Mon. Not. R. astr. Soc. (1985) **217,** 805–820

The statistics of maxima in primordial density perturbations

J. A. Peacock *Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*

A. F. Heavens *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*

Accepted 1985 July 18. Received 1985 July 2; in original form 1985 May 2

Summary. We have investigated the hypothesis that protogalaxies/protoclusters form at the sites of maxima in a primordial field of normally distributed density perturbations. Using a mixture of analytic and numerical techniques, we have studied the properties of the maxima, with the following results:

(i) Maxima are surprisingly overdense, with a fairly narrow distribution of overdensities. If the rms variation in $\delta(\equiv\delta\varrho/\varrho)$ is σ, then the mean value of δ for maxima is $\geqslant 2\sigma$ with a spread $\sim 0.7\sigma$, unless the perturbation spectrum is rather flat. This corresponds to a relatively small range of collapse times.

(ii) The shapes of the maxima depend on their heights. For asymptotically high maxima, all principal axes become equal. The typical maximum, however, is more nearly prolate with axial ratios 1 : 0.70 : 0.55, again quasi-independent of spectrum.

(iii) The clustering properties of maxima can differ from those of the density field as a whole. Maxima need not form good tracers of mass.

These results provide a natural mechanism for biased galaxy formation in which galaxies do not necessarily follow the large-scale density. Methods for obtaining the true autocorrelation function of the density field and implications for Microwave Background studies are discussed.

1 Introduction

In the gravitational instability picture for galaxy formation, structure in the Universe arises from the existence at early times of a spectrum of small density perturbations. A considerable literature now exists on deducing the present-day consequences of a given form for these initial irregularities. The most problematical step of this process is making the translation from the density field to the distribution of galaxies. The assumption that the galaxies trace the mass is widely employed in e.g. N-body simulations, where the initial distributions of test particles follow the large-scale density field. In this paper, we consider one reason why this assumption may be invalid: in situations where there exist density fluctuations on galaxy scales, bound structures will

eventually form at *maxima* of the density field. The properties of these maxima (overdensity, mass, clustering, etc.) are determined simply by the form of the initial irregularities; we investigate these for a variety of possible spectra. Furthermore, even in cases where there are no small-scale perturbations (the adiabatic or large-scale damping picture), a study of the properties of large-scale density maxima (identified as protoclusters or protosuperclusters) is of great interest. In particular, in addition to properties such as the height of the maxima (peak value of $\delta\varrho/\varrho$, which determines the collapse time) we can determine the shape, i.e. degree of triaxiality of the peaks. This is an important initial condition which determines the degree of elongation of present-day superclusters.

The need for such a study has been implicitly recognized for some time (e.g. Peebles 1980, p. 124), but it has not been pursued in Western literature. Our investigation is more similar in spirit to the work of Doroshkevitch & Shandarin (1978). The subject of density maxima has recently attracted interest through the idea of biased galaxy formation, where the chance of finding an object is supposed to be non-linearly enhanced in regions of high density, either through uncertain physics (e.g. Rees 1985; Davis *et al.* 1985) or as a result of our observational selection (Kaiser 1984). Our approach is distinct from these authors: maxima turn out to be already biased in the sense of having overdensities high by comparison with the 'typical' rms level. Nevertheless, maxima do not show enhanced clustering unless a further threshold is applied. The relation of this work to biased galaxy formation in general is discussed in Section 4.

The plan of the paper is as follows. Section 2 discusses the distribution of densities and masses of maxima; Section 3 considers shapes of maxima and the dependence of these on other properties; Section 4 investigates the clustering of maxima.

2 Scalar properties of maxima

2.1 HEIGHTS OF MAXIMA

We begin with some definitions. The fractional density perturbation $\delta(\equiv\delta\varrho/\varrho)$ is specified as a Fourier series within some large volume V_u, and the series approximated by an integral:

$$\delta(\mathbf{x})=\frac{V_u}{(2\pi)^3}\int\delta_{\mathbf{k}}(\mathbf{k})\exp(-i\mathbf{k}\cdot\mathbf{x})\,d^3k,$$

$$\delta_{\mathbf{k}}(\mathbf{k})=\frac{1}{V_u}\int\delta(\mathbf{x})\exp(i\mathbf{k}\cdot\mathbf{x})\,d^3x.$$

The usual assumption is that the phases of the δ_k are random – a hypothesis of maximum ignorance. The central limit theorem then requires that δ is a random Gaussian process. (This assumption may eventually be proved false – see Peebles 1983). In this case, the statistics of δ are specified by the power spectrum, which is often considered to be a power law $|\delta_k|^2\propto k^n$, where $k\equiv|\mathbf{k}|$.

The simplest property of a maximum in such a noise field is its height, $\delta_{\max}$. We can calculate the distribution of heights by extending the one-dimensional calculation of Rice (1954). Consider the joint probability density function of δ and its first two derivatives δ_i' and δ_{ij}'': for Gaussian noise, this will be a multivariate normal distribution in 10 variables (considering six independent δ_{ij}'' only), which we place in a vector $\mathbf{V}$:

$$f(\mathbf{V})=\frac{1}{(2\pi)^5}\frac{1}{|\mathbf{M}|^{1/2}}\exp(-\tfrac{1}{2}\tilde{\mathbf{V}}\cdot\mathbf{M}^{-1}\cdot\mathbf{V})$$

where $\mathbf{M}$ is the covariance matrix $M_{ij}=\langle V_iV_j\rangle$ (see Appendix). Following Rice, the number

density of maxima with heights in the range δ, $\delta+d\delta$ is $\eta(\delta)\,d\delta$, where

$$\eta(\delta)=\int f(\mathbf{V}\,|\,\delta_i'=0)\,|\mathbf{H}|\,d^6\delta''$$

where $\mathbf{H}$ is the Hessian matrix $H_{ij}=\delta''_{ij}$, so $|\mathbf{H}|$ is simply the product of the three principal curvature components. This expression arises by considering the probability of obtaining the derivatives δ_i' in the ranges $d\delta_i'$ in the volume $\Pi_i dx_i$; this is simply $\Pi_i d\delta_i'$ times the six-dimensional integral of f over δ''_{ij}. Now, near a maximum, $d\delta_i'=\delta''_{ij}\,dx_j$ and we can replace the δ_i' integration by one over x_i by means of the Jacobian $|\mathbf{H}|$. Finally, division by $\Pi_i dx_i$ gives the number density of maxima, $\eta(\delta)$. The evaluation of this integral is complicated by the regime of integration; for maxima, all three principal curvatures must be negative. Fortunately, it is straightforward to carry out the six-dimensional integration numerically. In δ'' space, the correlation of δ and δ''_{ij} means that $f(\mathbf{V}\,|\,\delta_i'=0)$ peaks at values of principal curvatures proportional to δ. The spread about the peak remains constant, so in the limit $\delta\rightarrow\infty$ the maxima will become round (all curvatures equal) and $\eta(\delta)$ will have the asymptotic form

$$\eta(\delta)\propto\delta^3\exp(-\delta^2/2\sigma^2)$$

where σ is the rms fractional density variation (see Appendix), *cf.* Doroshkevich & Shandarin (1978).

