Mon. Not. R. astr. Soc. (1992) **256**, 477–499

The topology of the QDOT *IRAS* redshift survey

Ben Moore,[1,2] Carlos S. Frenk,[1]★ David H. Weinberg,[2] Will Saunders,[3,4] Andy Lawrence,[5] Richard S. Ellis,[1] Nick Kaiser,[6] George Efstathiou[4] and Michael Rowan-Robinson.[3]

[1] *Department of Physics, University of Durham, South Road, Durham*
[2] *Astronomy Department, University of California, Berkeley, CA 94720, USA*
[3] *Astronomy Unit, Queen Mary and Westfield College, Mile End Road, London E1 4NS*
[4] *Department of Physics, University of Oxford, Keble Road, Oxford*
[5] *Department of Physics, Queen Mary and Westfield College, Mile End Road, London E1 4NS*
[6] *CIAR Cosmology programme and CITA, St George Street, Toronto, Canada*

Accepted 1991 December 20. Received 1991 December 9

SUMMARY

We have used an all-sky redshift survey of galaxies detected by the *IRAS* satellite to investigate the topology of the Universe to a depth of 200 h^{-1} Mpc. Qualitatively, the distribution of galaxies out to this distance resembles a Gaussian density field with a sponge-like topology: high- and low-density regions are topologically similar and surfaces of constant density are interconnected. Quantitatively, we have used the genus-threshold density relation of Gott *et al.* to test the hypothesis that the galaxy distribution grew out of initially Gaussian density fluctuations and to measure the effective slope of the power spectrum of fluctuations over a range of length-scales between 10 and 50 h^{-1} Mpc. To estimate random and systematic uncertainties in our analysis, we have used artificial 'galaxy' catalogues constructed from N-body simulations and a variety of Monte-Carlo techniques. We find that the observed genus curves are consistent with the Gaussian hypothesis. As an example of a non-Gaussian field, we have tested a Voronoi foam model with ~ 100 h^{-1} Mpc cells. We find that such a model can be ruled out at $\sim 5\sigma$. Our topological analysis is consistent with a power spectrum of galaxy fluctuations of the form $P(k) \propto k^n$, with $n \simeq -1$, over the full range of scales considered. On scales $\lesssim 15$ h^{-1} Mpc, the QDOT power spectrum has a similar slope to that of the mass distribution in the standard cold dark matter model, but it falls off less steeply on larger scales; the maximum discrepancy occurs at ~ 30 h^{-1} Mpc and is significant at about 2σ. Our power spectrum results are consistent with previous counts-in-cells analyses of the same survey, but the present method (which is sensitive to the slope of the spectrum rather than to its amplitude) weights the data quite differently.

Key words: surveys – galaxies: clustering – galaxies: distances and redshifts – galaxies: formation – large-scale structure of Universe – infrared: galaxies.

1 INTRODUCTION

Quantifying the large-scale structure of the Universe is a difficult task and a variety of techniques have been developed to approach it. The first large data sets were the two-dimensional catalogues of Shapley & Ames (1932) and Zwicky *et al.* (1961–68), and the Shane & Wirtanen (1967) galaxy counts. Large structures and clustering were readily apparent and statistics such as the angular correlation function were developed to quantify deviations from a random distribution. The advent of complete redshift surveys (Davis *et al.* 1982) revealed a complex network of voids and filamentary structures in the galaxy distribution. As new redshift surveys were completed, structures approaching the size of the survey volume were suggested (Geller & Huchra 1989). Only deeper surveys with well-defined selection criteria can clarify this issue. The clustering of galaxies in three dimensions has proved to be one of the most important constraints

★ Nuffield Foundation Science Research Fellow.

on models of galaxy formation. A successful model must reproduce the observed clustering pattern of galaxies as measured by statistics such as the two-point correlation function (Peebles 1980), match the observed large-scale velocity flows (Burstein, Faber & Dressler 1990), contain voids as large as Bootes (Kirshner *et al.* 1981) and have features as extended as the 'Great Wall' (Geller & Huchra 1989). One important characteristic of the galaxy distribution is its topology: does it consist of isolated high-density clumps in a low-density background (an isolated cluster or 'meatball' topology), isolated voids surrounded by walls of galaxies (a 'swiss cheese' or 'bubble' topology), or mutually interlocking, multiply connected high- and low-density regions (a 'sponge' topology)?

In the standard inflationary models for galaxy formation, present-day structures grew from small-amplitude, random-phase (or, equivalently, Gaussian) quantum noise in the early Universe (Bardeen, Steinhardt & Turner 1983). The statistical properties of a random phase distribution are completely specified by its power spectrum. In the linear regime, overdensities grow as $\delta\rho/\rho \propto a$, where a is the expansion factor. Therefore, in comoving coordinates, positive and negative fluctuations grow in place, increasing only in amplitude, and contours of the density field selected by volume fraction do not change with time. When $\delta\rho/\rho \approx 1$, the high-density regions begin to collapse rapidly and the initial power spectrum, which was frozen into the density field, becomes distorted. Biased galaxy formation (the preferential formation of galaxies in regions of high background density) can also distort the relation between initial mass density and final galaxy density. However, moderate dynamical evolution and biased galaxy formation both tend to preserve a monotonic relation between initial mass density and final galaxy density, i.e. they map high-density regions of the initial conditions into high-density regions of the final galaxy distribution, low-density regions into low-density regions and so on. If we smooth the data over a scale larger than the characteristic clustering length and define isodensity contours in terms of the fractional volume they enclose, we can recover properties of the initial density fluctuations, as verified by Weinberg, Gott & Melott (1987; hereafter WGM).

Until recently, studies of the topology of the galaxy distribution were limited to qualitative remarks about the high- and low-density regions (Gott, Melott & Dickinson 1986; hereafter GMD), although it was apparent from N-body simulations that the visual impression of the density field depends strongly upon the initial mass fluctuations (Efstathiou *et al.* 1988). Hamilton, Gott & Weinberg (1986) rederived an expression showing that the relation between the integrated curvature (or, equivalently, the genus) of density contours and the contour threshold density has a universal form for random-phase density fields, with a normalization determined by the slope of the power spectrum (see also Doroshkevich 1970; Adler 1981; Bardeen *et al.* 1986). In subsequent studies, Gott, Weinberg & Melott (1987; see also WGM and Melott, Weinberg & Gott 1988) applied this relation to numerical simulations and other models and showed that it could be used to test the hypothesis of random-phase initial fluctuations. Gott *et al.* (1989) used various redshift surveys to measure the topology on a range of length-scales between 3 and 50 h^{-1} Mpc. (We denote the Hubble constant by $H_0 = 100\ h$ km s^{-1} Mpc^{-1}.) On smoothing scales less than 10 h^{-1} Mpc, they used the CfA survey (Huchra *et al.* 1983), Tully's (1987) all-sky catalogue of nearby galaxies, the Schneider *et al.* (1990) diameter-limited survey of dwarf galaxies and the Giovanelli & Haynes (1985; hereafter GH) catalogue. A deeper subset of the GH survey and the Abell (1958) cluster catalogue were used to measure the topology on scales of 24 and 50 h^{-1} Mpc respectively. All of these studies taken together provide a strong test for any theory of galaxy formation. Gott *et al.* (1989) compared the data with heavy neutrino models, the cold dark matter model (CDM) and 'bubble' models and concluded that CDM provided the best fit to the data over the scales considered.

Although the GH and Abell cluster samples gave the most information on the large scales of interest to us here, they must be regarded with caution for the following reasons. The GH sample could be biased towards a 'meatball' topology because it deliberately targets the massive filamentary Perseus–Pisces supercluster which dominates the survey, so it might not constitute a fair sample of the local Universe. (However, while the GH sample shows a 'meatball' topology at 6 h^{-1} Mpc, it is consistent with a Gaussian topology at 12 and 24 h^{-1} Mpc.) The Abell catalogue should also be used with caution because it has been selected by eye and may be affected by contamination and projection effects (Lucey 1983; Frenk *et al.* 1990). It should also be noted that only a few per cent of galaxies actually lie in Abell clusters and it is not clear that the topology of the cluster distribution should match that of the galaxy distribution.

The Queen Mary and Westfield College, Durham, Oxford and Toronto (QDOT) 'one-in-six' redshift survey of galaxies detected by the *IRAS* satellite (Lawrence *et al.*, in preparation) samples the density field in the Universe to a redshift $z \approx 0.07$, well beyond the Local Supercluster. There are 2163 galaxies in the survey, randomly sampled at a rate of one-in-six from the *IRAS* point source catalogue of sources with 60-μm flux greater than 0.6 Jy; the survey covers almost the whole sky at galactic latitudes $|b| > 10°$. [See Lawrence *et al.* (in preparation) for a detailed description of the survey.] Regions of sky contaminated by interstellar dust (the infrared 'cirrus') and a patch of sky not covered by the *IRAS* satellite are coded into a mask which excludes about 17 per cent of the whole sky. The random sampling strategy allows the maximum information on large scales to be obtained from a

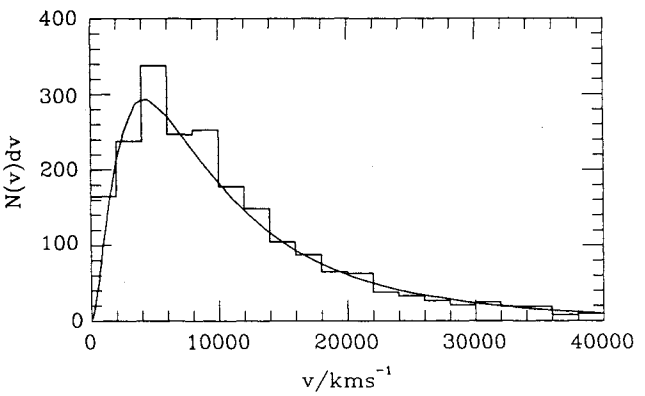


Figure 1. The redshift distribution of galaxies in the QDOT survey. The histogram represents the actual data and the solid curve is the result of integrating the QDOT luminosity function (solution 19 in table 3a of Saunders *et al.* 1990).

given amount of telescope time (Kaiser 1986). The QDOT redshift survey is well suited to investigate structure on large scales because it provides near full sky coverage with galaxies selected in a uniform way. The distribution of recession velocities, $n(v)$, shown in Fig. 1 has a median of 9000 km s^{-1} and a long tail which samples a fairly large range of scales. A separate survey of all *IRAS* galaxies brighter than 1.94 Jy has been carried out by Strauss *et al.* (1990).

Several results have already emerged from analyses of the QDOT data. The survey has been used to constrain the cosmological density parameter, Ω, by mapping the local peculiar gravity field and comparing the velocity predicted for the Local Group (Rowan-Robinson *et al.* 1990) and for a sample of 1000 galaxies (Kaiser *et al.* 1991) with observations. Efstathiou *et al.* (1990) and Saunders *et al.* (1991) have estimated the variance of QDOT galaxy counts-in-cells and find fluctuations on large scales in excess of those predicted in the standard cold dark matter (SCDM) model. Saunders *et al.* also obtained a marginal detection of skewness in the count distribution on large scales. The great depth of the QDOT survey allows a comparison with the topological analyses of the GH and Abell cluster surveys on scales greater than 10 h^{-1} Mpc. The QDOT survey provides higher 'signal-to-noise' (smoothing elements per survey volume) for this type of study than any previous galaxy survey on scales greater than $\sim 12\ h^{-1}$ Mpc.

We shall compare our results with theoretical predictions, particularly with those of the standard cold dark matter theory. This is the best-studied and, in many ways, the most successful cosmogonic model to date. Certainly on scales less than $\sim 10\ h^{-1}$ Mpc, the theoretical predictions are roughly consistent with many observed properties of galaxies and their spatial distribution (see Frenk 1991 for a recent review). On larger scales the model predicts significantly less superclustering than has been measured in recent surveys, including the QDOT survey. It is perhaps worth summarizing here the main tenets of the SCDM model. The SCDM model is based on four premises: (i) the dark matter consists of weakly interacting, 'cold' elementary particles; (ii) the Universe has critical density; (iii) the primordial fluctuations, which seed the formation of structure, are of the type predicted by the inflationary theory of the early Universe, i.e. Gaussian, adiabatic and scale-invariant. A less fundamental element of the SCDM model is the hypothesis (iv) that the distribution of galaxies is biased relative to the distribution of mass according to the 'high-peak model'; galaxies are assumed to form only near high peaks of the suitably smoothed linear density field so that, on large scales, the rms fluctuations in the galaxy distribution are a constant multiple of the corresponding mass fluctuations. (This model is sometimes referred to as the 'linear biasing model'.)

The QDOT survey has already proved extremely valuable in testing directly some of the tenets of the SCDM model. The dynamical analyses of Rowan-Robinson *et al.* (1990) and Kaiser *et al.* (1991) provide support for assumption (ii) above. Since standard Big Bang nucleosynthesis limits the density of baryons to less than about 10 per cent of the critical value (Olive *et al.* 1990), this result also provides indirect support for assumption (i). The counts-in-cells analyses of Efstathiou *et al.* (1990) and Saunders *et al.* (1991) test a combination of model assumptions, particularly assumption (iii) concerning the shape of the primordial density fluctuation spectrum and assumption (iv) concerning the way in which these fluctuations are manifested in the galaxy distribution. As mentioned above, both of these studies indicate large-scale power in excess of that predicted by the SCDM model. The present analysis is sensitive to the shape of the fluctuation spectrum and to the assumption of random phases, but it is less sensitive than the counts-in-cells analyses to the details of the assumed biasing prescription.

In this paper we use a combination of analytic and numerical tools to compare the topology of the QDOT survey with predictions of the SCDM model. For the latter, we rely heavily on artificial 'QDOT' surveys constructed from N-body simulations. Details of the simulations are given by Frenk *et al.* (1990), while the general procedure for constructing flux-limited 'galaxy' catalogues from simulations is described by White *et al.* (1987). For the present analysis, we matched this procedure to the particular circumstances of the QDOT survey. We assumed the luminosity function given by solution (19) of Saunders *et al.* (1990) (which includes a model for density evolution and k-corrections). The artificial 'QDOT catalogues' have the same luminosity function, redshift distribution and sky coverage as the real data and were masked in exactly the same way. They assume a biasing parameter $b=2$, in the notation of Frenk *et al.* (1990).

In the next section we describe qualitatively the density field traced by *IRAS* galaxies and we make visual comparisons with artificial surveys constructed from the SCDM model and from power-law fluctuation spectra. In Section 3 we review the basic theory necessary for a topological analysis of galaxy redshift surveys, namely the genus–density relation and its connection to other estimators of the power spectrum. Section 4 describes the analysis of the QDOT data and an estimate of the errors in the genus curve using our artificial 'QDOT' surveys. We also compare the genus curve of the data with Gaussian and non-Gaussian models. In Section 5 we estimate the slope of the power spectrum over a range of length-scales. The errors in the slope are calculated using the results of Section 4 and Monte-Carlo simulations. In Section 6 we summarize our results and their relation to the counts-in-cells analyses of the same data.

2 QUALITATIVE TOPOLOGY

Whilst the most luminous *IRAS* galaxies are predominantly interacting, starbursting disc systems (Joseph & Wright 1985; Sanders *et al.* 1988; Lawrence *et al.*, in preparation), most of the galaxies in the QDOT survey appear to be 'normal' late-type spirals, similar to the Milky Way, which avoid the cores of rich clusters but appear to trace the same large-scale structure as optically selected galaxies (Babul & Postman 1989).

