirit of a fractal Monte Carlo simulation of the galaxy distribution. Using the \( \beta \)-model, we simulate universes with different clustering properties, considering pure hierarchical distributions over all scales as well as distributions with an upper homogeneity cut-off. We 'observe' these synthetic universes by extracting several finite samples with the same depth and shape as the volume-limited version of the PP survey used by G91, and address the following three questions. (1) Is the density correlation function \( g(r) = 1 + \xi(r) \) of a pure fractal distribution forced to a constant value at large separations by the standard treatment of boundary effects? (2) If not, how is the true homogeneity scale reflected by the flattening scale \( r_c \) of \( g(r) \)? (3) Is a scale-dependent distribution with \( D_2 = 1.2 \) at small scales and \( D_2 = 2 \) at intermediate scales consistent with the observational data? With regard to the last point, we discuss the possible origins of the intermediate scaling regime, and we introduce a simple model of small-scale fractal clustering superimposed on to larger scale 'walls' or 'sheets' capable of reproducing the observed scaling.

The remainder of this paper is structured as follows. In Section 2 we discuss the basic definitions of correlation function and correlation dimension and recall the standard methods for correcting boundary effects. In Section 3 we discuss the results of the analysis of simulated fractal distributions and consider a possible alternative origin for the scaling regime at intermediate scales. In Section 4 we give conclusions and perspectives.

## 2 THE TWO-POINT CORRELATION FUNCTION

The classic two-point correlation function \( \xi(r) \) is an important statistical descriptor of a point distribution (see e.g. Peebles 1980 for a detailed discussion of this approach). The correlation function of a point set can be defined as

\[
\delta P = \langle n(r) \rangle \left( 1 + \xi(r) \right) \delta V,
\]

where \( \delta P \) is the probability of finding a point in the volume \( \delta V \) placed at a distance \( r \) from another point of the sample. That is, \( \xi(r) \) is a two-point probability that measures the clustering properties of the sample; \( \xi(r) = 0 \) for a Poisson random distribution.

The definition of \( \xi(r) \) explicitly depends on the average density \( \langle n \rangle \) of the sample under study. If the latter is not a well-defined quantity, then the correlation function may depend on the size of the sample (see e.g. Pietronero 1987). In this case, one may define the quantity

\[
\Gamma(r) = g(r) / \langle n \rangle,
\]

where

\[
g(r) = 1 + \xi(r).
\]

The function \( \Gamma(r) \) does not depend on \( \langle n \rangle \); in this case no dependence on the sample size is observed even for pure fractal sets. The quantity \( \Gamma(r) \) was originally introduced with the aim of explaining the difference between the amplitudes of the galaxy and cluster correlation functions as being spuriously generated by the depth dependence of the average matter density in a pure fractal universe (Pietronero 1987; Coleman & Pietronero 1992); See also Lemson & Sanders (1991) for a comparison between \( \xi(r) \) and \( \Gamma(r) \) in the case of two different synthetic point distributions.

The function \( g(r) \) is related to the correlation integral \( C(r) \) defined by Grassberger & Procaccia (1983), i.e.

\[
C(r) = \frac{4\pi \langle n \rangle}{N} \int_0^r g(r') r'^2 \, dr',
\]

where \( N \) is the total number of galaxies in the sample. For a fractal point distribution,

\[
C(r) \propto r^{D_2},
\]

on small scales, where \( D_2 \) is the correlation dimension of the set under study.

From equation (5), it is easy to see that for a pure fractal set the function \( g(r) \) has a power-law dependence at all

© Royal Astronomical Society • Provided by the NASA Astrophysics Data System
separations:
\[ g(r) \propto r^{-\gamma}, \]
where \( \gamma = 3 - D_2 \). Conversely, for a random uniform distribution \( g(r) = 1 \) and \( D_2 = 3 \). For a distribution showing a transition to large-scale homogeneity, the function \( g(r) \) has a power-law behaviour on the scales of fractality and flattens to a constant value for increasing \( r \). In general, the value \( r_c \) for which \( g(r) \) or \( \Gamma(r) \) becomes constant provides an estimate of the scale beyond which the density fluctuations become negligible with respect to the average density. Note that a pure power-law behaviour of \( \xi(r) \) would not imply fractality on all scales, but rather a precise form of the transition from small-scale fractal behaviour [where \( 1 + \xi(r) - \xi(r) \) to the large-scale homogeneous regime.