Fig. 1 shows the results of this procedure for several power spectra. We consider $|\delta_{\mathbf{k}}|^2=ak^n$ with sharp cut-offs at some $k_{\min}$ and $k_{\max}$, corresponding to wavelengths $\lambda_{\max}$ and $\lambda_{\min}$. For most values of n, the results depend only on $\lambda_{\min}$ (if $\lambda_{\max}\gg\lambda_{\min}$); for $n\leqslant-3$, however, the location of the upper cut-off becomes important (see below). This is a crude approximation to the situation in reality, where the power spectrum is damped at small wavelengths according to some transfer function, which in general has a complicated form. Simple approximations for the transfer function include $\exp(-k/k_D)$ for standard adiabatic fluctuations (Doroshkevich & Shandarin 1978) and $\mathrm{dexp}\{-(k/k_\nu)^{1.5}\}$ for large-scale damping with massive neutrinos (White, Frenk & Davis 1983). In practice, it is more convenient to characterize the spectrum by a single wavenumber. We have chosen to define k_c as the wavenumber at which $k^3|\delta_{\mathbf{k}}|^2$ (the power per unit $\log k$ for an isotropic spectrum) is a maximum. In this case, the true spectrum can be

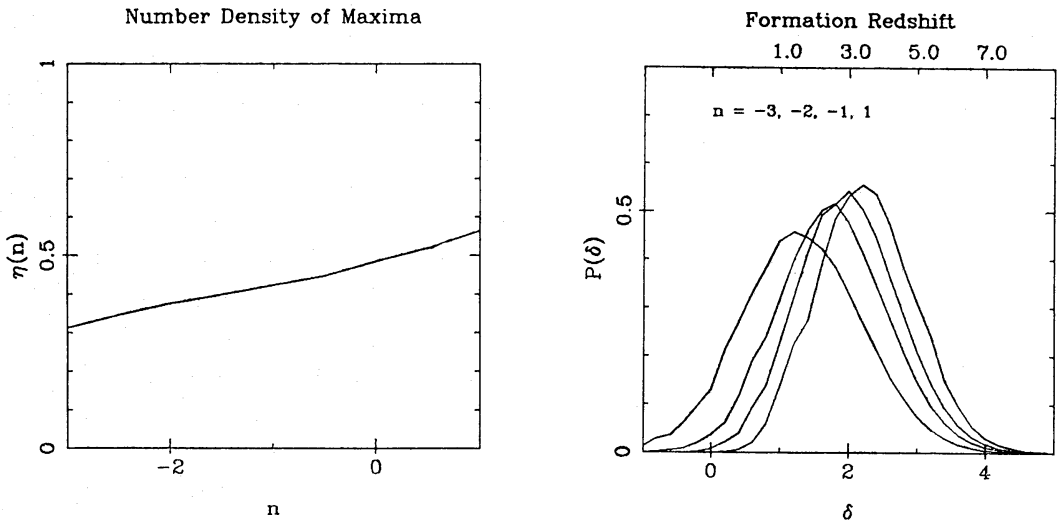

Figure 1. (a) The total number density of maxima for power-law spectra $|\delta_{\mathbf{k}}|^2\propto k^n$, $\lambda>\lambda_{\min}$ (density normalized to a volume $\lambda^3_{\min}$), as a function of n. (b) The probability distributions of overdensities of maxima, for varying n. The units of δ are the small-scale rms variation in $\delta\varrho/\varrho$, σ. Note that for all but the flattest spectra ($n\simeq-3$) the peak in $f(\delta)$ is at $\delta\simeq2\sigma$. The upper scale labels collapse redshift on the assumption that the 2σ point goes non-linear at $z=3$.

modelled by a power law truncated at k_{max}. where $k_{max}=f(n)k_c$ and $f(n)$ is rather insensitive to the precise form of the transfer function. This model cannot be exact, as we need to fit two moments of $|\delta_k|^2$ (see Appendix), but in most cases these are both given correctly to within ~10 per cent by $f(n)=2.3\,(n+3)^{-0.34}$ (except for $n\leqslant -2$). This procedure breaks down for the case of a Universe dominated by non-baryonic cold dark matter, where the power spectrum is not well approximated by a single power-law. There is a change in slope of $\Delta n=4$ near scales corresponding to the horizon at the onset of matter domination, but this happens quite gradually, and the spectrum is noticeably curved over scales $\sim 10^{10}$–$10^{16}\,M_\odot$ (Blumenthal *et al.* 1984). This power spectrum presents special problems for the approach we are taking here. The case usually considered is a Zel'dovich spectrum, which breaks to $n=-3$ at small scales; this spectrum is sufficiently powerful on large scales that it is less correct to consider small-scale maxima as isolated distinct entities – small maxima exist within large-scale ones of comparable strength. Further, there is no damping cut-off at short wavelengths in this model: if we consider a CDM spectrum truncated at λ_{min}, this will correspond to observational smoothing of the density field by some window function. A sensible choice for the window to be relevant to galaxies now would be $\lambda_{min}\sim$10–100 kpc. For the various CDM spectra given by Bond & Efstathiou (1984), this requires $\lambda_{max}/\lambda_{min}\sim$10–100 to model the spectrum as a doubly truncated $n=-3$ power law. As an illustrative case, we have considered $\lambda_{max}/\lambda_{min}=25$ (or a range of $\sim 10^4$ in mass). We plot in Fig. 1 the total number density of maxima (normalized to $\lambda_{min}=1$) and the probability distributions for the heights separately. There are several important points here:

(i) For $n\geqslant -1$, there are effectively no maxima with $\delta<0$. This corresponds to spectra which are dominated by power on small scales.