Visual impressions of the three-dimensional galaxy distribution of the QDOT survey have been given by Saunders (1991) and Kaiser *et al.* (1991). These pictures portray the wealth of structure that can be identified in the catalogue. Topological analysis uses contours of constant density and these provide an interesting visual representation of the structures that make up the high and low regions of the density field, in a way which complements the previous representations.

We obtain a smoothed density field by weighting the observed galaxy distribution with the inverse of the selection

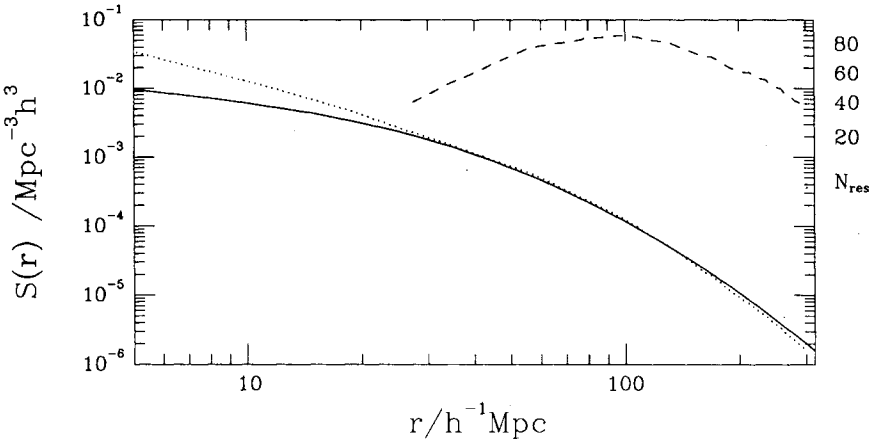


Figure 2. The selection function of the QDOT survey. The solid line is obtained from equation (2), integrating the luminosity function of Saunders *et al.* (1990); the dotted line is the result of applying equation (3). The dashed line gives the number of independent resolution elements (equation 11) out to distance r.

function and then smoothing with the Gaussian function

$$W(r,\lambda)=\frac{1}{\pi^{3/2}\lambda^3}\,e^{-r^2/\lambda^2}, \tag{1}$$

where r is distance and λ is the smoothing length. W is normalized so that its integral over all space is unity. [To maintain consistency with the conventions of Gott *et al.* (1989), our definition of smoothing length differs from that of Saunders *et al.* (1991) by a factor of $1/\sqrt{2}$.] The selection function, $S(r)$, is defined as the expected number of survey galaxies at a distance r in an unclustered distribution,

$$S(r)\equiv\int_{L_{\rm lim}}^{\infty}\phi(L)\,dL, \tag{2}$$

where $L_{\rm lim}$ is the minimum observable luminosity at distance r and $\phi(L)$ is the galaxy luminosity function. The solid line in Fig. 2 shows the selection function we adopt, obtained from equation (2) using the luminosity function of Saunders *et al.* (1990; solution 19). For comparison, the dotted line in Fig. 2 shows the selection function calculated directly from the data via

$$S'(r)\equiv\frac{3}{\omega}\sum_{D_{{\rm max},i}>r}\frac{1}{D^3_{{\rm max},i}}, \tag{3}$$

where $D_{{\rm max},i}$ is the maximum distance out to which the ith galaxy can be detected and ω is the solid angle of the survey. Equation (3) overestimates the local density below 30 h^{-1} Mpc but agrees fairly well with the estimate from equation (2) beyond that distance.

We now make plots of the QDOT data by carving out a sphere from the survey, tabulating the smoothed density field on a lattice and drawing isodensity contours which enclose the regions above or below a given threshold. Our galaxy is located at the centre of the sphere with the galactic plane running horizontally across the plot. The Galactic Centre lies towards the right-hand side in the positive x-axis direction. In Fig. 3(a) we plot the density field out to a distance $r_{\rm max}=75\ h^{-1}$ Mpc using a smoothing length $\lambda=12\ h^{-1}$ Mpc. The surfaces in this figure correspond to high-density regions contoured at the level which encompasses one third of the total volume of the sphere. The dominant feature is the Hydra–Centaurus complex, consisting of the Hydra–Centaurus and Pavo–Indus superclusters, which joins smoothly to Coma and A1367 in the northern hemisphere. From this vantage point the large overdensity of the Local Supercluster (LS) is hidden partially by the clusters forming the Hydra–Centaurus complex. The surface in Fig. 3(b) encloses the low-density regions which encompass one third of the total volume. This map is dominated by the Local Void which joins to the Eridanus Void in the southern hemisphere. Tully Void 3 (Tully 1987) is also clearly seen.

Figs 3(c) and (d) show similar plots but now out to a larger scale, $r_{\rm max}=100\ h^{-1}$ Mpc, with $\lambda=15\ h^{-1}$ Mpc. As before, the surfaces encompass roughly one third of the total volume. In Fig. 3(c) the Local Supercluster now lies insignificantly at the centre of the picture with A1367, Coma, A2197 and Hercules joining together to form part of the Great Wall in the northern hemisphere. Joined to this huge structure with a bridge of galaxies passing through the galactic plane are Perseus–Pisces and N1600 which make up the frightening Pisces–Cetus complex. The Hydra–Centaurus and Pavo–Indus clusters are still visible in this plot. The superclusters S2 and S6 were identified by Saunders *et al.* (1991). Fig. 3(d) shows the corresponding low-density regions on the same scale and with the same smoothing. The Tully Voids are all present along with the Eridanus and Local Voids which join together smoothly to form a continuous surface. The structures in these pictures have an overdensity of about $\pm1\sigma$ from the mean density.

In Fig. 4(a) the threshold is now such that the surface encompasses about a tenth of the total volume of the sphere. These high-density regions correspond to roughly 1.7σ peaks of the density field. The interconnected structures visible in Fig. 3(c) have mostly broken up and only isolated overdense superclusters, such as Hercules and Pavo–Indus, remain. Fig. 4(b) shows the high-density regions out to $r_{\rm max}=150\ h^{-1}$ Mpc smoothed with $\lambda=24\ h^{-1}$ Mpc. Hercules, the dominant feature in this plot and in the whole QDOT survey, extends from 100 h^{-1} Mpc to beyond the edge of this diagram. The superclusters Aquarius–Capricorn, Near–Horologium and S2 dominate the southern

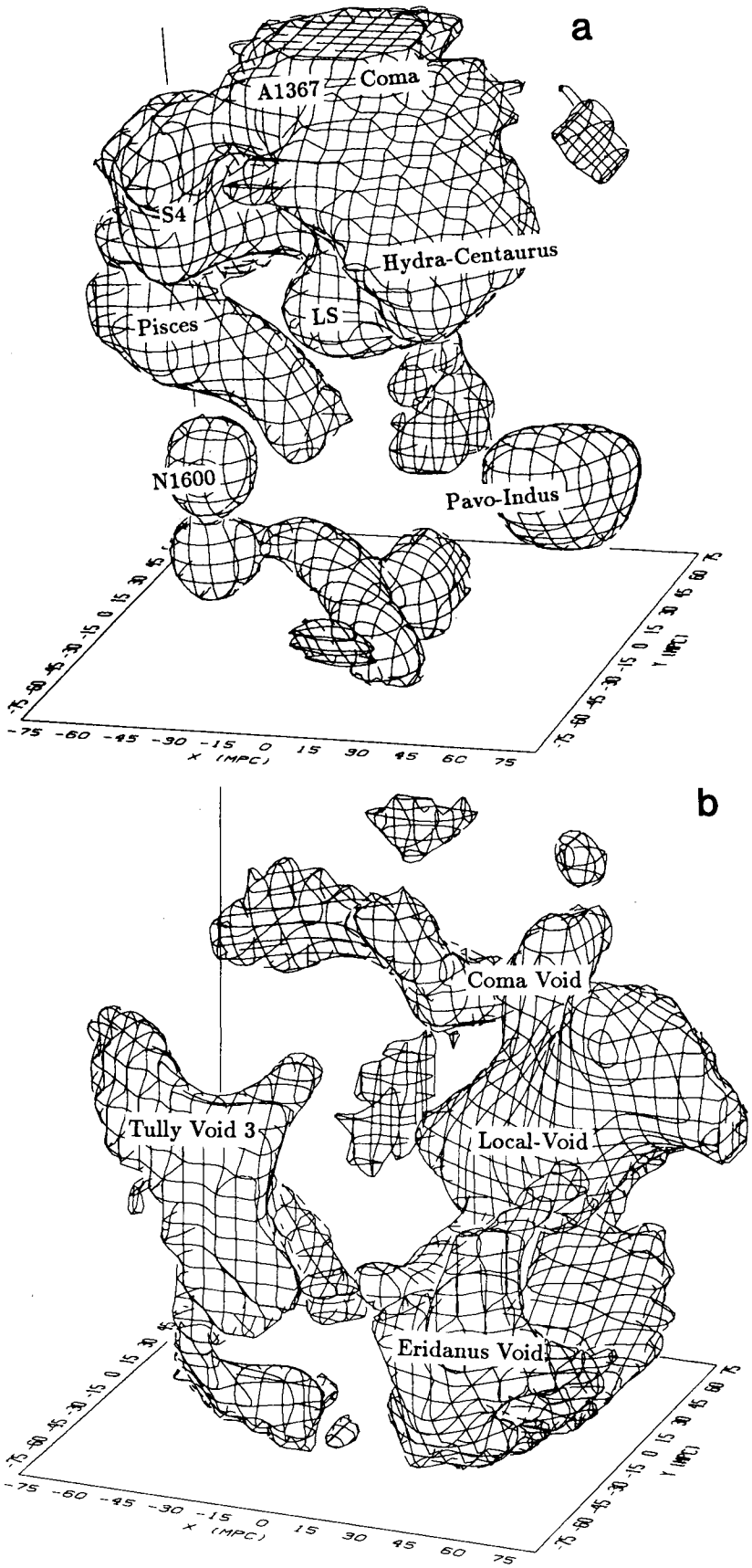
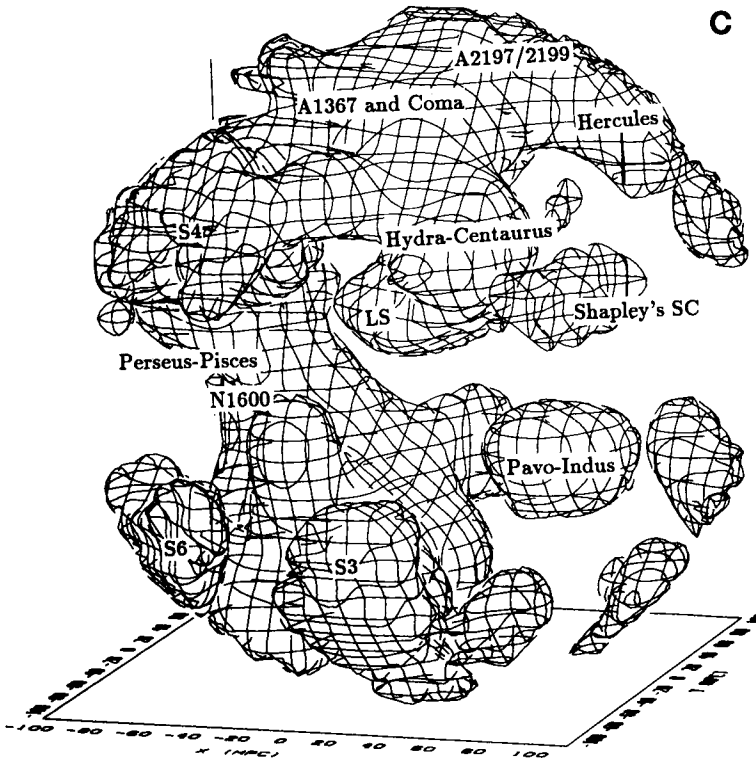


Figure 3. Isodensity contours in the QDOT survey enclosing roughly one third of the total volume. The Galactic Centre points towards the right-hand side of the plot, along the positive x-axis, with galactic longitude running anticlockwise around the sphere. (a) High-density regions within a sphere of radius 75 h^{-1} Mpc smoothed with a Gaussian of width $\lambda = 12\ h^{-1}$ Mpc. (b) Low-density regions on the same scale and with the same smoothing as (a). (c) High-density regions to a depth of 100 h^{-1} Mpc smoothed on scale $\lambda = 15\ h^{-1}$ Mpc. (d) Low-density regions on the same scale and with the same smoothing as (c).