It is important to recall that the two-point correlation function provides information only on the second-order moment of the probability distribution of the sample. In fact, the correlation function is the Fourier transform of the power spectrum \( P(k) \) of the density [or, apart from a Dirac \( \delta(k) \), the fluctuation] field defined by the smoothed galaxy distribution. Features observed in \( \xi(r) \) or \( g(r) \) are thus necessarily related to features present in \( P(k) \). In Fourier space, the transition from a scale-free clustering regime to a homogeneous distribution on larger scales corresponds to a transition from a power-law behaviour, \( P(k) \propto k^n \) (with \( n < 0 \)), at large \( k \) to a null or positive index for smaller wavenumbers (see e.g. Branchini, Guzzo & Valdarnini 1993).

The higher order moments of the probability distribution of the sample are related to the \( n \)-point correlation functions, to the spectrum of generalized fractal dimensions and to the correlations in the distribution of Fourier phases. Since we are herein interested only in the behaviour of the two-point correlation function, in this paper we consider only the case of monofractal distributions where a single value of the dimension (e.g. \( D_2 \)) is sufficient to characterize the sample. This approach is valid provided that we use point distributions with the appropriate value of \( D_2 \). For an introduction to the more general case of multifractal distributions and to the measurement of their properties, see e.g. Paladin & Vulpiani (1987), Martinez et al. (1990) and Borgani et al. (1993). A study of the multifractal properties of the PP Redshift Survey will be reported elsewhere.

The definition of the correlation function can be safely applied only when the set under study is physically limited to a finite region of space. For example, the correlation integral \( C(r) \) is widely used for determining the dimension of chaotic attractors, which are bounded objects in phase space (see e.g. Eckmann & Ruelle 1985). In the case of the galaxy distribution, however, the sample under study is a portion of a larger distribution; its boundlessness is due to observational limitations and not to its intrinsic nature. For this reason, boundary effects must be taken into account when computing the correlation function and/or the correlation integral. To this end, a few methods have been proposed. A first possibility is to discard, in the evaluation of \( g(r) \), all points that are closer than \( r \) to one of the edges of the sample. This method has two main disadvantages. First, it heavily limits the statistics and the range of scales over which the correlation function may be obtained. Secondly, the fraction of the sample contributing to the estimate of \( \xi(r) \) is decreasing with increasing \( r \). Such that only the central part of the sample is actually used at large separations (see discussion in Lemson & Sanders 1991).

An alternative method, and the one most commonly used, is based on normalizing the correlation function to a random sample with the same shape and selection function as the data set under study (e.g. Davis & Peebles 1983). That is, \( g(r) \) is obtained as

\[ g(r) = \frac{\Delta N_{sr}(r)}{\Delta N_{sr}(r)} \frac{2n_{sr}}{n_{sr} - 1}, \]

where \( \Delta N_{sr}(r) \) and \( \Delta N_{sr}(r) \) are respectively the number of independent galaxy–galaxy and galaxy–random pairs with separation between \( r \) and \( r + \Delta r \), and \( n_{sr} \) and \( n_{sr} \) are the expected numbers of pairs in the two cases. Typically, \( n_{sr} \gg n_{sr} \). This method has been used by G91 to obtain \( g(r) \) from the PP catalogue. In the next section we analyse synthetic samples obtained from simple fractal and geometric models of the galaxy distribution in order to verify whether the use of equation (7) may affect the behaviour of \( g(r) \) on large scales.

### 3 SIMULATING THE PROPERTIES OF GALAXY CLUSTERING

The results recently obtained by G91 indicate that galaxy clustering seems to be well described by two different scaling laws before the transition to homogeneity. In general, a power-law form of \( g(r) \) in a given range of scales suggests that it is possible to simulate the properties of the galaxy distribution on those scales by considering an appropriate scale-invariant model. One possible approach in the framework of fractal distributions is provided by the cosmological \( \beta \)-model (Castagnoli & Provenzale 1990, 1991; Provenzale 1991). This type of fractal cascade, based on the \( \beta \)-model of turbulence introduced by Frisch, Sulem & Nelkin (1978), provides a phenomenological description of the distribution of bright galaxies; the \( \beta \)-model can be used as a simplified ‘fractal machine’ in order to test data analysis methods and to verify the inferences obtained from the analysis of real samples. Along these lines, here we use the distributions generated through this algorithm for studying the role of boundary effects in the evaluation of \( g(r) \) at large separations. Note also that the \( \beta \)-model produces a monofractal distribution; a multifractal extension of this approach, designed for describing the process of non-linear gravitational clustering at small scales, has been discussed by Provenzale et al. (1992).