(ii) The distributions generally peak at a value of δ considerably greater than σ, the rms variation in δ for the density field as a whole.

(iii) The spread in peak values is quite small by comparison to the mean, typically $\sim 0.7\sigma$.

In particular, we may quantify these points by giving a simple analytic approximation for the height distribution:

$$\eta(\delta)\propto(\delta+\alpha)^3\exp\frac{-(\delta-\beta)^2}{2\sigma^2}.$$

Good fits are obtained with $\alpha=0, \beta=0.8\sigma$ for $n=1$, and $\alpha=2\sigma, \beta=0.9\sigma$ for $n=-3$. Note that the peak in the distribution occurs at a somewhat larger value than β.

These results have important implications for the calculation of collapse times for perturbations. If we take this to occur when linear theory predicts δ to reach a value $\delta_c\sim 1$ (1.68 on the linear spherical model: Kaiser 1984), then with δ growing as $(1+z)^{-1}$, it makes a large difference to the collapse time depending on whether we require 2σ (a typical maximum) or σ (a typical fluctuation) to grow to δ_c. Furthermore, apart from altering the mean collapse time, the above picture implies a small *range* of collapse times. For example, if we take the median δ for $n=1$ to collapse at $z=3$, then half the maxima will collapse in the redshift range $z=2$–4. This contrasts most strongly with previous work which considered the relevant $\eta(\delta)$ to be simply the positive half of the Gaussian obtained by averaging the primordial density field on some length-scale of interest (see e.g. Efstathiou, Fall & Hogan 1979; Schaeffer & Silk 1985). In this latter case, the spread in δ is ~the mean value; if we require the median collapse time to correspond to $z=3$, then 13 per cent of perturbations would remain uncollapsed at the present epoch, even for $\Omega=1$. In our model with $n=1$, essentially all maxima will have collapsed by the present epoch, provided the median collapse redshift is $\geqslant 2.5$.

Fig. 1 also shows a scale of formation redshift; this is defined by the epoch at which linear theory predicts that δ reaches some critical value δ_c. The scale is normalized so the formation redshift for

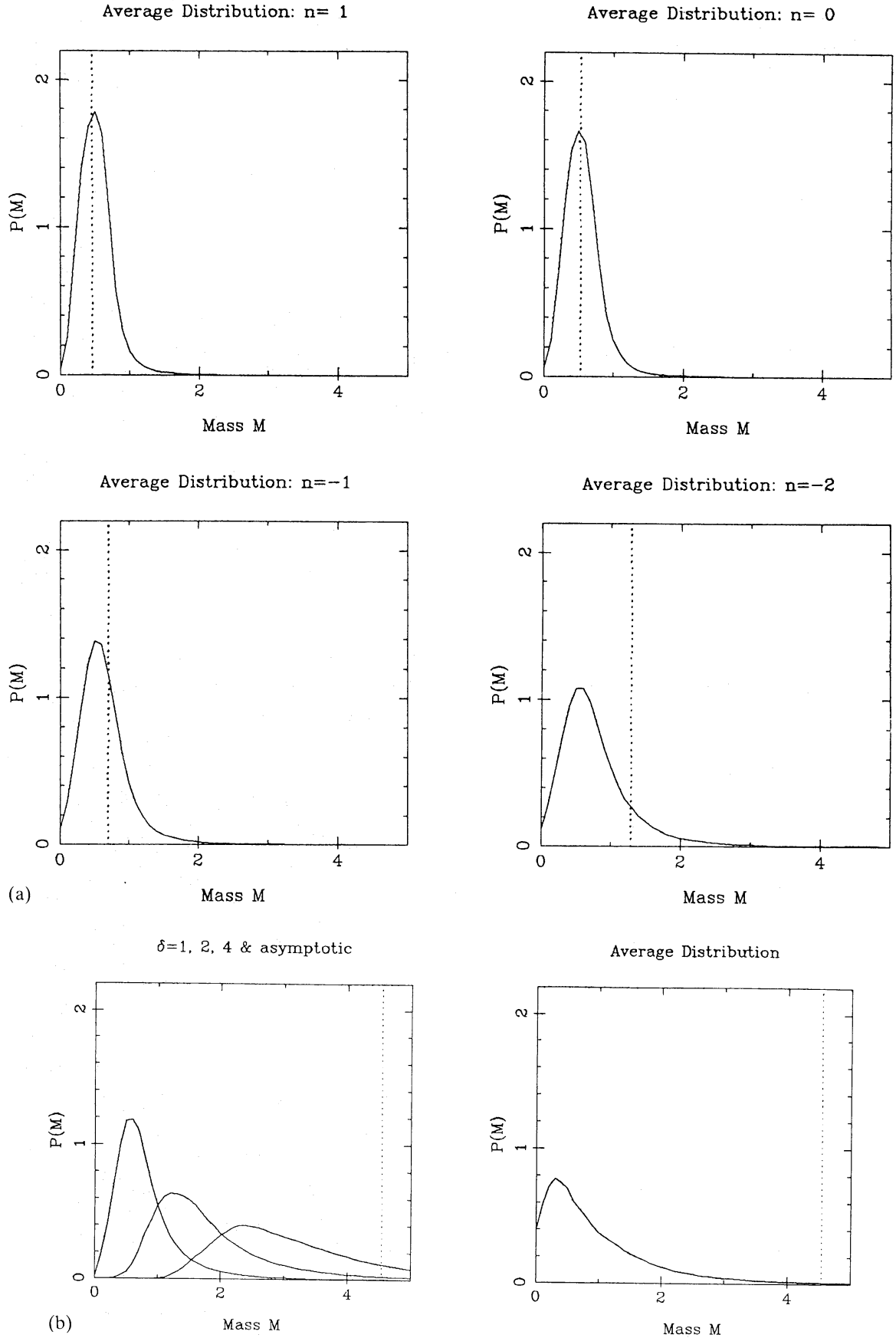

Figure 2. The probability distributions of the mass parameter for maxima, for a variety of spectra. (a) The full lines are the probability distributions for all maxima. The dotted lines are the asymptotic limit as $\delta \to \infty$. (b) The probability distributions for masses for maxima of different height, for an approximate 'cold dark matter' spectrum with $n=-3$. Note that higher peaks generally have greater mass parameters.

a 2σ peak (near the median for most spectra) is 3. It is straightforward to normalize to other 2σ formation redshifts $\hat{z}_f$ by

$$1+\hat{z}_f=\left(\frac{1+\hat{z}_f}{4}\right)\{1+\hat{z}_f\,(\text{graph})\}.$$