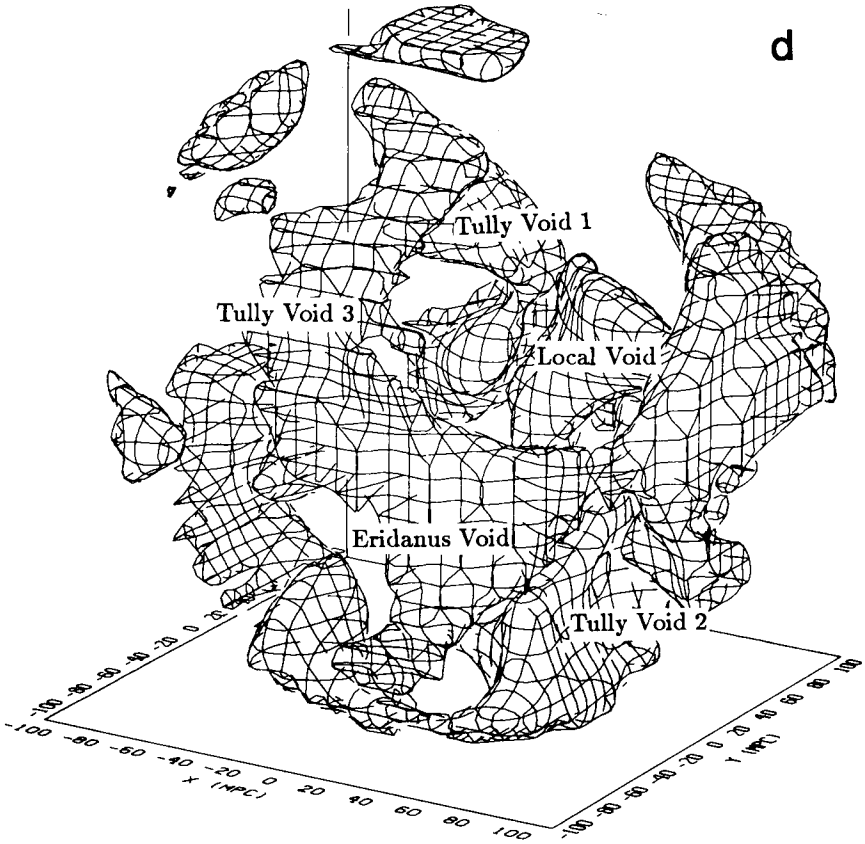


Figure 3 – *continued*

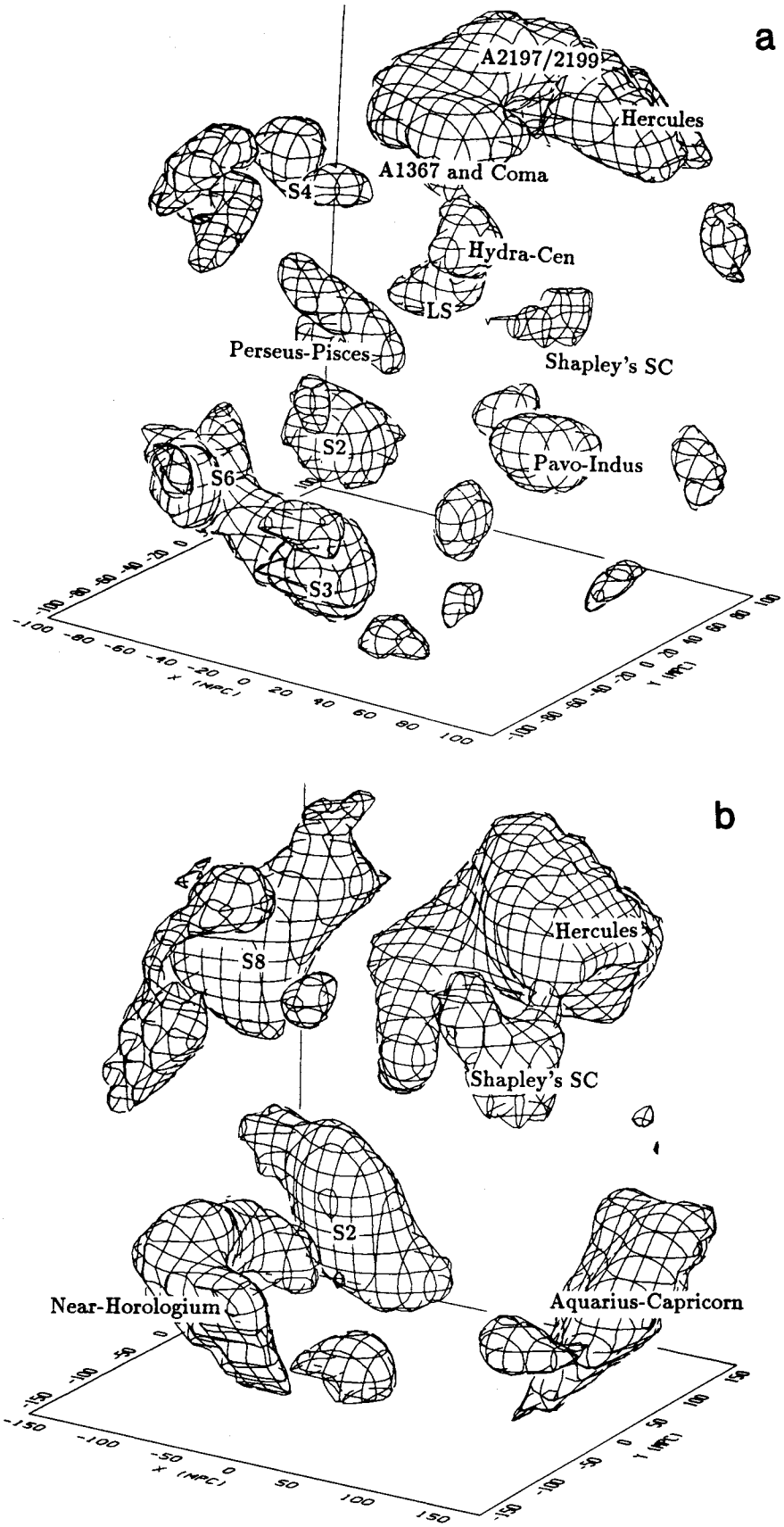
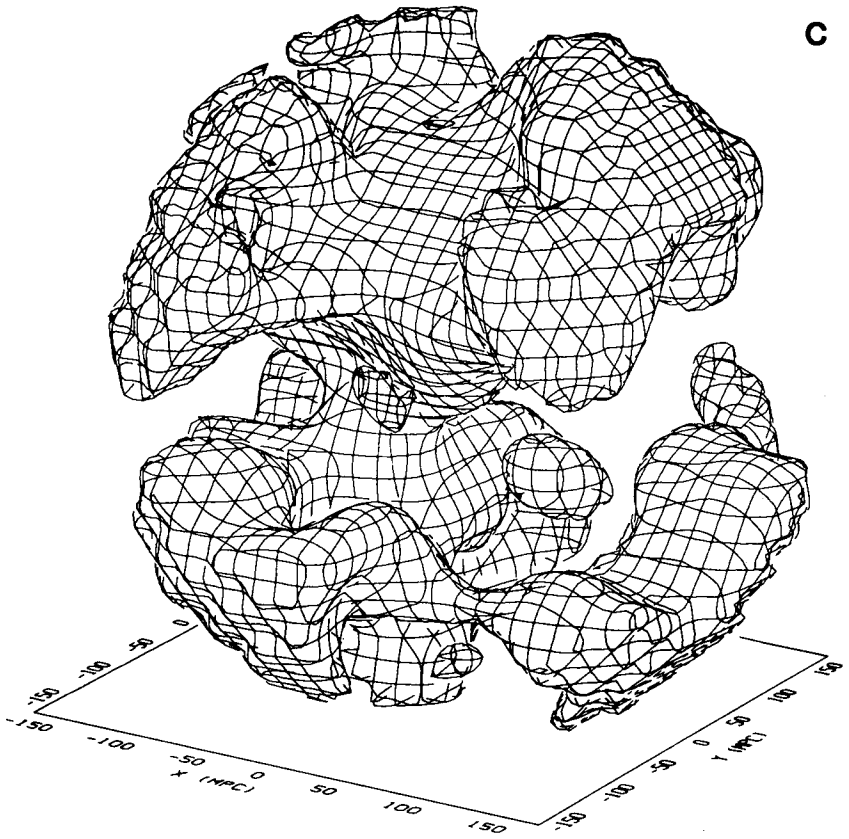


Figure 4. High-density regions of the QDOT survey. The coordinates are as in Fig. 3. (a) The high-density field of Fig. 3(c) but at a higher threshold so that roughly only one tenth of the total volume is enclosed. (b) High-density regions enclosing the same volume as in (a), but in a sphere of radius 150 h^{-1} Mpc smoothed with $\lambda = 24\ h^{-1}$ Mpc. (c) and (d) The density field of (a) but at the median density contour so that each plot shows one half of the total volume. The structures are interlocking and sponge-like, as expected in a Gaussian random field.



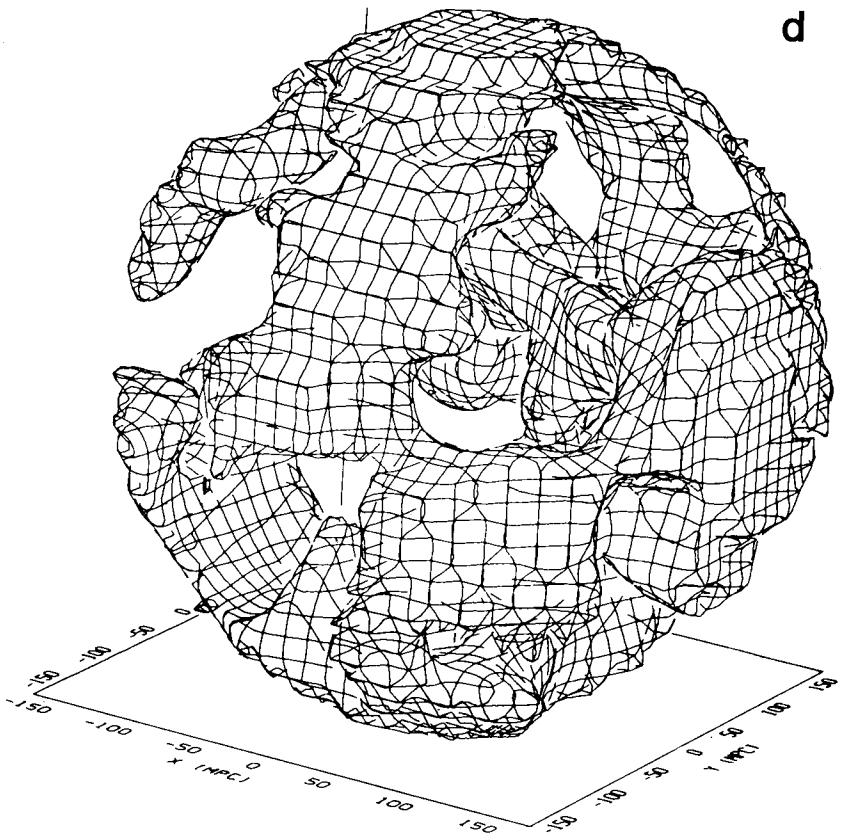


Figure 4 – *continued*

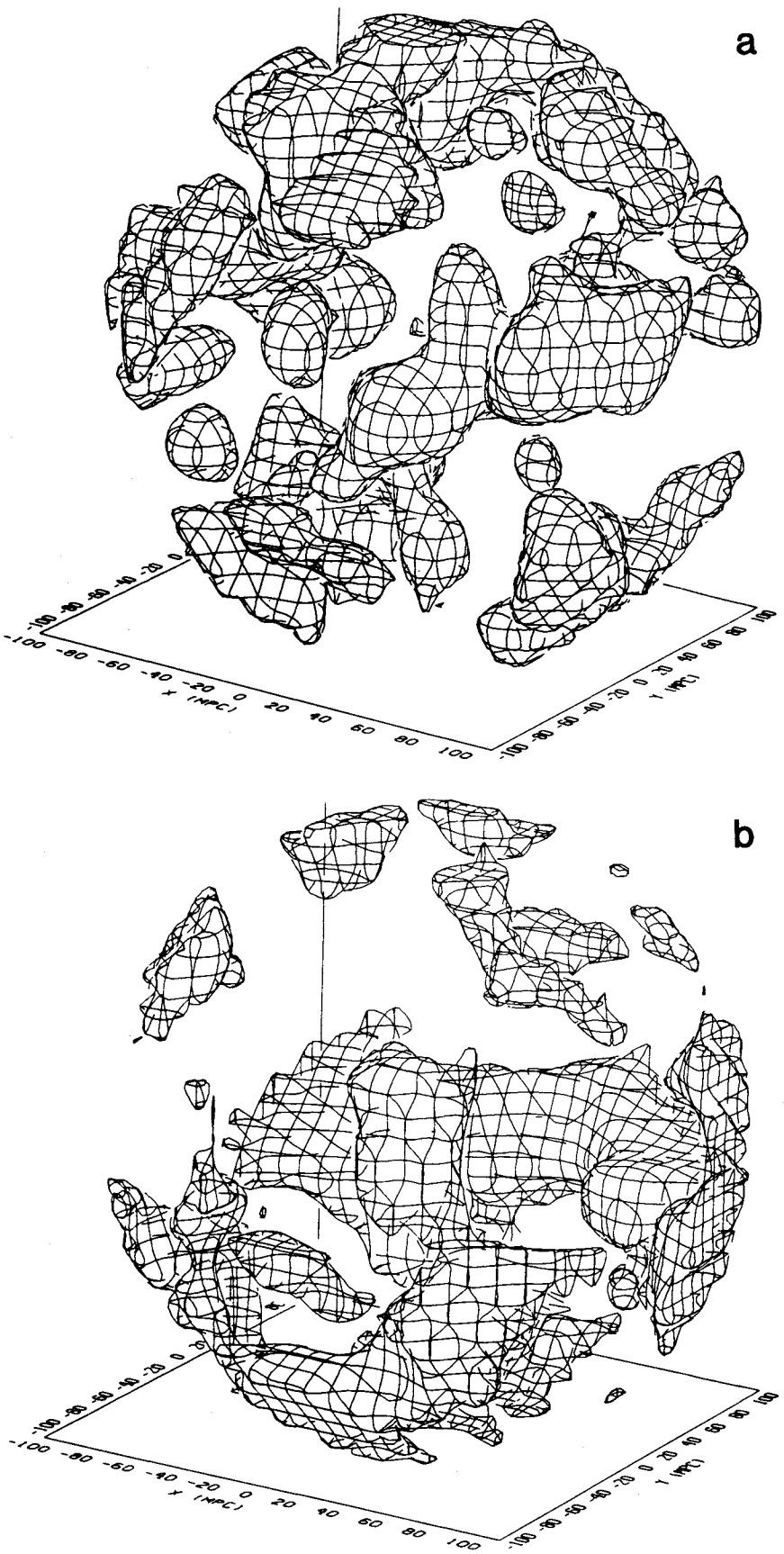


Figure 5. (a) High- and (b) low-density regions of an artificial 'QDOT' survey constructed from an SCDM N-body simulation with the same selection function, flux limit and volume as the QDOT survey. The smoothing lengths and volumes are the same as those in Figs 3(c) and (d), with which these plots should be compared.

hemisphere. The structures shown in Fig. 4(b) have overdensities similar to those in Fig. 3, approximately 1σ from the mean. Over 200 galaxies lie within these superclusters. If we raise the threshold so that only 2σ peaks remain visible, then we only see the Hercules supercluster, represented by 27 galaxies. Finally, on the same scale and with the same smoothing length as Fig. 4(b), Figs 4(c) and (d) show the interlocking median density contour of the high- and low-density regions. The similarity between the high- and low-density regions and their apparent connectedness are typical of the sponge-like topology expected for the median density contour of a Gaussian random field (GMD).

For comparison with the QDOT data depicted in Fig. 3(c), Fig. 5(a) shows a plot of the density field in an artificial catalogue constructed from an SCDM N-body simulation (see Section 1). Fig. 5(b) shows the corresponding low-density regions. The galaxy distribution in the artificial catalogues is treated in an identical way to the real QDOT survey data. Visually, the topology of the high- and low-density regions of the model appears quite similar to that of the QDOT survey. On this smoothing scale the SCDM power spectrum can be approximated by a power law, $P(k) \propto k^n$, with $n = 0.25$.

It is illuminating to compare plots of the QDOT survey and SCDM model with plots of distributions with different power spectra. Fig. 6(a) shows high-density regions in a smoothed Gaussian density field laid down using a power-law fluctuation spectrum with $n = 1$. The absence of large-scale overdensities is characteristic of such a power spectrum which produces lots of small-scale power or 'choppiness'. Similarly, Fig. 6(b) shows a smoothed density field with $n = -2$. This spectrum has large coherent fluctuations that extend beyond the 100 h^{-1} Mpc sphere depicted here. The power-law spectra produce scale-free density fields which we normalize by keeping $\lambda/r_{\max}$ fixed for the models and the data.

The plots in Figs 5 and 6 contrast the structures we would expect to observe in Gaussian models with different initial power spectra. Characterizing the differences between these density fields is the subject of the next two sections.

3 QUANTITATIVE TOPOLOGY

3.1 Basic theory

The genus of a contour surface at some arbitrary threshold density is defined as

$$\text{genus} \equiv (\text{Number of holes}) - (\text{Number of isolated regions}), \quad (4)$$

where we have adopted the definition of WGM in which a 'hole' means a hole like that in a doughnut. An isolated (i.e. compact) region may be above or below the threshold density. For example, a torus has a genus of zero and an isolated sphere has a genus of -1. GMD used the Gauss–Bonnet theorem to express the genus, G_s, in terms of the integral over the contour surface of the Gaussian curvature, K, defined as the reciprocal of the product of the two principal radii of curvature,

$$G_s = -\frac{1}{(4\pi)} \int_s K dA. \quad (5)$$

As an example, consider the simple case of a density contour surrounding two isolated spherical clusters of radius r_s. The Gaussian curvature at all points on each sphere is $K = r_s^{-2}$ and the total area is $A = 2(4\pi r_s^2)$, which gives $G_s = -2$. This accords with equation (4) since we have two isolated regions and no holes. A contour threshold can be specified by its overdensity with respect to the mean $\overline{\rho}$, $\delta\rho/\overline{\rho} = (\rho - \overline{\rho})/\overline{\rho}$. The mean density contour of a Gaussian random field has a positive genus since its surface is all in one piece and it is multiply connected with many holes. The integrated curvature is negative because the two principal radii of curvature tend to point in opposite directions.