The simplest realization of the \( \beta \)-model is a three-dimensional point distribution generated by a sequence of breaking iterations starting from a parent object (e.g. a cube) with linear size \( L_0 \) divided into \( M \) smaller objects with linear size \( L_1 = L_0/m \), where \( M = m^3 \). Of the \( M \) ‘daughter’ objects, only \( N_0 \) remain active and are able to break again at the second iteration. The number of active objects is taken to be a random variable; its mean value \( \langle N_0 \rangle \) is related to the survival probability \( p \) of each object, i.e. \( \langle N_0 \rangle = pM \). This must be understood in the sense that each object has a probability \( p \) of breaking again (of ‘surviving’) and a probability \( 1 - p \) of becoming quiescent. A quiescent object does not generate any point at the end of the cascade. Physically, active objects correspond to positive density fluctuations which undergo gravitational clustering, while...
quiescent objects correspond to negative density fluctuations.

In the simplest version of the model, \( p \) is constant for all objects and for all iterations. By continuing the breaking cascade, a fractal distribution of points is obtained. In this simple case, the average fractal dimension of the distribution is given by

\[
D = \lim_{k \to \infty} \frac{\log \langle N_k \rangle}{\log m^k} = \frac{\log \langle N_k \rangle}{\log m} = \frac{\log pM}{\log m},
\]

(8)

where the index \( k \) denotes the \( k \)th iteration in the breaking cascade. The value of the fractal dimension is completely determined by the average number of surviving objects \( \langle N_k \rangle \), i.e. by the value of the survival probability \( p \). A further simplification would be to take \( N_k \) to be a constant (i.e. not a random variable); in this case one would reobtain the classic definition of the fractal dimension (i.e. not only an average dimension as in equation 8). The case with \( D = 3 \) corresponds to an homogeneous and space-filling point distribution \( (p=1 \text{ and } N_k = m^3) \); this is equivalent to a uniform random distribution. At the end of the iteration cascade, a small random displacement (Gaussian-distributed with a standard deviation that is 1/4 of the size of the last-iteration cell) is added to the position of each point, in order to simulate the effects of small measurement errors and noise. Since the distribution generated by the \( \beta \)-model is a monofractal point set, in the present case the fractal dimension \( D \) defined by expression (8) coincides with the correlation dimension \( D_2 \) defined by equation (5).

### 3.1 Monofractal dust with \( D_2 = 1.2 \)

For the purpose of the analysis, we consider both pure self-similar point distributions and distributions with an upper homogeneity cut-off. Let us start with a pure fractal distribution with \( D_2 = 1.2 \). This is obtained by choosing \( m=2 \) and \( p=2^{-1.8} \) in the construction of the \( \beta \)-model. The correlation function \( g(r) \) for such a distribution has a pure power-law behaviour \( g(r) \propto r^{-1.8} \).

In order to verify whether the normalization to an homogeneous random distribution may induce a spurious flattening of \( g(r) \), we consider a set of PP-like samples extracted from pure \( D = 1.2 \) fractal distributions, and we correct for the boundary effects by using the procedure given by equation (7). The simulated sample is a wedge of \( 90^\circ \times 45^\circ \), with a depth of 79 \( h^{-1} \) Mpc, and is geometrically equivalent to the \( M = -19 \) volume-limited sample used by G91 to estimate \( g(r) \) from the PP survey. A total of 1092 points is randomly extracted from the fractal distribution, in order to have the same statistics as in the real sample. Fig. 1 shows the density correlation function \( g(r) = 1 + \xi(r) \) as obtained from an average over 20 different realizations of the \( \beta \)-model with \( D_2 = 1.2 \). The error bars are given by the standard deviation over the 20 different realizations. These error bars are quite large owing to the large fluctuations of \( g(r) \) from one realization to another, which are induced by the huge inhomogeneity of a pure monofractal dust with

\[1\text{ Here } M, \text{ absolute magnitude, should not be confused with the } M \text{ in equation (8).}\]

Figure 1. The density correlation function \( g(r) = 1 + \xi(r) \) averaged over 20 realizations of a pure fractal universe with \( D_2 = 1.2 \). A PP-like slice from each distribution has been extracted and analysed. Boundary effects have been corrected by normalizing to a random catalogue distributed within the sample volume; the error bars are the standard deviations obtained from the 20 realizations.
$D_2 = 1.2$. However, none of the realizations considered here displays any spurious flattening up to separations comparable to the sample size. We also note that the correlation amplitude for this model is much larger than the observed one, due to the evidently too inhomogeneous distribution of the simulated sample; that is, a pure fractal universe with $D_2 = 1.2$ is not consistent with the observed flattening and with the amplitude of $g(r)$ obtained by G91.