It is interesting to note that if $\hat{z}_f=3$ and $n=1$ a 'typical fluctuation' with $\delta=\sigma$ will not collapse until $z=0.8$, and yet 10 per cent of the objects form before $z=4.6$

2.2 MASSES OF MAXIMA

After collapse, the single most useful parameter of an object will be its mass. It is not completely straightforward to obtain the final mass of a primordial maximum as this will be modified by infall of surrounding material during the non-linear stages of collapse. Nevertheless, it is useful to consider a mass parameter for the maxima which may correlate with the final mass. Near a maximum in δ, we have

$$\delta(\mathbf{x})\simeq\delta_m-\tfrac{1}{2}\delta''_{ij}\,dx_i\,dx_j$$

or, diagonalizing the Hessian matrix δ''_{ij}, we obtain the principal axes a, b, c:

$$\delta(\mathbf{x})\simeq\delta_m\left\{1-\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\right)\right\}.$$

Defining the boundary as $\delta=0$, the volume of the maximum is $V=(4\pi/3)\,abc$ and the mass is therefore ϱV (as δ is very small initially). In Appendix 1, it is shown that $\delta''_{ij}\propto\delta_m$ for large δ_m. Thus, the scale factors, a, b, c should be approximately independent of δ_m. Indeed, asymptotically, the mass distribution tends to a delta function. At a fixed δ_m, it is straightforward to carry out the analogous integration to Section 2.1 and obtain the mass distribution. This is shown in Fig. 2, which gives for the same spectra as Fig. 1, both the asymptotic $f(M)$ ($\delta\to\infty$) and the average $f(M)$ (all maxima). The units of mass are $\varrho\lambda_{\min}^3$ where $\lambda_{\min}$ is the cut-off wavelength. Again, for steep spectra, these distributions are similar [confirming the approximate independence of $f(M)$ and δ] and quasi-Gaussian. There is in fact a slight tendency for the high maxima to be more massive, but only for flat spectra. Note that for the highest maxima of practical relevance ($\delta_m\simeq4\sigma$) the mass distributions are still very different from the asymptotic prediction.

3 Shapes of maxima

We saw in the previous section that very high density maxima ($\delta\gg\sigma$) tended towards spherical symmetry. Such fluctuations are extremely rare. It is of more interest to ask about the shapes of more typical ($\delta\simeq2\sigma$) maxima. It is possible in principle to extend the above techniques to find the joint probability density function of the three principal axes $f(a,b,c)$ via further numerical integration. In practice however, it is both more straightforward and a good deal faster in CPU terms to use a Monte Carlo simulation instead. This has the added advantage of providing a check on our 'analytic' results.

Such a simulation proceeds by filling a three-dimensional grid with a summation of many waves of specified amplitude and random phase. The only way of obtaining a reasonable number of grid points is to use a three-dimensional Fast Fourier Transform. In practice, this allowed us to generate an array of size 100^3 sampling rather fewer waves in ~1000 s on a VAX 11/780. Grid maxima (those points higher than their 26 nearest neighbours) were located. The grid maxima do

not, of course, necessarily correspond to the 'true' maxima of the density field. In order to estimate the properties of the true maxima, the 27 grid points were used to estimate δ_i' and δ_{ij}''. If the density field is well-sampled, these can be used to estimate the position and height δ_m of the true maximum. The Hessian matrix was then diagonalized to find the principal axes of the true maximum. It was found that sampling of four grid points per minimum wavelength was the limit for reliable results. One run of the simulation produced a sample of $\sim 10^4$ maxima. We found good agreement between the distributions of δ_m and mass produced by the simulation and the calculations in the previous section.

The simplest way of displaying the shape information is to plot the various axes directly. Fig. 3 gives plots of short axis against middle axis, normalized to long axis = 1. The major point to note is