A formula for the genus per unit volume, g_s, was independently derived for Gaussian density fields by Doroshkevich (1970), Adler (1981), Bardeen *et al.* (1986) and Hamilton *et al.* (1986):

$$g_s = N(1 - \nu^2) e^{-\nu^2/2}, \quad (6a)$$

where

$$N = \frac{1}{(2\pi)^2} \left(\frac{\langle k^2 \rangle}{3} \right)^{3/2} = \frac{1}{(2\pi)^2} \left[\int k^2 P'(k)\, d^3k \Big/ 3 \int P'(k)\, d^3k \right]^{3/2}. \quad (6b)$$

Here ν is the number of standard deviations by which the contour threshold overdensity is above or below the average density and $P'(k)$ is the power spectrum of the density field smoothed using equation (1), $P'(k) = P(k) \exp(-k^2\lambda^2/2)$. Smoothing introduces a short-wavelength cut-off in the power spectrum.

For a Gaussian random field positive and negative fluctuations are statistically indistinguishable; hence the genus curve is symmetric about the mean density. The transition from positive to negative genus occurs at $\nu = \pm 1$ which marks the change from multiply connected contours to isolated clusters or voids. The shape of the $g_s(\nu)$ curve is independent of the power spectrum but its amplitude depends on the form of the smoothed power spectrum through equation (6b). The curve is independent of the normalization of $P(k)$, since any multiplicative constant cancels out.

For typical power spectra the normalization factor N depends mainly on the logarithmic slope of the spectrum at the smoothing scale, λ. For the simple case of a power-law spectrum, $P(k) \propto k^n$, N can be calculated explicitly:

$$N = \frac{1}{(2\pi)^2 \lambda^3} \left(\frac{3+n}{3} \right)^{3/2}, \qquad n \geq -3. \quad (7)$$

By studying the topology using different smoothing lengths, we constrain the shape of the power spectrum over a range of scales.

If the density fluctuations are non-Gaussian then the genus curve may be skewed. For example, if galaxies sat mainly in isolated clusters which resided in a connected, low-density background, the genus curve would be skewed to the left (WGM); one would have to look at a very small volume fraction of the data before isolated voids could be detected. On the other hand, if galaxies sat on the surfaces of 'bubbles', surrounding large voids, the genus curve would be skewed to

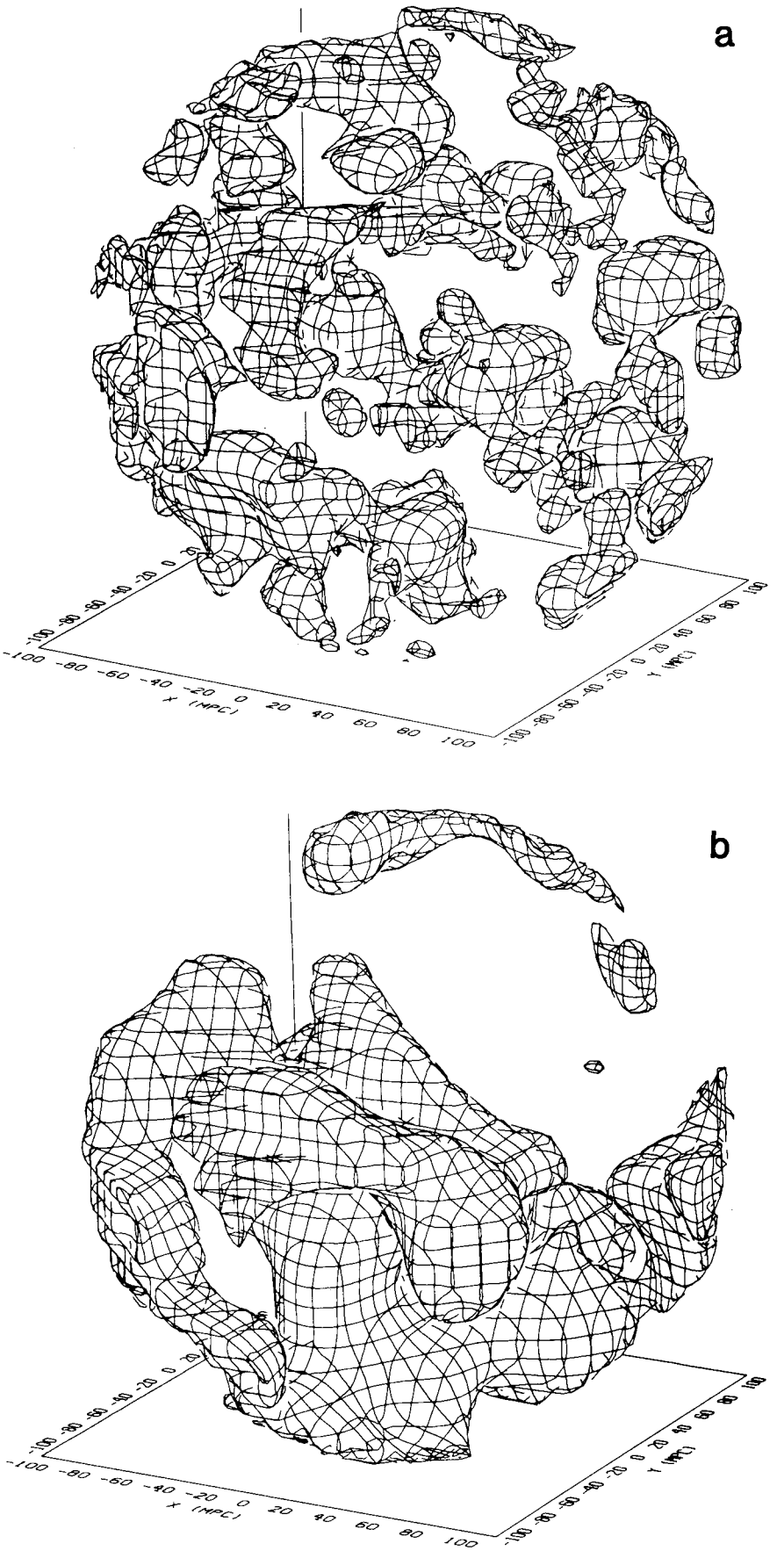


Figure 6. High-density regions of random density fields with power-law fluctuation spectra of spectral index $n = 1$ (a) and -2 (b). The smoothing lengths and volumes are the same as those used in Fig. 3(c).

the right. In this case the characteristic asymmetry in the genus curve could be detected if the typical void diameter were $\gtrsim 2\lambda$ (WGM). A symmetric genus curve does not guarantee that the density field is Gaussian, but any Gaussian field must have, on average, the symmetric genus curve of equation (6a).

For non-linear density fields WGM suggest specifying contour thresholds through the fractional volume f occupied by the high-density regions and relating these thresholds to the corresponding values of ν in a Gaussian field:

$$f=(2\pi)^{-1/2}\int_{\nu}^{\infty} \mathrm{e}^{-t^2/2}\, dt. \tag{8}$$

In a linear Gaussian density field this method yields the same result as specifying the threshold density directly in standard deviations from the mean. Non-linear gravitational evolution quickly distorts the Gaussian relation between density and fractional volume, but it tends to maintain a monotonic relation between initial and final density, mapping high-density regions of the initial conditions into high-density regions of the final conditions, low-density regions into low-density regions and so on. Furthermore, most proposed models of biased galaxy formation make the mass-to-light ratio a steadily increasing function of mass density, so contours of constant galaxy density are very close to the contours of constant mass density that enclose the same fractional volumes. Defining contour thresholds in terms of fractional volume takes advantage of these monotonic relations and allows one to recover contours of the smoothed, linear density fluctuations from the smoothed galaxy distribution. One can thereby address the question of whether the primordial fluctuations were Gaussian, even though the galaxy distribution today is non-Gaussian. The monotonic relations between initial mass density and final galaxy density are not exact. The effects of non-linearity and certain types of biasing on the genus curves of initially Gaussian models are discussed by Melott *et al.* (1988) and Park & Gott (1991). These studies show that non-linear evolution depresses the amplitude of the genus curve and that the combination of non-linearity and biasing tends to introduce a small shift in the direction of a 'meatball' topology. On the smoothing scales used in our QDOT analysis, we expect that at least the effects due to non-linearity should be small.

In our analysis we tabulate the smoothed density field on a cubic lattice. The curvature of a particular contour is concentrated at the vertices where square faces meet. (The faces have zero curvature.) The curvature at a vertex i on the surface is equal to its angle deficit $D_i=2\pi-\sum A_i$ where A_i are the surrounding vertex angles (see GMD). The genus of the surface can then be calculated directly as $G_s=-\sum D_i/4\pi$. We use the program CONTOUR described by Weinberg (1988) which uses this technique to measure the curvature of density contours.

3.2 Relation to counts-in-cells analysis

Efstathiou *et al.* (1990) and Saunders *et al.* (1991) have performed counts-in-cells analyses of the QDOT survey in cubical and Gaussian cells respectively. The variance of cell counts is related to the galaxy power spectrum, $P_g(k)$, by

$$\Delta^2(r)=\int_0^{\infty} P_g(k)\,\tilde{W}^2(kr)\,d^3k, \tag{9}$$

where $\tilde{W}(kr)$ is the Fourier transform of the cell window function. The variance is directly proportional to the amplitude of the power spectrum and, by measuring the variance in cells of several physical scales, r, one can measure the shape of the spectrum. In the linear biasing model, the galaxy power spectrum is just a constant multiple, b^2, of the underlying mass power spectrum. Since the variance is a mean-squared quantity, it is most sensitive to the regions of highest galaxy density, e.g. the Hercules supercluster in the QDOT survey.

A crucial feature of our topology analysis is that contours are defined in terms of enclosed fractional volume, or the corresponding value of ν from equation (8). As discussed in Section 3.1, the goal of this strategy is to recover contours of the initial density fluctuations from the present-day smoothed galaxy density field. One consequence is that the genus curve is independent of the amplitude of density fluctuations, so it provides no direct information on the amplitude of the power spectrum. However, the amplitude of the genus curve depends on the slope of the power spectrum near the smoothing length, or, more precisely, on the coherence scale, $\langle k^2\rangle^{-1/2}$, which determines the typical size of contour holes in a smoothed Gaussian field. Our analysis never assumes a linear relation between galaxy and mass fluctuations, only that a higher galaxy density corresponds to a higher mass density. It therefore assigns much less weight than counts-in-cells analyses to the highest density cells of the galaxy distribution. We show in Section 6 that the topology and counts-in-cells methods yield consistent results on the shape of the power spectrum. Since they measure the spectrum in different ways and weight the data quite differently, this consistency is reassuring.

The variance of cell counts alone gives no information on whether the underlying density field is Gaussian. However, the distribution $f(N)$ of counts-in-cells of fixed size [equivalent to the one-point distribution $P(\delta)$ of the field smoothed with the cell window function] can reveal non-Gaussian behaviour. Saunders *et al.* (1991) report marginal detections of skewness in the count distributions of the QDOT data at smoothing lengths of 14 and 28 h^{-1} Mpc (converting to our definition of Gaussian filter length). As pointed out by Coles & Frenk (1991) and Juszkiewicz (in preparation), skewness in the present-day galaxy distribution does not necessarily imply that the primoridal fluctuations were non-Gaussian, since non-linear evolution and biased galaxy formation naturally generate skewness from Gaussian initial conditions.

Defining contours in terms of fractional volume automatically eliminates any information on the one-point distribution of the smoothed density field. Topology and cell-count analyses are therefore sensitive to entirely independent sorts of non-Gaussian behaviour. While non-linearity and biasing can distort the genus curves of initially Gaussian models on scales smaller than the galaxy correlation length (Melott *et al.* 1988; Park & Gott 1991), significant departures from Gaussian topology on the 10–50 h^{-1} Mpc scales studied in this paper would probably indicate non-Gaussian features in the initial conditions. It would be perfectly consistent to find a

Gaussian topology and a non-Gaussian count distribution on these scales if the primordial fluctuations were Gaussian and the non-Gaussian $P(\delta)$ developed from some combination of non-linear evolution and biased galaxy formation.

4 ANALYSIS

4.1 The data

Our techniques for analysing redshift data are similar to those of Gott *et al.* (1989) with two exceptions: we treat the boundary differently when smoothing and we estimate error bars on the genus curve from Monte-Carlo simulations rather than bootstrap analysis. To construct a smooth density field, we bin the galaxies onto a 64^3 array and convolve it with the Gaussian filter of equation (1), setting the width λ equal to the mean intergalaxy separation at the edge of the survey, at distance $r_{\max}$:

$$\lambda = [S(r_{\max})]^{-1/3}, \tag{10}$$

where $S(r)$ is the selection function defined in Section 2. This choice of λ ensures that the density field is not undersampled.

Smoothing in Fourier space requires a periodic box; to accommodate this we embedded the survey region within a larger box. The space outside $r_{\max}$ must be treated in such a way as to minimize the effects of the boundary on the genus curve $g_s(\nu)$. If this space is left empty, then smoothing tends to lower the density at the edge of the survey, skewing the genus curve to the right. We tried several ways of dealing with this problem. (i) Include the galaxy distribution outside $r_{\max}$. (This is possible because the $n(v)$ distribution extends well into high redshifts.) (ii) Set the density in all the cells with $r > r_{\max}$ equal to $S(r_{\max})$ and then smooth the data. (iii) Smooth the data assuming zero density outside $r_{\max}$ and divide this array by the smoothed selection function $S(r)$, also with zero outside $r_{\max}$. This division corrects for the fraction of the smoothing volume outside the boundary. We tested each of these methods on artificial 'QDOT surveys' constructed from CDM simulations by comparing their genus curves with the theoretical curves obtained using equation (6) and the SCDM power spectrum. The least satisfactory method was (ii) which tended to overestimate the mean density at the edge of the array. Method (i) worked well for distances up to 100 h^{-1} Mpc, but beyond that Poisson noise in the survey outside $r_{\max}$ rendered the results unreliable. We adopted method (iii) which worked well at all distances. The galaxy selection function tabulated in a 64^3 lattice was multiplied by the volume of each cell visible to the observer, taking into account the *IRAS* mask (which excludes about 8 per cent of the sky above and below the galactic plane).

To obtain the maximum signal-to-noise ratio in the genus curve, $r_{\max}$ should be chosen to give the largest number of smoothing volumes, $V_{sm} = \pi^{3/2}\lambda^3$, per survey volume. The dashed line in Fig. 2 shows the number of resolution elements, N_{res}, as a function of distance,

$$N_{res} = \frac{\omega_s r^3}{3\pi^{3/2}\lambda^3}, \tag{11}$$

where ω_s is the solid angle of the survey. The number of resolution elements reaches a maximum of 80 at 100 h^{-1} Mpc where the mean galaxy separation is 20 h^{-1} Mpc. This value of N_{res} is larger than that of any previous redshift survey. (The CfA1 survey, for example, has a maximum of 54; see the table in Gott *et al.* 1989.) We examined the topology on a range of smoothing scales between 10 and 50 h^{-1} Mpc. Table 1 lists the smoothing lengths we considered, the distance, $r_{\max}$, at which we cut off the data, the values of $S(r_{\max})$ and the number of resolution elements within the survey volume.