The second question, then, is whether the observed flattening of the correlation function $g(r)$ provides a good measure of the true homogeneity scale. To address this issue, we consider a simulated PP-like sample with $D_2 = 1.2$ up to an homogeneity scale $L_h = 40 \ h^{-1}\ Mpc$, and with $D = 3$ (homogeneity) beyond $L_h$. The homogeneous regime is obtained by imposing $p = 1$ in the $\beta$-model cascade. Fig. 2(a) shows $g(r)$ as computed from an average over 20 different realizations of this model. The error bars are now much smaller, due to the fact that homogeneity on large scales induces much smaller fluctuations from one realization to another. The flattening of $g(r)$ is observed on scales of the order of 20–25 $h^{-1}\ Mpc$, roughly half the value of $L_h$. Note that the value of $L_h$ refers to the size of the breaking objects in the model, and thus to the diameter of a sphere centred on a generic galaxy of the distribution. Conversely, the scale indicated in the evaluation of the correlation function refers to the radius of the same sphere. Thus a flattening in $g(r)$ at $r_c \approx 20–25\ h^{-1}\ Mpc$ correctly indicates an homogeneity scale $L_h \approx 2r_c$ in the framework of the $\beta$-model. These results indicate that the scale of the flattening of $g(r)$ provides a reliable estimate of the radius of crossover to homogeneity, when this is smaller than the characteristic sample size. Note again, however, that the correlation amplitude for this simulated sample is still too high, indicating that a pure model with $D_2 = 1.2$ at small scales and homogeneity at larger scales cannot reproduce all the characteristics of the observed $g(r)$.

An interesting point concerns the possibility of estimating the homogeneity scale when this is close to the size of the sample. Fig. 2(b) shows the function $g(r)$ for a distribution with an homogeneity scale $L_h = 60\ h^{-1}\ Mpc$, as computed from a PP-like sample with a depth of 79 $h^{-1}\ Mpc$. A flattening is observed even in this case. However, a noticeably stronger anticorrelation on large scales [$g(r) < 1$] is produced due to the integral constraint on the two-point correlation function (Peebles 1980), which can be expressed as

$$4\pi \int_0^{r_{\text{sample}}} \xi(r) r^2 dr = 0.$$  \tag{9}

Note, however, that it is still possible to infer a fair estimate of the homogeneity scale using the flattening radius of $g(r)$ even in the presence of a strong integral constraint (i.e. with a catalogue that is a fair sample of the Universe only marginally). On the other hand, looking only at a plot of $\xi(r)$ the effect of the integral constraint would have been that of forcing a break in the power law on scales smaller than the true homogeneity radius, i.e. at that separation where $g(r)$ first becomes 1. This effect has been widely discussed by Pietronero (1987) in terms of the function $\Gamma(r)$.

As an extreme example, Fig. 2(c) shows the function $g(r)$ for a distribution with $L_h = 75\ h^{-1}\ Mpc$; thus in this case $L_h$ is as large as the size of the sample. Even though some

Figure 2. Density correlation functions $g(r) = 1 + \xi(r)$ averaged over 20 realizations of a fractal universe with $D_2 = 1.2$ and different upper homogeneity thresholds $L_h$. The values of $L_h$ are (a) $L_h = 40\ h^{-1}\ Mpc$, (b) $L_h = 60\ h^{-1}\ Mpc$, and (c) $L_h = 75\ h^{-1}\ Mpc$. Details are as in Fig. 1.
Figure 2 – continued
flattening in \(g(r)\) may also be detected in this case, the strength of the anticorrelation is now so large that a clear indication of the homogeneity scale cannot be obtained, and the correlation function is not much different from that of a pure fractal distribution (such as that shown in Fig. 1). For each survey a test of this kind should be pursued in order to obtain an estimate of the maximum homogeneity scale that may be safely detected. For the volume-limited PP sample with 1092 galaxies, the maximum homogeneity scale that may be approximately detected is of the order of 60 \(h^{-1}\) Mpc.

3.2 Monofractal dust with \(D_2 = 2\)

An additional result provided by the above analysis is that no intermediate scaling regime with \(D_2 \approx 2\) has been spuriously generated by the treatment of boundary effects either in the case of a pure monofractal distribution or for a fractal distribution with transition to homogeneity at large separations. This indicates that the intermediate scaling regime with \(D_2 = 2\) detected by G91 should have an origin that is physical, and not simply related to boundary effects. In the framework of fractal distributions, the regime with \(D_2 = 2\) can be simulated by a \(\beta\)-model with \(m = 2\) and \(p = 1/2\).