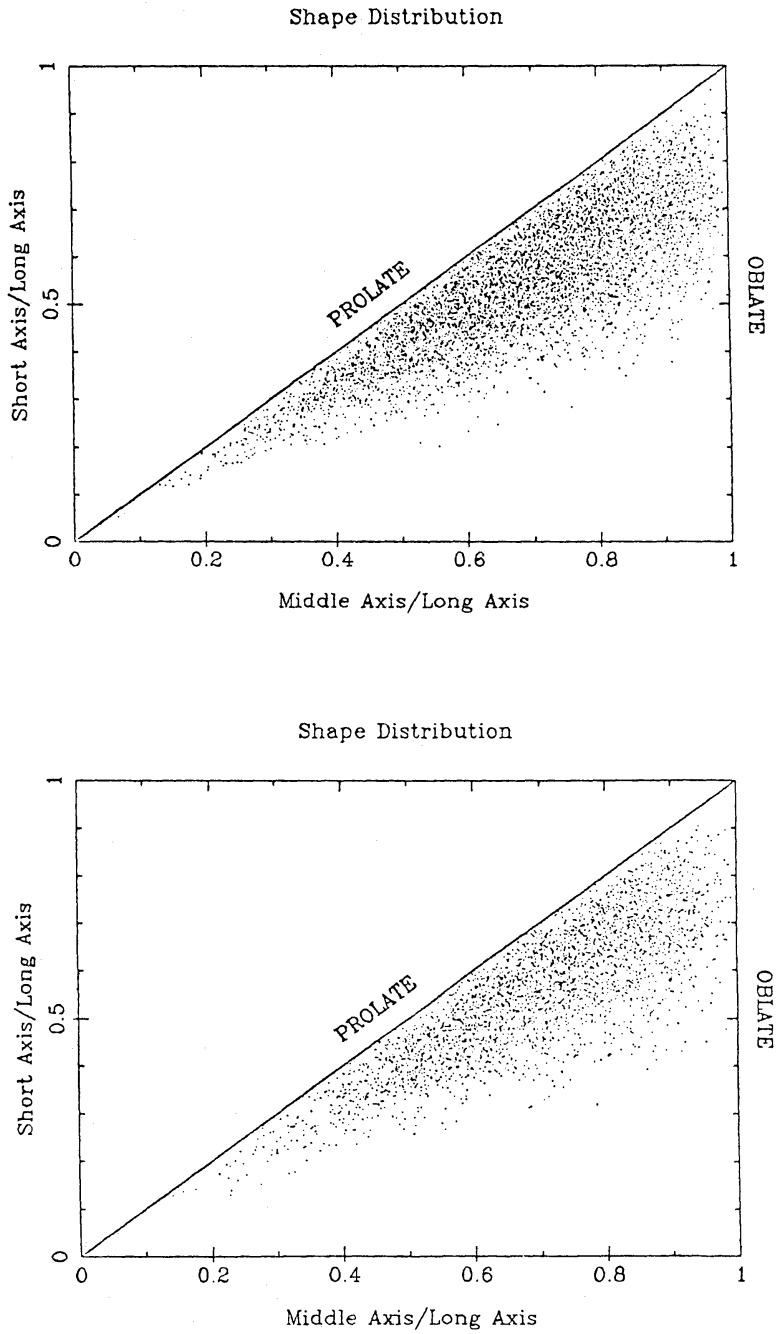

Figure 3. A plot of short axis against middle axis (normalized to long axis = 1) for two of the Monte Carlo simulations. Each maximum is represented by one point on the plot. The top diagram is $n = 1$; the bottom $n = -3$. Note the tendency of points to lie near the prolate line.

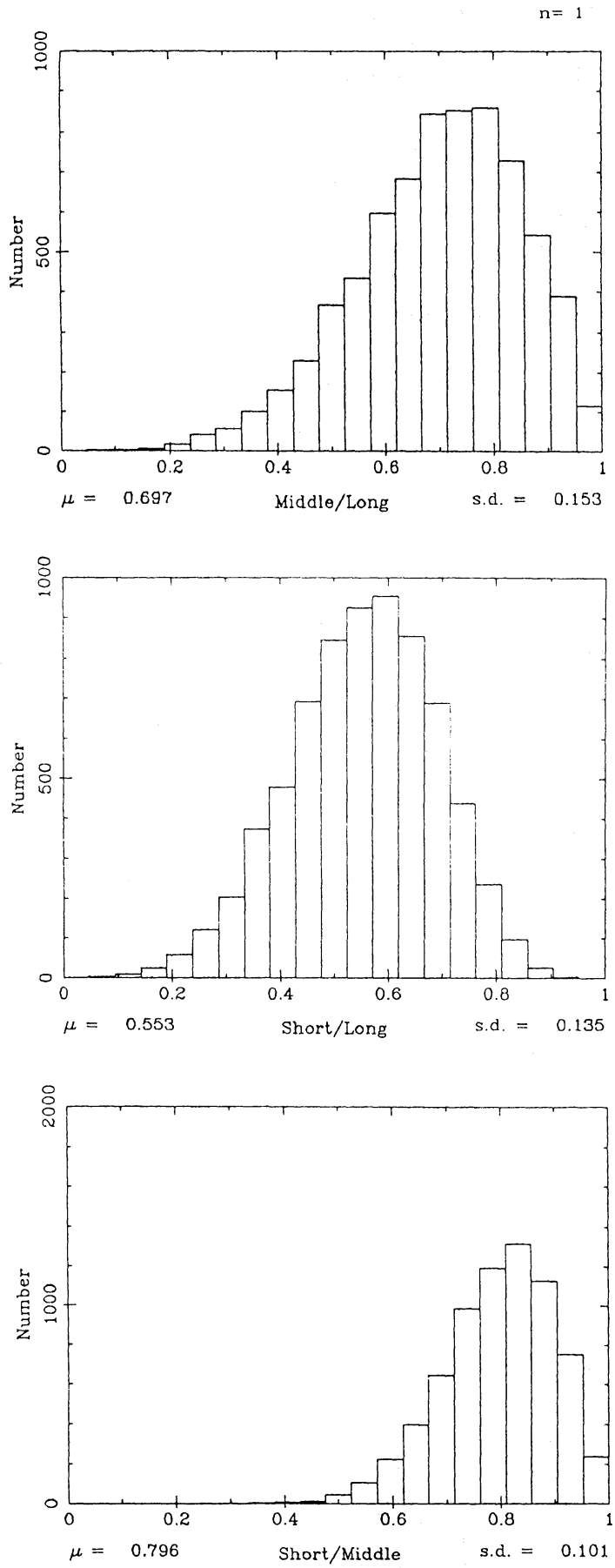

Figure 4. Histograms of axial ratios for the $n=1$ data shown in Fig. 3. Note the relative equality of short and middle axes.

that the predicted roundness of perturbations is reached only slowly. In particular, for maxima with $\delta_m \sim 3\sigma$, the mean axial ratios (for $n=1$) are $b/a=0.7$ and $c/a=0.6$. Higher maxima are rounder, but they are very rare – typical maxima are markedly aspherical. They are typically slightly more prolate than $1:x:x^2$. Fig. 4 illustrates this point in another way by plotting histograms of axial ratios. For $n=1$, the median values are $b/a=0.7$, $c/a=0.55$. Again, the dependence on spectrum is not extreme.

The significance of these results depends on the precise scheme of galaxy formation being considered. We have focused on two cases: power-law spectra with $n=1$ and $n=-3$, with sharp cut-offs at long and short wavelengths. (As noted before, smooth cut-offs or breaks in the spectrum do not change the results significantly, provided $\lambda_{\rm break} \gg \lambda_{\rm min}$.) The natural application of the $n=1$ case is to the adiabatic or large-scale damping model of galaxy formation. In this case, the maxima should be identified as protosuperclusters and minima as protovoids. The shape information is important because it determines the final degree of flattening of the supercluster – i.e. whether it has collapsed along one axis to become a 'pancake' or along two to become a 'cigar' or 'filament'. These issues are considered by More, Heavens & Peacock (1986, in preparation); the essential point is that the triaxial nature of the maxima makes eventual collapse along two axes seem more likely to occur.

The interpretation of the $n=-3$ case is in the context of the cold dark matter model with small-scale damping. We are then effectively smoothing the power spectrum on scales corresponding to a galaxy, and then locating maxima. This is a less clear-cut procedure than in the large-scale damping model.

In the small-scale damping model, the presence of a considerable amount of large-scale power then influences the clustering properties of the galaxies, which we address next.