Figs 7(a)–(d) show our estimates of the genus for the QDOT data, for four different smoothing scales. The filled squares give the values of the genus at 15 threshold overdensities between $\pm 2\sigma$ from the mean density contour. The error bars on the QDOT data are 1σ and will be discussed in Section 4.2. The smooth curves show the best-fitting theoretical curves to the data using equation (6a) with the normalization factor N as a free parameter. The corresponding values of the effective spectral index are discussed in Section 5. The one-in-six sampling strategy has resulted in a loss of information on small scales which is apparent from the large scatter in Fig. 7(a). However, we gain on larger scales and Figs 7(b)–(d) show that the topology of the QDOT survey is similar to that of a Gaussian random field. As expected from the number of resolution elements, we obtain the best results using a smoothing length of 20 h^{-1} Mpc. On this scale, there is a slight shift in the genus curve to the left in the direction of an 'isolated cluster' topology, but not at a very significant level. The consistency of the data with the Gaussian hypothesis is discussed further in Section 4.3.

4.2 Error estimates from N-body simulations

There are two kinds of sampling errors which contribute to the uncertainties in our measurement of the genus: errors due to sparsely sampling the galaxy distribution and errors due to having surveyed only one region of space. We estimate these uncertainties using the artificial catalogues constructed from our SCDM simulations. From five different simulations we generated five fully sampled 'galaxy' catalogues, each

Table 1. Characteristics of the QDOT and CDM predictions.

λ /h^{-1}Mpc	r_{max} /h^{-1}Mpc	$S(r)$ /h^{3}Mpc^{-3}	N_{res}	n (QDOT)	n_{eff} (SCDM)
10	40	1.0×10^{-3}	48	-1.61 ± 0.80	-0.93(0.8σ)
14	70	3.0×10^{-4}	70	-1.04 ± 0.45	-0.59(1.0σ)
20	100	1.2×10^{-4}	80	-0.79 ± 0.35	-0.25(1.6σ)
28	135	4.2×10^{-5}	66	-0.88 ± 0.40	0.07(2.4σ)
40	175	1.6×10^{-5}	52	-0.62 ± 0.50	0.34(1.9σ)
50	210	8.1×10^{-6}	46	-0.71 ± 0.65	0.48(1.8σ)

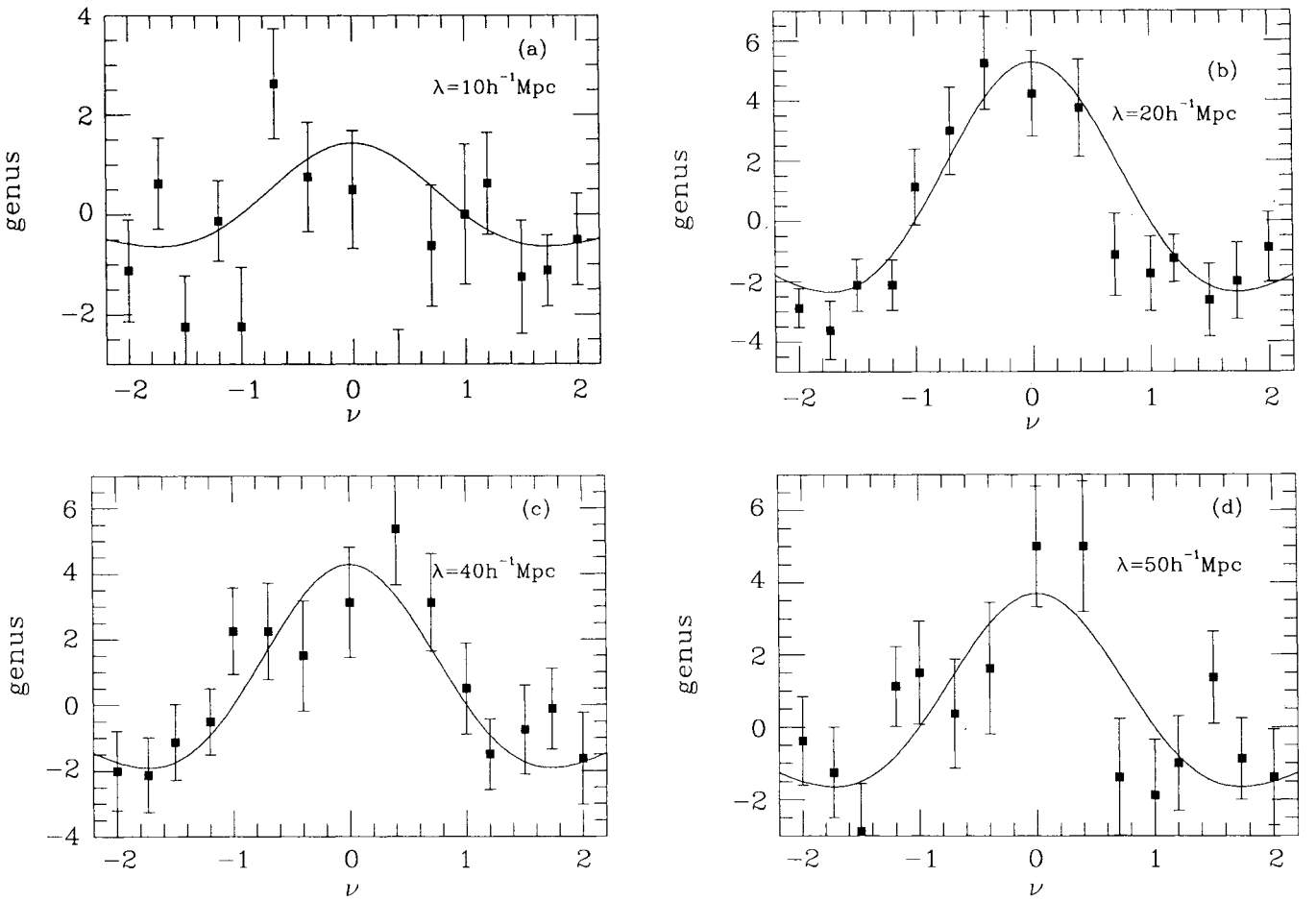


Figure 7. The genus of the QDOT survey for four different smoothing scales. On the horizontal axis, ν is the number of standard deviations from the mean density contour. The solid curves are the best-fitting theoretical curves of the form given by equation (6a).

containing over 12 000 galaxies with the selection function of the QDOT survey. From each of these we extracted six one-in-six randomly sampled subcatalogues similar to the QDOT survey.

The effect of shot noise in our sparsely sampled catalogues can be visualized by comparing contour plots of them with those of the fully sampled parent catalogue. Fig. 8(a) shows one of the parent catalogues contoured using $r_{\max} = 100\ h^{-1}$ Mpc and $\lambda = 20\ h^{-1}$ Mpc; Figs 8(b)–(d) show three one-in-six daughter catalogues. The global structure and dominant features of the contours do not change appreciably although shot noise introduces differences in detail. For a more quantitative measure of shot noise, we compare the density per pixel in the parent and daughter catalogues using smoothing lengths equal to $\lambda/2$, $3\lambda/4$, λ and $4\lambda/3$, where $\lambda = 20\ h^{-1}$ Mpc and $r_{\max} = 100\ h^{-1}$ Mpc. Fig. 9 shows that if the data are smoothed on a scale smaller than λ, the mean intergalaxy separation at $r_{\max}$, the agreement between the fully sampled and sparsely sampled density fields is poor, but the agreement improves steadily as the smoothing length is increased and is quite acceptable when it is set equal to λ. Choosing a smoothing length of $4\lambda/3$ would further decrease the shot noise, but this would lower the number of independent smoothing volumes. We also compared contour plots for the fully sampled and sparsely sampled distributions at these smoothing lengths and confirmed that λ should be at least as large as the mean intergalaxy separation at the edge of the survey. As λ was reduced, the number of spurious structures started to increase rapidly.

In Fig. 10 we show as filled circles the mean genus curves for our five fully sampled catalogues and as open circles the average genus curve of the 30 randomly sampled catalogues, with parameters $\lambda = 20\ h^{-1}$ Mpc and $r_{\max} = 100\ h^{-1}$ Mpc. Comparing these, we can see that random sampling has not introduced any strong systematic bias and that we have successfuly recovered the topology of the fully sampled simulations. The smooth curve in Fig. 10(a) is the theoretical genus curve for the SCDM model in the linear regime, calculated from equations (6) and the expression for the power spectrum given in equation (2) of Davis *et al.* (1985). The evolved simulations are a good fit to this curve, demonstrating that the non-linear effects and effects due to biasing have not significantly perturbed the topology of the initial fluctuations. The error bars in Fig. 10(a) show the scatter between the five fully sampled catalogues and represent realization-to-realization differences. The error bars in Fig. 10(b) show the mean of the 1σ scatter obtained from sparsely sampling each simulation six times. Fig. 10(c) shows both of these

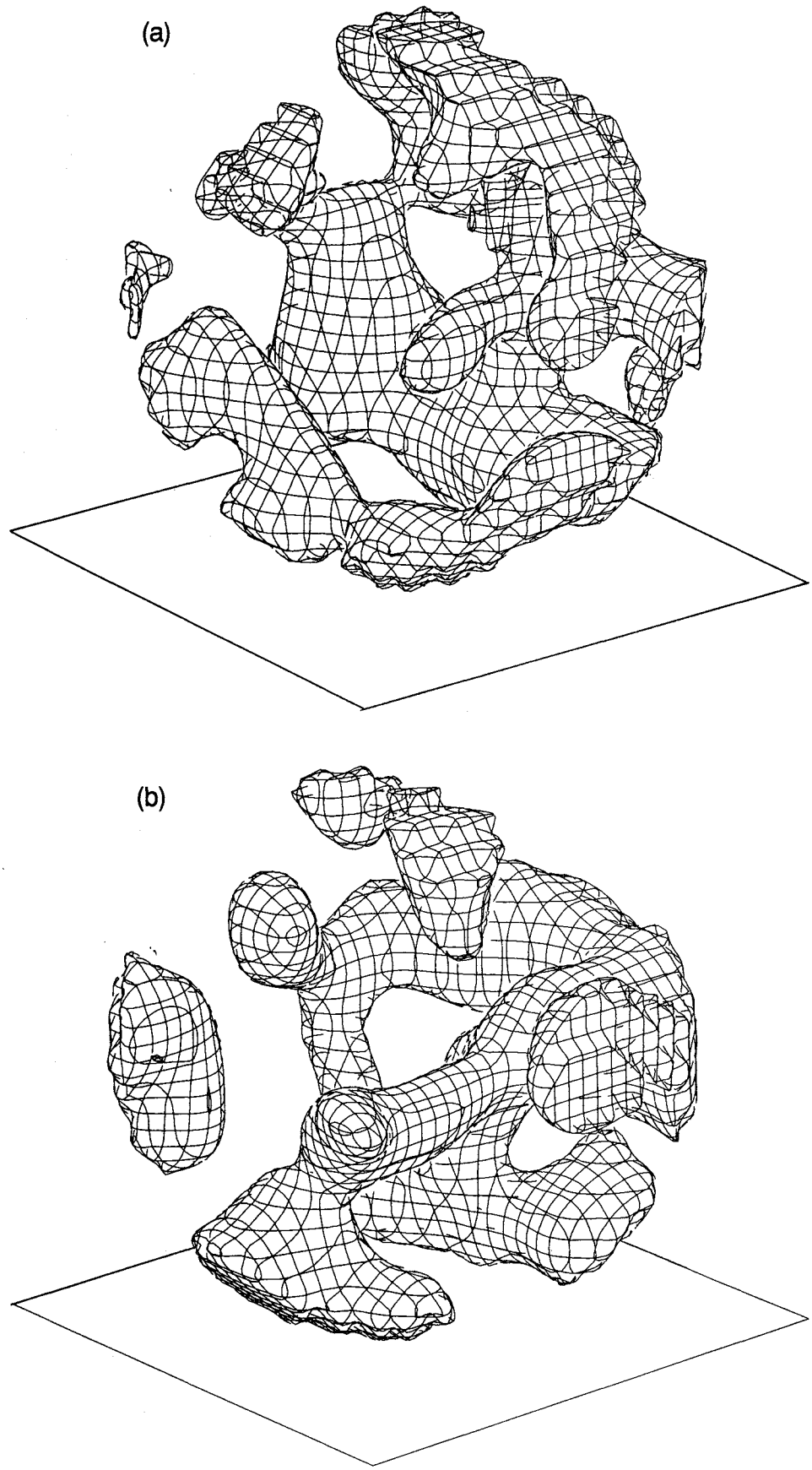


Figure 8. (a) High-density regions of a fully sampled artificial catalogue constructed from a cold dark matter N-body simulation. The smoothing scale is $\lambda = 20\ h^{-1}$ Mpc and the catalogue extends out to $r_{max} = 100\ h^{-1}$ Mpc. (b–d) High-density regions of subcatalogues extracted from that depicted in (a) by randomly sampling the galaxy distribution at a rate of one-in-six.