Along these lines, we first verify whether a spurious homogeneous regime may be observed for a monofractal distribution with \(D_2 = 2\). Fig. 3 shows the correlation function \(g(r)\) as obtained from an average over 20 realizations of the process described above, where \(D_2 = 2\) at all scales. The error bars are now much smaller than those shown in Fig. 1, due to the fact that pure \(D_2 = 2\) distributions display less dramatic density fluctuations than those with \(D_2 = 1.2\). Again, the average \(g(r)\) does not saturate to a constant value by correction of boundary effects.

For the \(D_2 = 2\) distributions, however, in some cases a spurious flattening of \(g(r)\) may be observed. Fig. 4(a) shows as an example the correlation function \(g(r)\) for one realization where \(g(r) \approx 1\) on large scales. The behaviour of \(g(r)\) indicates an apparent homogenization of the distribution, at least in the slice considered in the analysis. Fig. 4(b) shows the function \(g(r)\) for a different realization of the same random monofractal dust where no flattening is present and \(g(r)\) falls off even more rapidly than \(r^{-1}\) at large scales.

The results reported in Figs 4(a) and (b) require an explanation, and, in particular, verification of whether they may be generated by the correction of boundary effects. In regard to this question, we note that, in the \(\beta\)-model with \(D = 2\), four cubes out of eight survive on the average \((\langle N_0 \rangle = 4)\). This fact, and the possibility of fluctuations with \(N_s > \langle N_s \rangle\), in some cases lead to a situation where each of the \(N_s\) surviving cubes with size \(L_s\), generated by an object with larger size \(2L_s\), is adjacent to other such cubes on one or more sides. If a PP-like sample with maximum size of the order of \(L_s\) is extracted by such a group of adjacent (surviving) objects, then the correlation function may detect an apparent homogeneity on separations of the order of \(L_s/2\). This homogeneity is not spurious, however, and nor is it due to the correction of boundary effects; the flattening of \(g(r)\) in this case is simply the detection of a ‘homogeneity fluctuation’ on the largest scales of the sample in an otherwise random fractal universe. This is none other than the standard statement that single fluctuations may be very important on large scales, where the sampling is not sufficient to provide good statistics. Clearly, by increasing the

![Figure 3](https://academic.oup.com/mnras/article-abstract/266/3/555/1059906/1059906)  

Figure 3. The density correlation function \(g(r) = 1 + \xi(r)\) averaged over 20 realizations of a pure fractal universe with \(D_2 = 2\). Details are as in Fig. 1.
size of the sample well beyond \(L_x\) and/or by averaging different samples, the correlation function for a pure fractal universe with \(D = 2\) should again take its average power-law form on the scale \(L_x\). This effect accounts for the flattening of \(g(r)\) in Fig. 4(a), which is due to a random homogeneity fluctuation of a size comparable to the size of the PP-like sample extracted from this particular realization of the fractal dust. Analogously, the rapid fall-off of \(g(r)\) in Fig. 4(b) is due to a ‘sparseness fluctuation’ in the distribution extracted from a different realization of the \(\beta\)-model.

To show that the above interpretation is correct, in Fig. 4(c) we show the function \(g(r)\) for a geometrical fractal dust with \(N_x = \langle N_x \rangle = 4\), with the constraint that the four surviving cubes are disposed so as not to have any adjacent side. In this case, no flattening in \(g(r)\) is observed in the PP-like sample extracted from the geometrical fractal dust, even though it has been treated with the standard correction of boundary effects. As a conclusion of this analysis, we may thus say that the treatment of boundary effects does not generate any spurious flattening of \(g(r)\) even for a fractal dust with \(D = 2\). Even a random fractal dust, however, may generate homogeneity fluctuations on the scale of a sample, which in this case should not be interpreted as evidence of a global transition to homogeneity. The probability of such local homogeneity fluctuations increases with the value of the fractal dimension \(D\). Clearly, the only way to ensure that a flattening of \(g(r)\) really indicates a global transition to homogeneity and not a local fluctuation is to compare different samples and/or to increase the size of the sample. In the case of the galaxy distribution, the fact that different samples (e.g. CFA145 and PP; cf. fig. 2 of G91) give a similar amplitude and flattening scale of \(g(r)\) suggests that the observed homogeneity is not just a local fluctuation.

### 3.3 A ‘staircase’ \(\beta\)-model

From the results reported above, one may infer that a three-slope shape of \(g(r)\) really does indicate the presence of three different scaling (although not necessarily fractal; see below) regimes in the galaxy distribution. The results reported by G91 thus indicate that the galaxy distribution has to be described by a distribution with \(D_x = 1.2\) at scales smaller than \(r_x \approx 3.5 \, h^{-1} \text{Mpc}\), with \(D_x = 2\) for scales between \(r_x\) and \(r_c \approx 25 \, h^{-1} \text{Mpc}\), and with \(D_x = 3\) beyond \(r_c\). In order to simulate this three-scale distribution, here we consider a monofractal dust generated by the \(\beta\)-model with \(m = 2\) and \(p = 2^{-1.8}\) at scales smaller than \(L_\text{nl}\), by \(m = 2, p = 1/2\) at scales larger than \(L_\text{nl}\) and smaller than \(L_h\), and by \(m = 2, p = 1\) at scales larger than \(L_h\).