4 Clustering of maxima

4.1 CORRELATION FUNCTIONS

We now wish to ask how the spatial distribution of maxima is related to the overall density field. In particular, what information about the large-scale non-uniformity of the Universe can be extracted from the clumping of tracer particles viewed as density maxima? There are two relevant measures of the clustering; these are the autocorrelation function $W(\mathbf{r})$ of the fractional density perturbation field $\delta(\mathbf{r})$

$$W(\mathbf{r})=\langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle$$

and the two-point correlation function $\xi(\mathbf{r})$ of the set of tracer particles considered. This is defined as the excess probability of finding a particle a distance r from another, relative to the probability for an unclustered distribution dP_0

$$dP=\{1+\xi(\mathbf{r})\}\, dP_0.$$

If we consider a set of particles scattered with a density proportional to matter density, then $\xi(r)=W(r)$, except on very small scales (Peebles 1980, p. 147). Thus, subject to the common assumption that galaxies trace the mass distribution, we could deduce the autocorrelation function of $\delta(\mathbf{r})$ from $\xi(r)$ for galaxies. However, this assumption may not hold in the case we are considering. In linear theory with $\Omega=1$, δ increases with time, so that $W(r) \propto (1+z)^{-2}$ (if r is in comoving units). However, the positions of the maxima are independent of this growth and thus $\xi(r)$ for them does not change with time. We can illustrate this difference by evaluating directly the two-point correlation function for maxima in our simulations and comparing this with the analytic $W(r)$.

$$W(r)=\frac{V_u}{(2\pi)^3}\int |\delta_{\mathbf{k}}|^2 \exp(i\mathbf{k}\cdot\mathbf{r})\, d^3k.$$

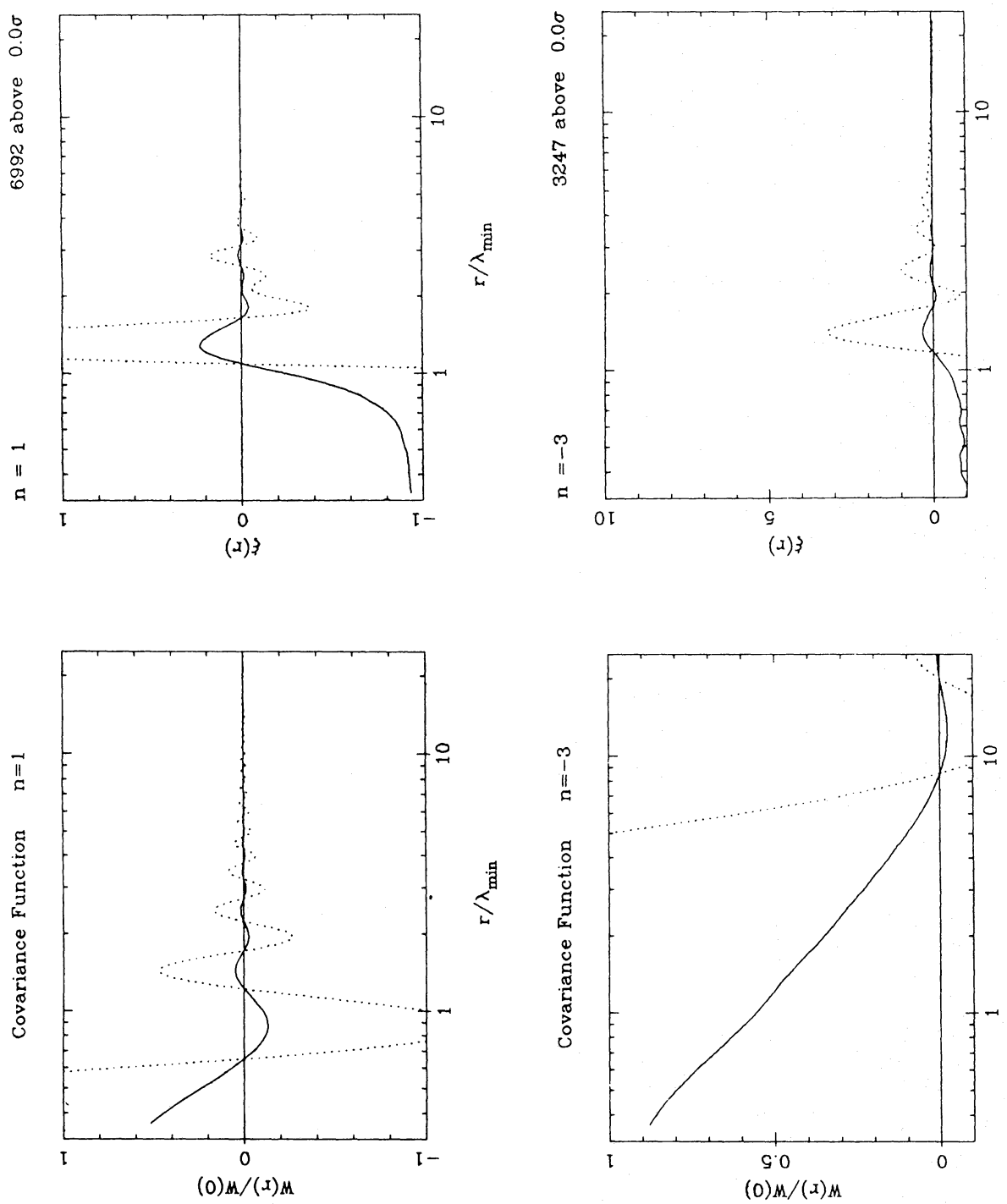

Figure 5. Comparison of the mass autocorrelation function $W(r)$ [normalized to $W(0)=1$] with the 2-point correlation functions for maxima $\xi(r)$ for various spectra. The dotted lines show the correlation functions multiplied by a factor of 10 to illustrate the behaviour at large r.

Fig. 5 makes this comparison for a range of spectra; note that for convenience $W(r)$ is normalized to $W(0)=1$. It is probably not meaningful to consider $W(0)\leqslant 1$ as then the maxima will not have undergone collapse and will not be identifiable as distinct objects. For $W(0)\gg 1$, conversely, the above argument does not apply. This is because, going beyond linear theory, the positions of small-scale maxima are affected by the large-scale density perturbations. In the Zel'dovich approximation (e.g. Efstathiou & Silk 1983, section 5.2), fluid displacements due to short- and long-wavelength disturbances grow independently and small-scale maxima will thus be approximately convected with any large-scale flow. This means that our correlation functions for maxima are not strictly accurate on scales $\gg\lambda_{\min}$, but for $W(0)\sim 1$ the clustering on these scales is in any case small. For $n\geqslant -3$ these effect of convection are negligible. Once the local peaks have collapsed, the resulting bound objects should continue to act as Lagrangian markers and trace the large-scale flow.