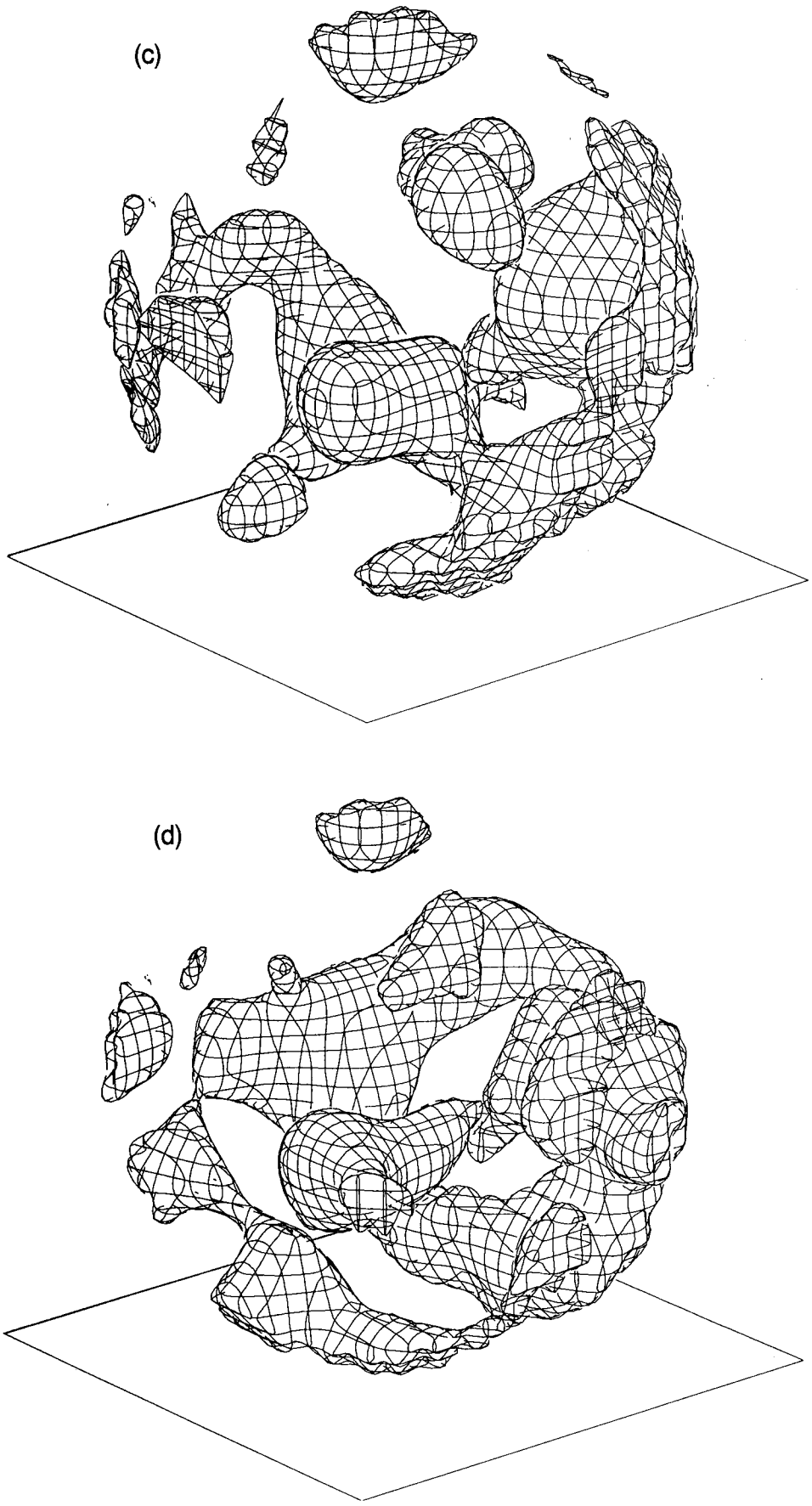


Figure 8 – *continued*

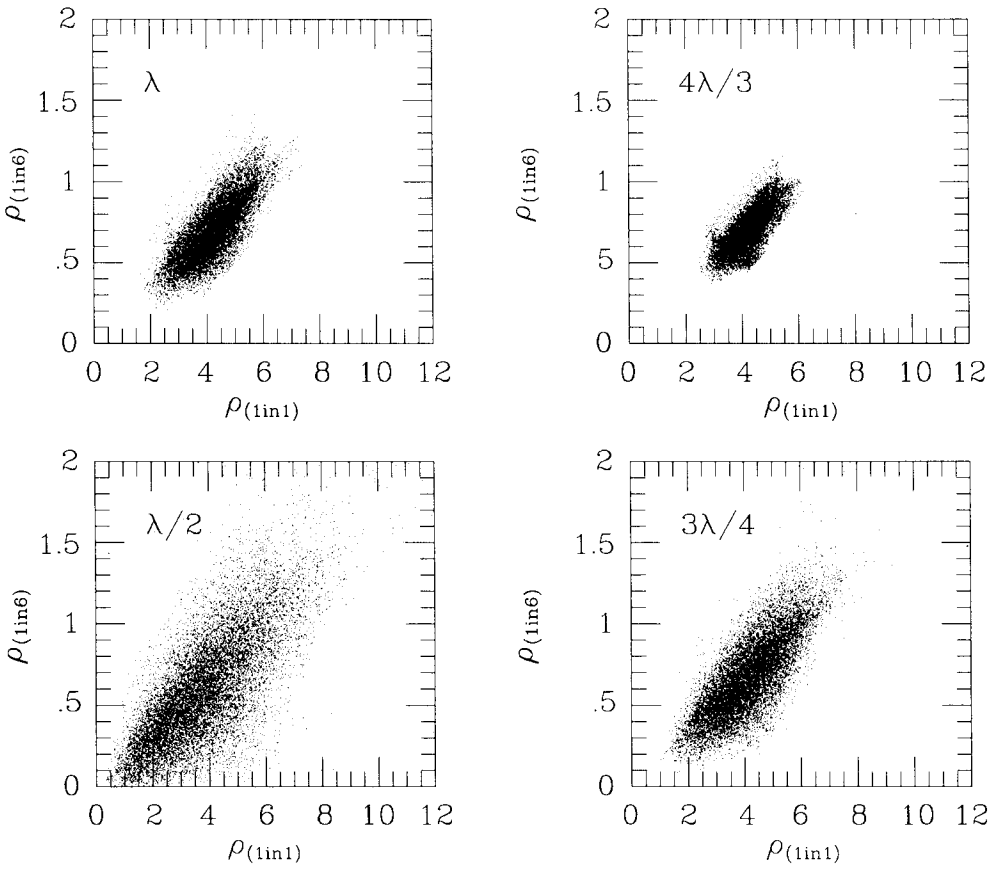


Figure 9. Pixel-by-pixel comparison of the density fields in a fully sampled catalogue and a sparsely sampled subcatalogue. The different panels correspond to different smoothing lengths; $\lambda = 20\ h^{-1}$ Mpc and $r_{\max} = 100\ h^{-1}$ Mpc.

errors added in quadrature and should be compared with Fig. 10(d) which shows 1σ errors obtained from randomly sampling each of the five simulations once only. These two methods give very similar error bars for the model, roughly a mean error per genus measurement of ± 1.2. It is also clear that the scatter introduced by the one-in-six sampling is comparable to the scatter between realizations. The errors plotted in Fig. 7 were derived from the model and are the sum in quadrature of the random sampling errors and the realization-to-realization errors, calculated on each smoothing scale.

Gott *et al.* (1989) used a bootstrap resampling procedure for estimating errors in the genus curve. [Note that our QDOT results in Fig. 7 should be compared to the 'raw' data of Gott *et al.* (1989) rather than to their 'bootstrap average' points which are less noisy because of the averaging.] It is interesting to compare these with the errors obtained above. To perform the bootstrap resampling, we assigned to each galaxy a position randomly chosen from the original list of positions. Some positions may have more than one galaxy and some may be left empty with probability given by the Poisson formula $P(N) = (N!)^{-1} e^{-N}$. Bootstrap resampling in the artificial catalogues gave a mean error per genus of ± 1.35, similar to the error found by resampling the original QDOT survey. This is about 10 per cent larger than the errors estimated from the simulations. It is interesting to note that according to the simulations, had we measured redshifts for all 13 000 galaxies in the fully sampled QDOT catalogue, the mean genus values could have been determined only to ± 0.8. However, a better dynamic range would be attainable by using smaller smoothing lengths.

4.3 Are the observed genus curves consistent with Gaussian initial conditions?

In this section we compare the genus curve of the QDOT data with the genus curves predicted for certain Gaussian and non-Gaussian random fields. A simple χ^2 test is not suitable for this comparison because points on the genus curve are not independent. As an alternative, we construct Monte-Carlo realizations of a particular model and calculate the likelihood of the data. Given the mean genus curve for a model, $G_{\rm m}$, we form the sum of absolute differences between the genus values of the data and the genus values of the model:

$$D_{\rm d,m} = \sum_{\nu_i = 1}^{15} \mathrm{abs}\,(G_{{\rm d},\nu_i} - G_{{\rm m},\nu_i}). \qquad (12)$$

By constructing many realizations of the model, we calculate the sampling distribution of this quantity and hence the significance level at which the model is rejected by the data. We

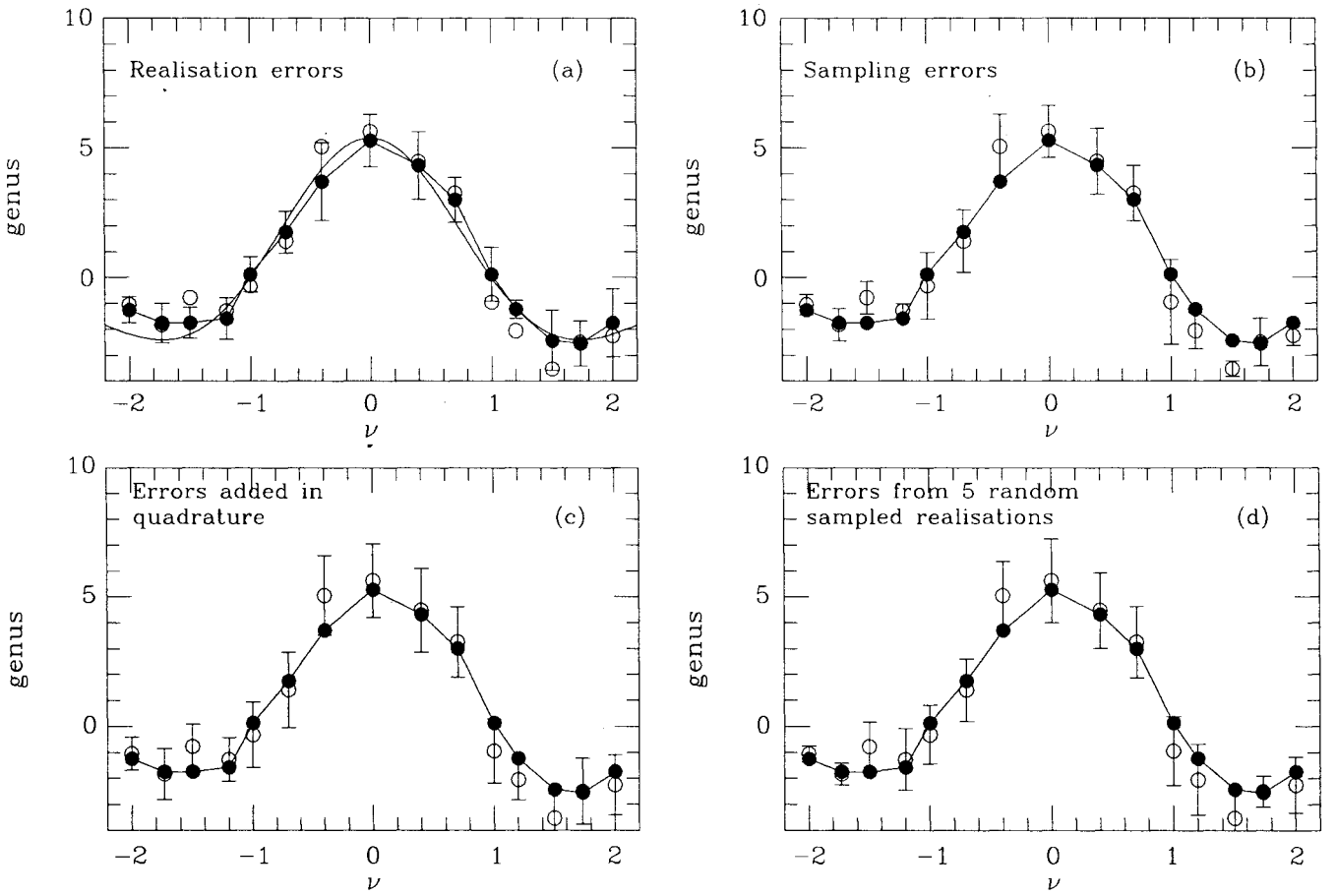


Figure 10. Mean genus curves for artificial 'QDOT' surveys constructed from five cold dark matter N-body simulations. The filled circles are mean values from five fully sampled catalogues. The open circles are mean values from 30 randomly sampled (at a one-in-six rate) subcatalogues, six from each parent catalogue. The parameters are $\lambda = 20\ h^{-1}$ Mpc and $r_{\max} = 100\ h^{-1}$ Mpc. The smooth curve in (a) is the theoretical genus curve for the cold dark matter model. The error bars are 1σ and arise from: (a) the scatter between the five fully sampled catalogues; (b) the scatter between the 30 sparsely sampled subcatalogues; (c) the errors from (a) and (b) added in quadrature; (d) the errors from five sparsely sampled subcatalogues, each from a different simulation.

consider simple examples of Gaussian and non-Gaussian fields.

We constructed a Gaussian model by generating a random-phase density field in a 64^3 grid with a power-law spectrum, $P(k) \propto k^n$, with $n = -1$. To calculate the genus curve of this model we divided the smoothed density field by the smoothed mask and used the same value of $\lambda / r_{\max}$ as for the QDOT data. We repeated this procedure 200 times using $\lambda = 20\ h^{-1}$ Mpc and $r_{\max} = 100\ h^{-1}$ Mpc. From the sum-of-differences method we find that the model can only be rejected at the 0.4σ level, indicating that the QDOT data are consistent with this particular realization of the Gaussian hypothesis. The shape of the QDOT genus curve is also similar to that for the SCDM model since the starting point of this model is, of course, a Gaussian random field. However, since the *amplitude* of the genus curve also depends on the shape of the power spectrum, we defer a detailed comparison with the SCDM model until Section 5.

As an example of a non-Gaussian field we consider galaxies distributed on the surfaces of large bubbles, as prescribed by the Voronoi foam model (van de Weygaert 1991 and references therein). The Voronoi foam model is a simple scheme for dividing space into a tessellation of cells. Points or 'seeds' are placed at random within the space and cell boundaries or walls are defined as the loci of points which are equidistant from two seeds and which are not closer to any other seed. The only free parameter is the mean separation of the seeds and choosing a value of $100\ h^{-1}$ Mpc gives rise to two interesting results. (i) If rich clusters form at the vertices of the cells, the observed abundance and correlation function of Abell clusters can be reproduced (van de Weygaert & Icke 1989). (ii) A random line of sight through the tessellation will intersect cell walls with an expected separation of $137\ h^{-1}$ Mpc, with quasi-periodic separations fairly likely (Coles 1990). This is close to the $128\ h^{-1}$ Mpc, quasi-periodic spacing found by Broadhurst *et al.* (1990) in deep pencil-beam surveys of the galaxy distribution. In a Voronoi foam model we expect the genus curve to be skewed to the right.

Within a cube of side $200\ h^{-1}$ Mpc we placed eight seeds at random. 1000 'galaxies' were distributed randomly on the surface of the tessellation. Using this number of galaxies mimics the sampling effects in the QDOT data. In the observed galaxy distribution, voids of this size are not com-

pletely empty; we therefore tested three Voronoi foam models, each with a fraction of the galaxies repositioned at random throughout the volume. We calculate the genus curve of the models in a similar way to the data after smoothing the density field and dividing by the smoothed mask; again we compare the topology on our most accurate smoothing length of 20 h^{-1} Mpc.

The closed circles, open circles and open squares in Fig. 11 show the mean genus curves from 100 realizations of the Voronoi foam model, with a fraction 0, 15 and 50 per cent respectively, of galaxies distributed at random throughout the volume. The error bars on the open squares represent the 1σ scatter for the 100 realizations of this particular model. We also tested a model with 32^3 galaxies to assess sampling effects and found a mean genus curve very similar to the model containing only 1000 galaxies. The genus curves are all skewed to the right, away from the symmetric shape expected for a Gaussian distribution. As more random galaxies are added, the genus curves become more symmetric. We applied the same test to the Voronoi foam model as we previously applied to the Gaussian random-phase model. The Voronoi foam model for large-scale structure with $\approx 100\ h^{-1}$ Mpc cells can be ruled out at $\sim 7\sigma$ if the voids are truly empty, at $\sim 5\sigma$ if 15 per cent of galaxies populate the voids, and $\sim 3\sigma$ if 50 per cent of galaxies populate the voids.