Fig. 5(a) shows a PP-like sample of 1092 ‘galaxies’ extracted from a \(\beta\)-model with three scaling regimes, with \(L_h = 43.75 \, h^{-1} \text{Mpc}\) and \(L_\text{nl} = 5.5 \, h^{-1} \text{Mpc}\). Fig. 5(b) (filled circles) shows the average \(g(r)\) calculated from 20 realizations of the above model. The error bars are the standard deviations over the 20 realizations. The simulated samples have a depth of \(79 \, h^{-1} \text{Mpc}\), and are geometrically equivalent to that used by G91 to estimate \(g(r)\) from the PP survey, shown by the open circles. The treatment of boundary effects has been conducted in the same way for both the original and the simulated samples. The agreement between the data and the simulation is impressive, confirming the proposal of the existence of different scaling regimes on different scales advanced by G91.

### 3.4 A simple model of a ‘walls-plus-fractal’ dust

In the above approach, we have used a multi-scale fractal distribution in order to reproduce the observed shape of \(g(r)\). At small scales, the numerical simulations discussed by Valdarnini, Borgani & Provenzale (1992), by Yepes, Dominguez-Tenreiro & Couchman (1992) and by Colombi, Bouchet & Schueffer (1992) show that the natural outcome of non-linear gravitational clustering is a multifractal dust with generalized dimensions \(D_q = 1\) for \(q \approx 2\) at the stage of strongly developed gravitational evolution. Thus the physical origin and the true fractal nature of the scaling regime at small separations are well assessed.

On the other hand, the existence of an intermediate fractal regime with \(D_x = 2\) is more difficult to understand. In fact, there is no evidence, at least at present, that such a regime could be generated by the gravitational evolution. A completely different origin for the scaling regime with \(D_x = 2\) is, however, available. The regime may in fact have a topological origin, being associated with the existence of large-scale flattened structures such as walls or sheets. These structures may be expected to form as a result of the weakly non-linear gravitational evolution of initial conditions charac-
terized by a power spectrum with enough large-scale power. A possible way to interpret the observation of two power-law regimes in the galaxy clustering could thus be in terms of a small-scale fractal dust with $D_2 = 1.2$ (produced by nonlinear hierarchical clustering), which is superimposed on to larger scale, quasi-two-dimensional structures still reminiscent of initial conditions. At still larger scales, the network of flattened structures generates an homogeneous distribution (when averaged over scales that are larger than the typical size of a wall). The transition from the fractal regime to the structured regime is related to the value of the scale of non-linearity (represented by $r_{nl}$), while the transition to homogeneity is presumably related to the correlation length of the initial gravitational potential.

To demonstrate quantitatively that a network of walls on intermediate scales and a fractal dust on small scales are able to reproduce the observed form of the density correlation function, we consider here a very simplified model with a given number density $n_g$ of randomly placed and randomly oriented walls of size $L_x \times L_y \times \sigma_w$; on these, a certain
number of clustering seeds are randomly placed. Each clustering seed originates a fractal dust with correlation dimension $D_2 = 1.2$ for scales below a non-linearity scale $L_{mlb}$, which is taken as the size of the fractal dust piles. Clearly, the simple model discussed here is not intended to provide a dynamical description of the galaxy distribution; this approach has been introduced here simply as a methodological tool to show that a topological explanation for the regime with $D_2 = 2$ is also acceptable (and, in our opinion, is physically preferred) for reproducing the observed shape of $g(r)$. A different topological model that is able to reproduce this behaviour is the Voronoi foam (Icke & van de Weygaert 1987; Lemson & Sanders 1991; van de Weygaert 1991), consisting of a distribution of bubbles, which it is thought can arise in some inflationary scenarios (Amendola & Occhionero 1992). Here we prefer to keep to the simplest planar model in order to show more clearly the relationship of the $D_2 = 2$ range with a topological character of the distribution.