A further complication is presented by the mass which is not initially incorporated into one of the first generation of bound objects. If the efficiency with which this material is incorporated into a neighbouring proto-object is a function of position, then the galaxies will not trace the mass in that the large-scale luminosity density need not follow the mass distribution. To give an idea of the scale of the problem, we can see how much of the mass of the Universe is accounted for by the mass estimates in Section 2.2. This generally accounts for only about 25 per cent of the total. Now, if infall was completely effective in assigning mass to one or other bound objects there would be an interesting consequence: given a constant M/L ratio, the maxima would again trace the mass in that the large-scale luminosity density would vary with the density field. However, in general the efficiency of infall might well be expected to vary. In regions of large-scale overdensity, there is both a higher density of background matter and a greater number density of potential accreting objects. This could lead not to a further increase in the number density, but to an increase in the *luminosity* of objects in these areas. In this case $\xi(r)$ from a magnitude-limited sample would not yield a correct estimate of the large-scale density contrast because it implicitly assumes that the luminosity function is constant in form throughout space. It is not straightforward to say for certain whether this must happen. However, it should be interesting to investigate any differences between normal and luminosity-weighted correlation functions for existing redshift surveys.

It is interesting to compare these results with those of Kaiser (1984). He computed the cross-correlation function of regions lying at thresholds $\delta>\nu\sigma$ in Gaussian noise. For linear clustering in the limit $\nu\gg 1$, Kaiser found

$$W_{>\nu}(r)=\left(\frac{\nu^2}{\sigma^2}\right)W(r).$$

This amplification occurs essentially because high fluctuations are more likely to occur in regions of greater large-scale density. Kaiser estimated the amplification for Abell clusters, assuming that $W(r)$ could be replaced by $\xi(r)$ for clusters and galaxies respectively, and showed that a threshold $\nu=2$–2.5 could account for the observed discrepancy between the galaxy–galaxy and cluster–cluster correlation functions. This process is related to a general idea now receiving much attention, called biased galaxy formation (see e.g. Rees 1984, 1985). The essential idea is that the efficiency of galaxy formation could be some non-linear function of density, leading to galaxies being found preferentially in regions of high threshold in δ. The major motivation for this idea is to allow galaxies to be more strongly clumped than the overall matter distribution, so that a density parameter of $\Omega=1$ would be allowed despite the smaller values apparently implied by the low peculiar velocities of galaxies (e.g. Efstathiou & Silk 1983). In this context, it is worth noting that the statistics of maxima provide a natural mechanism for biasing in that very few $<1\sigma$ peaks are found (Section 2.1); this is achieved without appealing to any non-linearity. However, this does not immediately yield Kaiser's result on clustering strengths. This is because he considers the

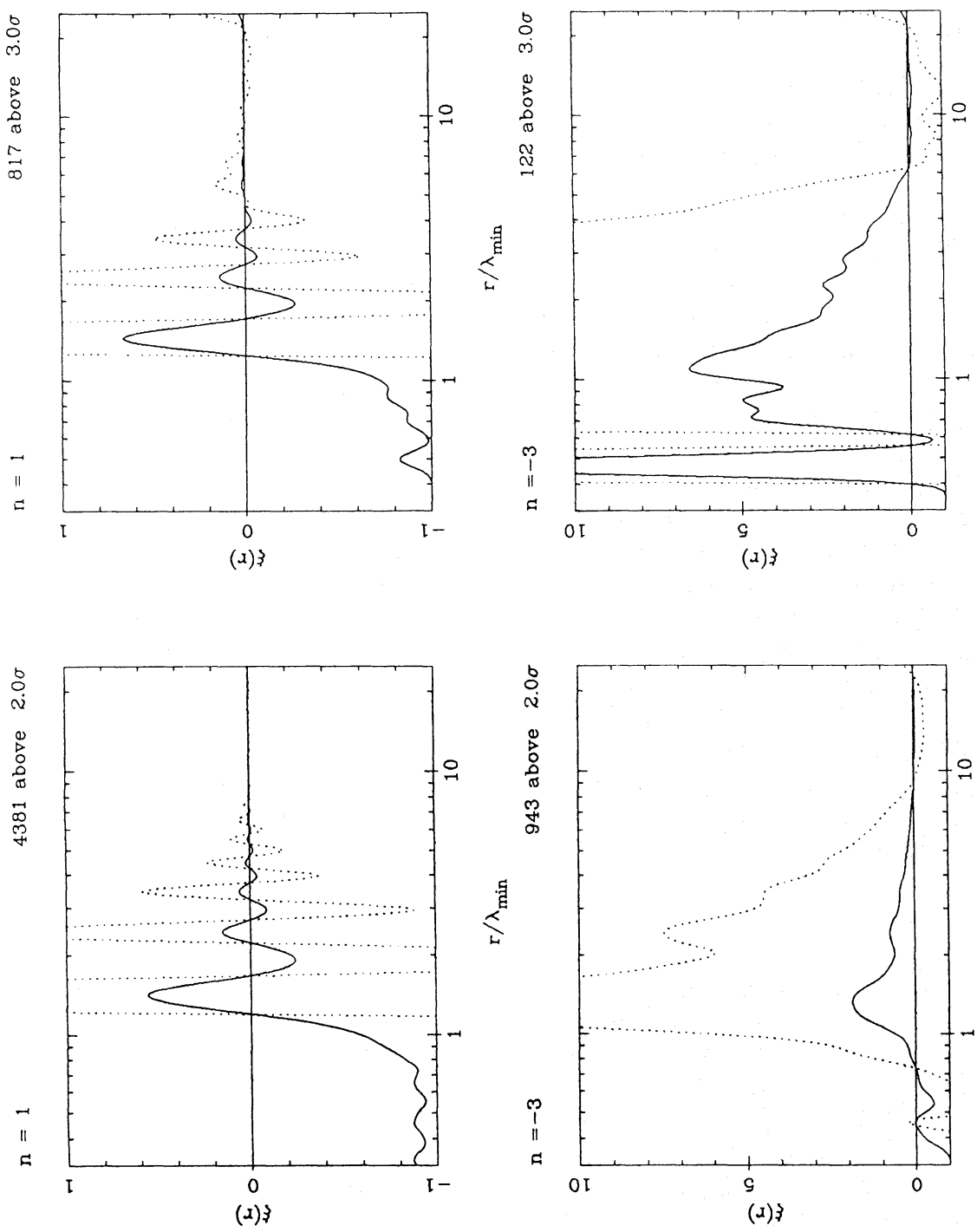

Figure 6. Thresholded 2-point correlation functions [i.e. $\xi(r)$ for high maxima only]. Thresholds are set at $\delta_{max}=\nu\sigma$ and the function found numerically from the Monte Carlo simulation. As essentially no peaks have $\delta_{max}>4\sigma$, only the values $\nu=2$ and 3 are considered ($\nu=0$ is shown in Fig. 5). For $n=-3$, the $\nu=3$ case has $\xi(r)$ amplified by approximately an order of magnitude. Conversely, for $n=1$, thresholding has little effect.