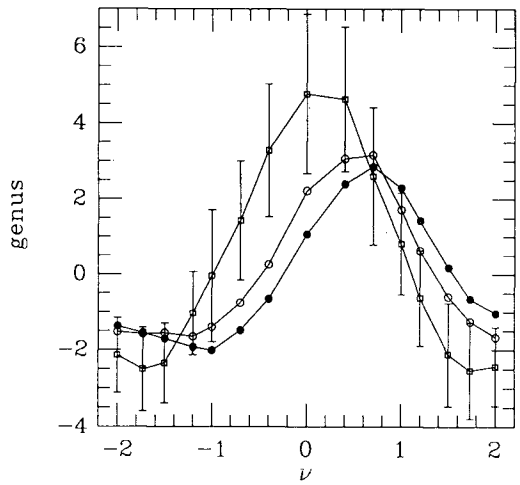


Figure 11. The mean genus curves from 100 realizations of Voronoi foam models, with 0 per cent (filled circles), 15 per cent (open circles) and 50 per cent (open squares) of the galaxies placed at random within the volume. The error bars show the run-to-run 1σ dispersion for the model with 50 per cent random galaxies.

Fig. 12(a) shows the 10 genus curves from the Gaussian model which agree best with the QDOT data, i.e. the curves with the 10 lowest values of $D_{\mathrm{d,m}}$ from equation (12). Similarly, Fig. 12(b) shows the 10 best genus curves from the foam model with 15 per cent random galaxies. These plots provide a simple visual comparison between the data and the models; while it is clear that each of the Gaussian genus curves provides a respectable match to the data, none of the Voronoi foam genus curves (and these are the best fits) comes close to providing a good fit.

As a final test, we compared the Gaussian model directly with the Voronoi foam model using a similarity scoring method. This method is akin to our first test except that we assign a score to the data based on the mean sum of the differences between the data and the foam model minus the mean sum of the differences between the data and the Gaussian model. (See Gott *et al.* 1989 for further details.) By comparing the Gaussian model directly with the three Voronoi foam models we can calculate the relative likelihood that the observed genus curve is representative of a Gaussian topology. We find $P_{\mathrm{Gauss}} > 99.99$, > 99.99, $= 99.96$ per cent for Voronoi foam models with 0, 15 and 50 per cent respectively of the galaxies distributed at random.

From these tests we conclude that the topology of the QDOT survey is inconsistent with the non-Gaussian Voronoi foam model. By contrast, a Gaussian random field provides a good match to the topology of the survey. An important corollary is that we can use the analytic relations given in Section 3 to constrain the *shape* of the power spectrum of density fluctuations.

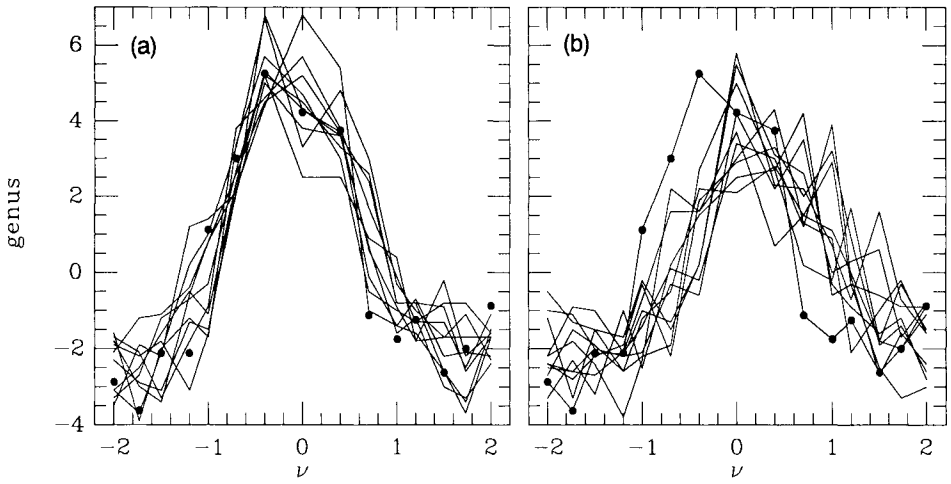


Figure 12. Comparison of the QDOT survey with a Voronoi foam model and a Gaussian random field with $P(k) \propto k^{-1}$. The filled circles give our estimates of the genus for the QDOT survey, with $\lambda = 20\ h^{-1}$ Mpc and $r_{\mathrm{max}} = 100\ h^{-1}$ Mpc. The lines show the 10 genus curves which are most like the data in (a) the Gaussian model and (b) the Voronoi foam model.

5 THE SHAPE OF THE POWER SPECTRUM OF DENSITY FLUCTUATIONS

We showed in Section 4 that the shapes of the genes curves for the QDOT survey are consistent with the characteristic shape for a Gaussian random field. In this section we use the *amplitude* of the genus curves as a function of smoothing length to constrain the shape of the power spectrum of the density field. We estimate the amplitude, $g_s(0)$, by least-squares fitting the random-phase form (equation 6) to the measured genus and substituting this estimate of N into equation (7) to obtain an 'effective spectral index', $n_{\rm eff}$. This index quantifies the coherence of features in the smooth density contours. For a Gaussian field, $n_{\rm eff}$ gives the local slope of the power spectrum at the smoothing length λ, but for a non-Gaussian field it can, in principle, be different to the value obtained from a direct estimate of $P(k)$. Because we use fractional volume to define density contours, our measure of the spectrum is less sensitive than direct estimates to non-linearities and to the assumed way in which *IRAS* galaxies trace the mass distribution.

Estimating the uncertainty in $n_{\rm eff}$ is not straightforward because points on the genus curve are not independent. A rough estimate of the uncertainty follows from simply taking the error in the amplitude of the genus curve to be the mean error per genus measurement (computed using the simulations as described in Section 4.2). The errors vary with the volume and smoothing length under consideration so the mean error per genus measurement must be recalculated for each scale (i.e. ± 1.2 for $\lambda = 20\ h^{-1}$ Mpc). The estimate of $n_{\rm eff}$ and the associated 1σ errors are listed in the fifth column of Table 1. Our most accurate estimate of the slope of the power spectrum corresponds to the scale with the maximum number of resolution elements, $\lambda = 20\ h^{-1}$ Mpc. On this scale we find $n_{\rm eff} = -0.79 \pm 0.35$, with values in the range $n_{\rm eff} > 0.2$ or < -1.8 ruled out at more than 3σ.

The above results are summarized in Fig. 13. The filled circles show our estimates of the effective power spectrum slope at a range of scales between 10 and 50 h^{-1} Mpc. Our data are consistent with a constant value, $n_{\rm eff} \approx -1$, over this range. The solid line gives $n_{\rm eff}$ for the SCDM spectrum calculated from equations (6) and (7). For comparison we also show $n_{\rm eff}$ for the Abell cluster sample (filled triangle) and the Giovanelli–Haynes galaxy sample (star) taken from Gott *et al.* (1989). The topology of the QDOT survey agrees well with that of the GH survey but is marginally inconsistent with that of the Abell cluster sample. This discrepancy could be due to systematic effects in the Abell cluster catalogue or to a difference between the way galaxies and clusters trace the large-scale mass distribution.

In the last column of Table 1 we list the slope of the SCDM spectrum on each scale we have studied. In brackets we give the significance level at which this slope is rejected by the QDOT data, assuming the error estimate described above. On our most accurate scale of 20 h^{-1} Mpc, the slope of the SCDM spectrum is $n_{\rm eff} = -0.25$, 1.6σ away from the QDOT result. The largest discrepancy occurs on 28 h^{-1} Mpc, scales where the SCDM spectrum has $n_{\rm eff} \approx 0$ and the QDOT data have a spectrum of slope $n_{\rm eff} = -0.88 \pm 0.4$.

To obtain an alternative estimate of the uncertainties in our determination of $n_{\rm eff}$, we carried out a series of Monte-Carlo simulations similar to those described in Section 4.3, but with various values of the power-law exponent n; we also calculated Monte-Carlo realizations of density fields with the power spectrum of the SCDM model in the linear regime. Fig. 14 shows mean genus curves from 200 simulations with $n = 1, 0, -1, -2$ and with the SCDM power spectrum, all assuming $\lambda = 20\ h^{-1}$ Mpc and $r_{\rm max} = 100\ h^{-1}$ Mpc. Fig. 15(a) shows the distributions of fitted amplitudes of the individual genus curves, $g_s(0)$, for the various models. These distributions appear symmetric about the mean and they become broader for simulations with less power on large scales. Fig. 15(b) shows the best-fitting Gaussian to each distribution; the reduced χ^2 is less than unity in all cases. Comparison of the amplitude of the genus curve for the QDOT survey with these distributions gives rejection levels for each model.

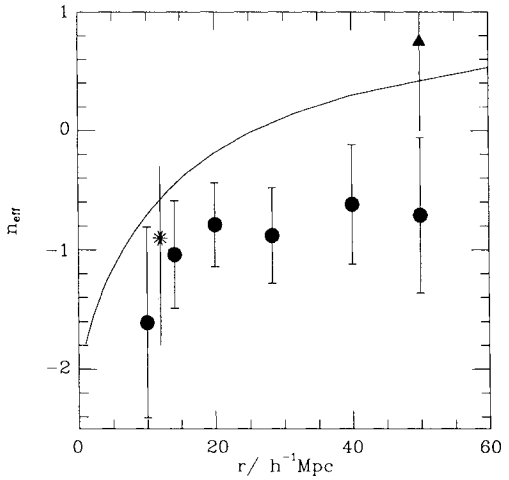


Figure 13. The effective spectral index, $n_{\rm eff}$, as a function of smoothing scale. The filled circles give results for the QDOT survey. The star and the filled triangle, taken from Gott *et al.* (1989), are for the Giovanelli–Hanes survey and the Abell cluster sample respectively. The solid line gives the effective index for the standard cold dark matter model.

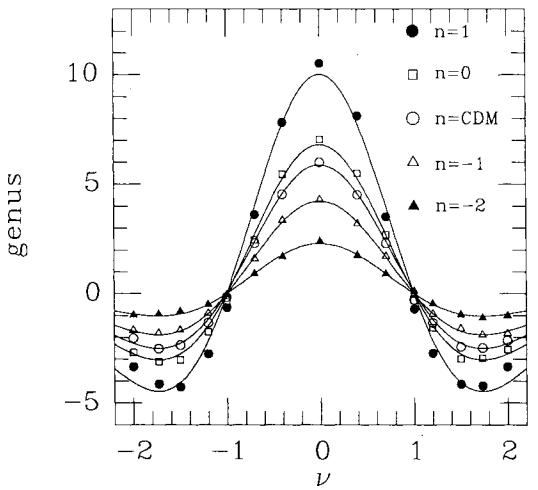


Figure 14. Mean genus curves from Monte-Carlo simulations of power-law fluctuation spectra of index $n = +1, 0, -1, -2$ and of the standard dark cold matter power spectrum. The averages are over 200 simulations of each model. The smooth curves are the theoretical curves calculated using equations (6) and (7).

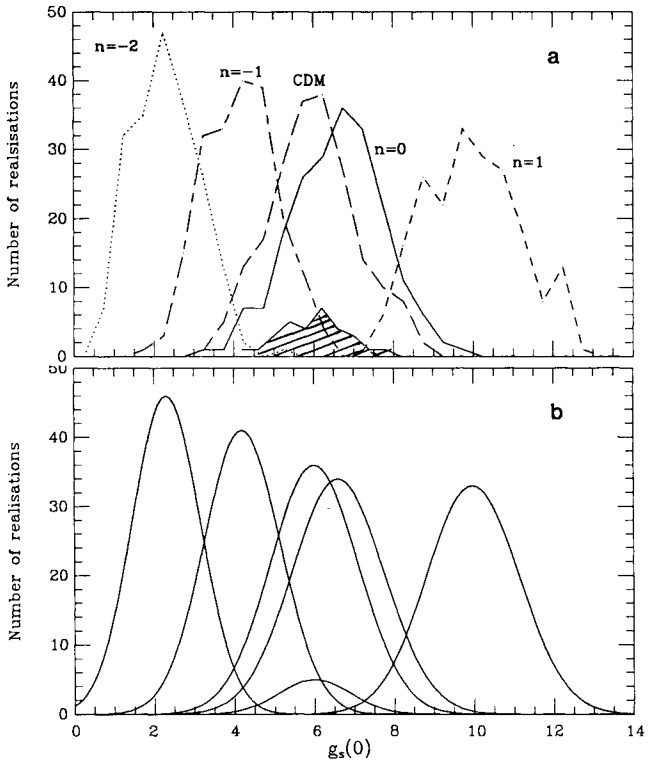


Figure 15. (a) The distribution function of the amplitude (from least-squares fitting) of the genus curve in the 200 Monte-Carlo simulations of Fig. 14. The distribution of sampling errors in cold dark matter N-body simulations is shown as the hatched region underneath the distribution of amplitudes for this model. (b) Best-fitting Gaussians to the distributions plotted in (a).

Table 2. Rejection levels for CDM and power-law models.

Model	$\lambda = 10h^{-1}$Mpc $r_{max} = 40h^{-1}$Mpc	$\lambda = 20h^{-1}$Mpc $r_{max} = 100h^{-1}$Mpc	$\lambda = 40h^{-1}$Mpc $r_{max} = 175h^{-1}$Mpc
CDM	$+0.7\sigma$	$+1.5\sigma$	$+1.7\sigma$
n = +1	$+3.0\sigma$	$+4.8\sigma$	$+2.9\sigma$
n = 0	$+1.5\sigma$	$+2.0\sigma$	$+1.1\sigma$
n = -1	$+0.3\sigma$	-0.2σ	-0.3σ
n = -2	-0.1σ	-3.1σ	-2.1σ

These are listed in Table 2 for three different smoothing scales, $\lambda = 10$, 20 and 40 h^{-1} Mpc. The results from these Monte-Carlo simulations agree well with our previous simpler estimates.

Our analysis of the artificial 'QDOT' surveys in Section 4.2 showed that the scatter in the genus measurements due to the one-in-six sampling is comparable to the scatter between different realizations of the survey volume. At least for the case of SCDM we can quantify the effect of the sparse sampling on the estimate of $n_{\rm eff}$ from the scatter in the amplitudes of the genus curves in our artificial catalogues. For each catalogue, we calculate the difference between the amplitude of the genus curve of the fully sampled simulation and those of its six one-in-six subsets. The small shaded histogram underneath the SCDM curve in Fig. 15(a) shows this distribution computed from five different SCDM simulations. A Gaussian fit gives a width of 0.9, whereas the distribution of amplitudes for the full SCDM realizations has a width of 1.1. A convolution of these two distributions would yield a broadened Gaussian of width 1.4; thus, the confidence levels listed in Table 2 are ~ 25 per cent too high.

Fig. 13 shows that the power spectrum of the QDOT data is steeper (more power in small wavenumbers) than that of the SCDM model on all the smoothing scales we have considered. We have calculated the significance of this difference at each scale, but if different smoothing scales were independent, then the overall significance would be higher. To test the interdependence of the genus curves at different smoothing scales, we generated a series of Monte-Carlo realizations of the SCDM power spectrum. We calculated the genus curves of the density fields as before, but we then extracted the 16^3 core of the lattice of each realization and recalculated the genus curve using a smaller smoothing length. The amplitude of each genus curve was compared with the theoretical expectation in order to ascertain whether genus curves on two different scales are independent. For example, from 100 realizations we measured the amplitudes of the genus curves for density fields smoothed on 20 and 28 h^{-1} Mpc. If these scales were independent then the amplitudes would be distributed evenly above and below the expected value given by equation (6). We found that 65 pairs lay either both higher or both lower than the expected value and 35 pairs had amplitudes one higher and one lower, indicating that the genus curves from the same underlying density field with these two smoothing lengths are not independent. However, repeating this test on smoothing scales of 20 and 40 h^{-1} Mpc, we found that 45 pairs had amplitudes both under or over the expected amplitude while 55 pairs had amplitudes one higher and one lower. This shows that, at least for SCDM, our measurements of the slope of the power spectrum are independent on smoothing scales of 20 and 40 h^{-1} Mpc. Thus, the combined significance of these 1.5σ and 1.7σ discrepancies (Table 2) is $(1.5^2+1.7^2)^{1/2}=2.3\sigma$. These numbers, however, are also subject to the ~ 25 per cent correction due to the sampling uncertainties discussed above so the overall significance is reduced to $\sim 2\sigma$.