The available observational constraints allow us to determine the parameters of the model. In particular, the knowledge of the average density of galaxies in the local Universe and the determination of the average surface density of galaxies on flattened structures (as given e.g. by de Lapparent, Geller & Huchra 1991) fix a combination of $L_{mlb}$ and of the plane number density $n_p$. The choice is then between a distribution of a few (small $n_p$), but extended (large $L_{mlb}$), flattened structures and many, but smaller, walls. Numerical simulations with different choices of $L_{mlb}$ and $n_p$ indicate that the distribution with a few large walls displays a larger homogeneity scale $r_h$, while that with many smaller flattened sheets has a smaller value of $r_h$. The observed value of $r_h$ is thus helpful for defining the appropriate combination of $L_{mlb}$ and $n_p$. Analogously, the observed value of the transition scale between the two slopes $\gamma = 1.8$ and $\gamma = 1$ of $g(r)$ is then used to fix the value of $L_{mlb}$. For simplicity, in a first approximation one may also suppose that the width $\sigma$ of the walls is mainly determined by the strongly non-linear dynamics and is related to the value of $L_{mlb}$.

As an explicit example of this type of distribution, in Fig. 6(a) we show a PP-like sample obtained from the ‘walls-plus-fractal’ model with the following choice of parameters: $L_{mlb} = 40$ $h^{-1}$ Mpc, $\sigma = 1$ $h^{-1}$ Mpc and $L_{mlb} = 2.5$ $h^{-1}$ Mpc. Clustering seeds are uniformly and randomly distributed on the walls. Each clustering seed is taken as the centre of a distribution generated by a $\beta$-model with $D_2 = 1.2$, whose larger scale is given by $L_{mlb}$. A total of 30 walls are randomly positioned and randomly oriented in a $100$ $h^{-1}$ Mpc$^3$ box. On each wall, an average of 70 clustering seeds is placed; the $\beta$-models centred on the seeds are iterated four times. A slice with the same shape as the volume-limited PP sample is then extracted from the distribution; from this, a subset of 1092 ‘galaxies’ is randomly extracted. Fig. 6(b) shows the average density correlation function $g(r)$ obtained by an average over 20 different realizations of the model; the error bars are the standard deviations on the set of 20 correlation functions.

Fig. 6 shows that even a very simplified model of ‘walls-plus-fractals’ such as that discussed here is able to reproduce the observed shape of the two-point correlation function. In particular, an intermediate regime with $D_2 = 2$ is clearly generated. If only walls were present, this would be the only scaling law in the correlation function. By contrast, the small-scale scaling regime with $D_2 = 1.2$ is due to the presence of the fractal dust at strongly non-linear scales. As a word of caution, we note that the ‘walls-plus-fractal’ model has one more parameter than the multi-scale $\beta$-model and requires some tuning of parameters that are not directly observable. On the other hand, the ‘walls-plus-fractal’ model probably has deeper physical motivations.

4 SUMMARY AND CONCLUSIONS

The main results of the present work may be summarized as follows.

(1) The observed flattening of $g(r)$ at scales larger than about $r_h = 30$ $h^{-1}$ Mpc is not due to the standard treatment of boundary effects. In fact, pure fractal dusts with no homogeneity transition do not display a flattening in the correlation function $g(r)$; i.e. a pure fractal universe is excluded.

(2) The scale of the flattening of $g(r)$ provides a reliable estimate of the scale of transition to homogeneity in sufficiently deep samples.

(3) The intermediate scaling regime with $D_2 = 2$ observed by G91 is not spuriously generated by the treatment of boundary effects.

(4) A simple fractal model with three scaling regimes (namely, a fractal dust with $D_2 = 1.2$ at small scales, a fractal dust with $D_2 = 2$ at intermediate scales and an homogeneous distribution at large scales) is able to reproduce the observed correlation function. This indicates that $\gamma = 1$ at intermediate scales really does indicate the presence of a scaling regime with $D_2 = 2$.

(5) As an alternative to the explanation of the $D_2 = 2$ scaling range in terms of a fractal-like distribution of matter, we have introduced a different model where a small-scale fractal dust with $D_2 = 1.2$, which is known to be generated by the strongly non-linear gravitational clustering, is superimposed on to a larger scale network of ‘walls’. These walls should be dynamically generated by the weakly non-linear evolution of initial long-wavelength perturbations. The intermediate scaling regime is in this case not fractal but rather topological, i.e. it is related to the existence of planar structures and not to a fractal dust. This is the same conceptual difference that exists between a plane and the trajectory of a particle performing a Brownian motion.