6 DISCUSSION AND CONCLUSIONS

The topological analysis discussed here complements previous studies of the distribution of *IRAS* galaxies using the QDOT survey. In particular, counts-in-cells analyses by Efstathiou *et al.* (1990) and Saunders *et al.* (1991) give the variance, $\Delta^2(r)$, of the density field smoothed on scale λ (see equation 9). While the topology is sensitive to the shape of the power spectrum, the variance of counts-in-cells depends both on the shape and the amplitude of the power spectrum. Fig. 16 gives this variance as a function of smoothing scale as estimated by Efstathiou *et al.* (squares) and Saunders *et al.* (circles). We have taken into account the different smoothing functions used in these studies, cubical cells in the first and Gaussian spheres in the second. The relation between these depends slightly on the spectral index; we have used the conversion appropriate to a power-law spectrum with $n=-1$. In this case, the variance in Gaussian spheres of radius $r\,h^{-1}$ Mpc is the same as that in top-hat spheres of radius $2.2\,r\,h^{-1}$ Mpc. [Recall that our definition of a Gaussian window in

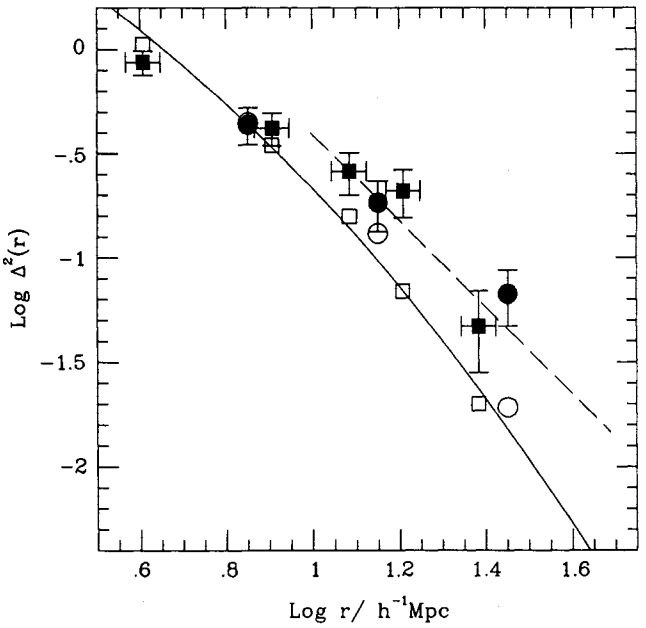


Figure 16. The variance of counts-in-cells in the QDOT survey as measured by Efstathiou *et al.* (1990; solid squares) and Saunders *et al.* (1991; solid circles). The open squares and circles are their respective measurements of the variance from the same N-body simulations of a cold dark matter universe used in this paper. The vertical error bars are 1σ. The horizontal error bars on the squares reflect the uncertainty in converting between different smoothing functions. The solid curve gives the variance calculated for the *linear* cold dark matter spectrum in *real* space using equation (9) normalized by assuming a variance in top-hat spheres of radius 8 h^{-1} Mpc equal to 1/1.4. The dashed line shows the variance for the power-law spectrum of index $n=-0.9$ found in this paper. This line was placed at the height that gives the best fit (by eye) to the results of the counts-in-cells analyses.

equation (1) differs from the definition of Saunders *et al.* by a factor of $1/\sqrt{2}$.] The predictions of the SCDM model, calculated from the same N-body simulations used in this paper, are shown as open symbols in Fig. 16.

The variance of galaxy counts in top-hat spheres of radius 8 h^{-1} Mpc is close to unity for optical galaxies (Davis & Peebles 1983). This scale corresponds to a 5.1 h^{-1} Mpc Gaussian window; on this scale Fig. 16 shows the variance of the *IRAS* galaxy counts to be well below unity. A least-squares fit to the first three data points in this figure gives Δ (5.1 h^{-1} Mpc) $=0.67\pm0.11$. This is consistent with a correlation analysis (in preparation) which shows that *IRAS* galaxies are somewhat less clustered than optically selected galaxies. The solid line gives the theoretical variance calculated directly from the *linear theory* cold dark matter power spectrum in *real* space, normalized so that the rms fluctuations in the *IRAS* galaxy distribution are 1.0/1.4 times the observed rms fluctuations of optical galaxies in top-hat spheres of radius 8 h^{-1} Mpc. The dashed line in Fig. 16 has the slope $n=-0.9$ obtained in Section 5, over scales of 10 to 50 h^{-1} Mpc. Since the topology analysis does not constrain the amplitude of the power spectrum, we have arbitrarily placed this line at the height at which it gives the best possible agreement (as judged by eye) with the results of the counts-in-cell analyses. It is reassuring that the topology and counts-in-cells techniques yield consistent results, even though they are sensitive to different aspects of the data.

On scales $\lesssim 15$ h^{-1} Mpc the QDOT power spectrum agrees well with the standard cold dark matter model, as can be seen either from the solid line or the open symbols in Fig. 16; on larger scales the data fall off less steeply than the model predictions. The largest discrepancy with the topology analysis occurs at ~ 30 h^{-1} Mpc and is significant at the $\sim 2\sigma$ level. This result is somewhat less significant than that obtained from the counts-in-cells analyses which weight the data in a very different way. We expect the topology test to be rather less sensitive than the counts-in-cells method to the details of the way in which the *IRAS* galaxies are assumed to trace the mass. Nevertheless, we cannot rule out the possibility that the large-scale distributions of galaxies and mass may not trace each other in the simple way posited by the linear biasing model. Indeed, recent work by Babul & White (1991) and by Bower *et al.* (in preparation) shows that alternative biasing schemes in the cold dark matter model can give rise to large-scale power in the galaxy distribution at the level detected in the QDOT survey. In these biasing models, the transformation between the density and galaxy fields is not purely local and this may preclude a simple relation between their respective topological properties. Alternative diagnostics, such as the structure of the microwave background radiation or the abundance of galaxy clusters, may help to decide whether the discrepancy we have uncovered reflects a shortcoming of the density field in the standard cold dark matter model or an inadequate prescription for the way in which the galaxies are assumed to trace this field.

A central aim of this work was to answer the question of whether or not the observed distribution of *IRAS* galaxies could have resulted from the gravitational growth of an initially Gaussian field of density fluctuations, as assumed, for example, in the standard cold dark matter model. Smooth maps of the galaxy distribution exhibit the topology characteristic of Gaussian density fields. High- and low-density regions have a remarkably similar structure. Voids are not isolated empty spaces but rather regions of low density which interconnect in a sponge-like fashion. Clusters are also joined together by walls and bridges of galaxies that form extended overdense regions. We found that the genus-threshold density relation developed by Hamilton *et al.* (1986) and Gott *et al.* (1987) provides a useful quantitative measure of the topology. For our analysis, we relied on mock 'QDOT surveys' constructed from cold dark matter N-body simulations and a variety of Monte-Carlo techniques to keep a tight grip on random and systematic uncertainties. Due to the great depth of the QDOT survey, we were able to probe a range of length-scales, 10–50 h^{-1} Mpc, not probed by previous galaxy surveys. Over this range, the topology of the QDOT survey agrees remarkably well with that expected for a distribution which has evolved from an initially Gaussian field of density fluctuations. As a counter example, we showed that our data rule out the non-Gaussian distribution in a Voronoi foam model of large-scale structure with 100 h^{-1} Mpc cells at about the 5σ level. It should be noted that our results do not conflict with the marginal detection of skewness in the count distribution of QDOT galaxies reported by Saunders *et al.* (1991). In fact, given the measured variance, this level of skewness is about what one

would expect from gravitational evolution from Gaussian initial conditions (Coles & Frenk 1991).

To summarize, although *IRAS* galaxies are somewhat less strongly clustered than optical galaxies, the QDOT survey contains a wealth of information on large-scale structure. The survey is useful out to at least 200 h^{-1} Mpc and many well-known clusters, superclusters and voids can be identified within this distance. The most prominent feature within 75 h^{-1} Mpc is the Hydra–Centaurus complex and the far field is dominated by the Hercules supercluster, extending between 100 and 150 h^{-1} Mpc. Our topological analysis has led to two main results: the detection of a 'sponge' topology characteristic of a Gaussian density field and a measurement of the slope of the power spectrum on large scales.

ACKNOWLEDGMENTS

BM acknowledges the support of an SERC/NATO fellowship. CSF acknowledges a Nuffield Foundation Science Research Fellowship. DW acknowledges support from the US National Science Foundation through a NATO postdoctoral fellowship and grant AST90-20506.

REFERENCES

Abell, G., 1958. *Astrophys.. J. Suppl.,* **3,** 211.
Adler, R. J., 1981. *The geometry of Gaussian random fields,* Wiley, Chichester.
Babul, A. & Postman, M., 1989. *Astrophys. J.,* **359,** 280.
Babul, A. & White, S. D. M., 1991. *Mon. Not. R. astr. Soc.,* **253,** 31P
Bardeen, J. M., Steinhardt, P. & Turner, M., 1983. *Phys. Rev. D.,* **28,** 679.
Bardeen, J. M., Bond, J. R., Kaiser, N. & Szalay, A. S., 1986. *Astrophys. J.,* **304,** 15.
Broadhurst, T. J., Ellis, R. S., Koo, D. C. & Szalay, A. S., 1990. *Nature,* **343,** 726.
Burstein, D., Faber, S. & Dressler, A., 1990. *Astrophys. J.,* **354,** 18.
Coles, P., 1990. *Nature,* **346,** 446.
Coles, P. & Frenk, C. S., 1991. *Mon. Not. R. astr. Soc.,* **253,** 727.
Davis, M. & Peebles, P. J. E., 1983. *Astrophys. J.,* **267,** 465.
Davis, M., Huchra, J., Latham, D. W. & Tonry, J., 1982. *Astrophys. J.,* **253,** 423.
Davis, M., Efstathiou, G., Frenk, C. S. & White, S. D. M., 1985. *Astrophys. J.,* **292,** 465.
Doroshkevich, A. G., 1970. *Astrophysics,* **6,** 320 (translated from *Astrofizika,* **6,** 581).
Efstathiou, G., Frenk, C. S., White, S D. M. & Davis, M., 1988. *Mon. Not. R. astr. Soc.,* **235,** 715.
Efstathiou, G., Kaiser, N., Saunders, W., Lawrence, A., Rowan-Robinson, M., Ellis, R. S. & Frenk, C. S., 1990. *Mon. Not. R. astr. Soc.,* **24,** 10P.
Frenk, C. S., 1991. *Physica Scripta,* **T36,** 70.
Frenk, C. S., White, S. D. M., Efstathiou, G. & Davis, M., 1990. *Astrophys. J.,* **391,** 2.
Geller, M. J. & Huchra, J. P., 1989. *Science,* **246,** 897.
Giovanelli, R. & Haynes, M. P., 1985. *Astr. J.,* **90,** 2445 (GH).
Gott, J. R., Melott, A. L. & Dickinson, M., 1986. *Astrophys. J.,* **306,** 341 (GMD).
Gott, J. R., Weinberg, D. H. & Melott, A. L., 1987. *Astrophys. J.,* **391,** 1.
Gott, J. R. *et al.,* 1989. *Astrophys. J.,* **340,** 625.
Hamilton, A. J. S., Gott, J. R. & Weinberg, D. H., 1986. *Astrophys. J.,* **309,** 1.
Huchra, J. P., Davis, M., Latham, D. & Tonry, J., 1983. *Astrophys. J. Suppl.,* **52,** 89.
Joseph, R. D. & Wright, G. S., 1985. *Mon. Not. R. astr. Soc.,* **214,** 87.
Kaiser, N., 1986. *Mon. Not. R. astr. Soc.,* **219,** 795.
Kaiser, N., Efstathiou, G., Ellis, R. S., Frenk, C. S., Lawrence, A., Rowan-Robinson, M. & Saunders, W., 1991. *Mon. Not. R. astr. Soc.,* **252,** 1.
Kirshner, R. P., Oemler, A., Schechter, P. L. & Schectman, S. A., 1981. *Astrophys. J. Lett.,* **248,** L57.
Lucey, J. R., 1983. *Mon. Not. R. astr. Soc.,* **204,** 33.
Melott, A. L., Weinberg, D. H. & Gott, J. R., 1988. *Astrophys. J.,* **328,** 50.
Olive, K. A., Schramm, D. N., Steigman, G. & Walker, T. P., 1990. *Phys. Lett. B,* **236,** 434.
Park, C. & Gott, J. R., 1991. *Astrophys. J.,* **378,** 457.
Peebles, P. J. E., 1980. *The Large Scale Structure of the Universe,* Princeton University Press, Princeton, NJ.
Rowan-Robinson, M. *et al.,* 1990. *Mon. Not. R. astr. Soc.,* **247,** 1.
Sanders, D. B., Soifer, B. T., Elias, J. H., Madore, B. F., Mathrews, K., Neugebauer, G. & Scoville, N. Z., 1988. *Astrophys. J.,* **325,** 74.
Saunders, W., Rowan-Robinson, M., Lawrence, A., Efstathiou, G., Kaiser, N., Ellis, R. S. & Frenk, C. S., 1990. *Mon. Not. R. astr. Soc.,* **242,** 318.
Saunders, W., Frenk, C. S., Rowan-Robinson, M., Efstathiou, G., Lawrence, A., Kaiser, N., Ellis, R. S., Crawford, J., Xiao-Yang Xia & Parry, I., 1991. *Nature,* **349,** 32.
Schneider, S. E., Thuan, T. X., Magri, C. & Wadiak, J. E., 1990. *Astrophys. J. Suppl.,* **72,** 245.
Shane, C. D. & Wirtanen, C. A., 1967. *Publ. Lick Obs.,* **22,** 1.
Shapley, H. & Ames, A., 1932. *A survey of the external galaxies Brighter than the Thirteenth Magnitude, Annals Harvard Coll. Obs.,* **88,** No. 2.
Strauss, M. A., Davis, M., Yahil, A. & Huchra, P. J., 1990. *Astrophys. J.,* **361,** 49.
Tully, R. B., 1987. *Astrophys. J.,* **323,** 1.
van de Weygaert, R., 1991. *PhD thesis,* University of Leiden.
van de Weygaert, R. & Icke, V., 1989. *Astr. Astrophys.,* **213,** 1.
Weinberg, D. H., 1988. *Publs astr. Soc. Pacif.,* **100,** 1373.
Weinberg, D. H., Gott, J. R. & Melott, A. L., 1987. *Astrophys. J.,* **321,** 2 (WGM).
White, S. D. M., Frenk, C. S., Davis, M. & Efstathiou, G., 1987. *Astrophys. J.,* **313,** 505.
Zwicky, F., Herzog, E., Wild, P., Karpowicz, M. & Kowai, C. T., 1961–1968. *Catalogue of Galaxies and Clusters of Galaxies,* California Institute of Technology, Pasadena.