In general, the results discussed in this paper show that both a multi-scale fractal dust and a distribution generated by a ‘walls-plus-fractal’ model are able to reproduce the observed correlation function. In both cases, the slope of $g(r)$ indicates a form $P(k) \propto k^{-2}$ for the power spectrum on intermediate scales. A numerical study of structure formation from initial conditions resembling a fractal dust with $D_2 = 2$, a situation that could be speculated to arise in an explosion scenario (Vicsek & Szalay 1987), is discussed in Branchini et al. (in preparation). Within a more canonical framework, a phenomenological model capable of reproducing the observed behaviour of $g(r)$ and taking into account the large-scale microwave background constraints from COBE is discussed by Branchini et al. (1993). This model correctly reproduces the observed power spectra from different surveys (in particular the CfA15, Vogeley et al. 1982), and is similar to a CDM model with ‘tilted’ primordial index $n = 0.7$ and bias factor $1/\sigma_8 \simeq 2$. 

© Royal Astronomical Society • Provided by the NASA Astrophysics Data System
In a more general framework, we note that the two geometric possibilities discussed here for the intermediate scales (fractal or topological) cannot be distinguished on the sole basis of a correlation analysis. In this regard, it is necessary to identify appropriate statistical tests capable of distinguishing between the two different types of distribution in the analysis of data. Presumably, a percolation approach, a careful evaluation of the filling factor or an evaluation of the lacunary properties should be of value for distinguishing between a fractal dust and a 'walls-plus-fractal' distribution.

The model of the fractal dust superimposed on to randomly placed walls introduced here is extremely simplified; it has been used only to show that even such a simple model is capable of reproducing the observed form of the correlation function. Physically, the most plausible origin of the large-scale flattened structures is expected to be provided...
by the weakly non-linear gravitational evolution of an initial density perturbation field with sufficient large-scale power, as modelled, for example, by the Zeldovich approximation or by the adhesion approach (Gurbatov, Saichev & Shandarin 1989).

It is interesting to recall that Dekel & Aarseth (1984), reploting $1 + \xi(r)$ for the Davis & Peebles (1983) correlation analysis of the CFA, noticed the appearance of a two-power-law behaviour which they interpreted as evidence for flat, pancake-like structures on large scales. These authors concluded that these structures were dynamically young; i.e. they could not be considered as being the result of a top-down scenario such as that produced by adiabatic fluctuations in a purely baryonic Universe, or, in more modern terms, by adiabatic fluctuations within a hot dark matter (neutrino) dominated Universe. A pure top-down scenario is indeed not necessarily required in order to produce large-scale sheet-like structures; the only condition required is that the initial power spectrum on the corresponding scales keeps enough relative power. As a matter of fact, large-scale structures such as filaments and sheets may be produced even in a nearly purely hierarchical model such as CDM (e.g. Davis & Efstathiou 1988), although the flat version of this model does not provide enough large-scale power to match the observations (Efstathiou et al. 1990; Maddox et al. 1990). An interesting alternative possibility is to consider the case of a Universe dominated by a mixture of hot plus cold dark matter (see e.g. Valdarnini & Bonometto 1985; Klypin et al. 1993). In this picture, the resulting initial spectrum might have the appropriate form for producing both strong non-linear evolution at small scales (hierarchism and $D_2 = 1.2$) and dynamically young flattened structures at intermediate scales (where $D_2 \approx 2$).

As a final remark, we note that, in both the models we have discussed here, the Fourier phases of the density field are quite far from being randomly and uniformly distributed. This fact is related to the existence of ‘structures’ at scales much larger than the correlation length $r_0$; in fact, the correlation function cannot distinguish between distributions with random phases or highly structured phases since it is related only to the power spectrum of the density field (i.e. to the second moment of the distribution). We recall that phase correlations are, for example, generated even by weakly non-linear evolution before pancake formation (e.g. Ryden & Gramann 1991; Scherrer, Melott & Shandarin 1991; Soda & Suto 1992). A careful study of the properties and evolution of phase correlations, by direct investigation, by the use of the bispectrum (e.g. Baumgarten & Fry 1991) or by $n$-point correlation functions (Bonometto et al. 1993) should thus be of value for better understanding the properties of the large-scale structure of the Universe.

ACKNOWLEDGMENTS

We are grateful to S. Bonometto, S. Borgani, E. Branchini and R. Valdarnini for helpful discussions and to B. Jones for several useful comments and for constructive refereeing of the paper. AP and GM are members of the ‘European Cosmology Network’ in the framework of the EEC contract ‘Large Scale Structure of the Universe: Evolution and Statistics’.

REFERENCES

Amendola L., Occhionero F., 1992, preprint
Castagnoli C., Provenzale A., 1990, Vistas Astron., 33, 323
Eckmann J.-P., Ruelle D., 1985, Rev. Mod. Phys., 57, 617
Grassberger P., Procaccia I., 1983, Physica, 9D, 189
Pietronero L., 1987, Physica, 144A, 257

© Royal Astronomical Society • Provided by the NASA Astrophysics Data System