cross-correlation function of regions of high threshold; such regions are found preferentially in areas of large-scale enhanced density and thus have a larger coherence length. However, this does not imply that there are vastly fewer maxima in large-scale voids – only that they are on average lower and will thus collapse later. Since Section 2.1 showed that there is a narrow range of collapse times for maxima, there is only a small range of time in which we will find significantly fewer bound structures in voids. Thus, we would not expect enhanced clustering of maxima as specified only by their two-point correlation function unless luminosity-biasing effects as discussed in Section 4.1 were important [see also Rees (1984, 1985) for mechanisms whereby the first maxima to collapse could prevent others forming galaxies]. The Kaiser effect operates to some extent if we now impose a threshold on maxima – i.e. considering abnormally high peaks only. However, it does this in a way which depends on the degree of large-scale power in the spectrum (not surprisingly). This is illustrated in Fig. 6: for flat spectra ($n=-3$) there is a large amplification, with $\xi(r)$ for $>3\sigma$ peaks increased by about an order of magnitude; conversely for $n=1$ there is no change in ξ with threshold, as there is little large-scale power. In any case, any amplication is with respect to $\xi(r)$ for maxima as a whole, which is not necessarily related to the large-scale $W(r)$ (see above). Kaiser associates 3σ peaks with Abell clusters, and compares their correlation function with that of galaxies instead of that of the matter. The only circumstances in which Kaiser's assumption of equating $W(r)$ with $\xi(r)$ for galaxies might seem to be valid is if we consider large-scale damping with infall going to completion: in this case the numbers of galaxies forming in clusters and superclusters could scale with overall mass. However, since clusters would then no longer correspond to peaks in the primordial density field, Kaiser's analysis of $\xi(r)$ for the cluster does not apply, as he points out.

From the point of view of estimating $W(r)$ it is interesting to note that the main effect of thresholding appears to be on the amplitude of $\xi(r)$: the positions of zeros in $\xi(r)$ are approximately preserved. Moreover, comparison of Figs 5(b) and 6(b) indicates that, on scales greater than the inter-maximum separation, $W(r)$ and $\xi(r)$ for maxima above a threshold $\nu\sigma$ agree approximately to within a multiplicative factor (this overestimates ξ slightly):

$$\xi(r)\simeq\nu^2\frac{W(r)}{W(0)},$$

as found by Kaiser (1984) and Politzer & Wise (1984) for a different but related problem. This means that, if we have a set of objects which we know to have been selected above some threshold $\nu\sigma$, then $\xi(r)$ for these objects may give us an estimate of the *form* of $W(r)$. The relative scaling can only be determined from the threshold ν and the epoch of formation, however. For $n=-3$, we know that, when a typical (1.5σ) maximum collapses ($\delta=1$ in linear theory) then $W(0)=(1.5)^{-2}$ at the redshift z_f when this occurs. If the large-scale correlation function is not a result of the formation of anisotropic structures (Heavens 1985), then on scales where the density contrast is still small, the correlation function of the maxima above $\nu\sigma$ should be related to the autocorrelation function of the matter by

$$\xi_{>\nu\sigma}(r, z=0)\simeq(1.5)^2\nu^2W(r, z=z_f).$$

The question is to which objects this relation should be applied. Galaxies appear to be a poor choice for two reasons. First, we need to consider scales where the large-scale structure is still linear, and $\xi(r)$ for galaxies become small and uncertain. Secondly, we do not know if galaxies are in fact selected above any fixed threshold. Conversely, if we consider Abell clusters, the signature of clustering is strong out to $\sim$100 Mpc. Moreover, there is a prospect of determining the correct threshold, because this is set by our observational selection, rather than by any of the poorly understood mechanisms for biased galaxy formation. If Abell clusters are viewed as maxima in $\delta(\mathbf{r})$ when smoothed through a window of size R_W, then a knowledge of the number density of

clusters determines the threshold ν. Kaiser (1984) shows that R_W in the range (20–10) h_{50}^{-1} Mpc yields $\nu=2$–3. Thus, there seems some prospect of regarding the clustering of Abell clusters not as biased by comparison with the 'true' $\xi(r)$ for galaxies, but as the best available diagnostic (when corrected as above) of the large-scale inhomogeneity of mass in the Universe.

4.2 IMPLICATIONS FOR MICROWAVE BACKGROUND ANISOTROPIES

Any possible difference between $\xi(r)$ and $W(r)$ is especially important in calculations of the expected anisotropies in the Microwave Background. These depend on inferring the present-day $\delta(\mathbf{r})$ from the galaxy distribution and then extrapolating back to the time of recombination. This can be done in essentially two ways. Wilson & Silk (1981) and Vittorio & Silk (1984) normalize by requiring $W(r)$ for the matter distribution now to equal $\xi(r)$ for galaxies, at the point where $\xi(r)=1$ $\{10(H_0/50)^{-1}$ Mpc, where H_0 is in km s^{-1} Mpc$^{-1}\}$. This procedure is also followed by Bond & Efstathiou (1984) for the case of small-scale damping. For large-scale damping (e.g. a Universe dominated by massive neutrinos), however, they instead normalize to a pancake collapse epoch when $W(0)=1$, requiring that this happen at a redshift >3.

In the latter case, the effect of our results is easy to see; applying Bond & Efstathiou's criterion for collapse, the correct collapse epoch is when δ for a typical maximum reaches 1, i.e. $2.2\sigma=1$ rather than $\sigma=1$. Thus, the predicted $\Delta T/T$ values of Bond & Efstathiou are reduced by a straightforward factor of 2.2. Their $\Omega=1$ massive neutrino prediction already lies a factor ~2 below the best observational upper limit (Uson & Wilkinson 1984), so it would seem possible to reduce Ω significantly without running into difficulties with the background anisotropy.

The impact on the cases of purely baryonic or cold non-baryonic dark matter is less easy to quantify. On the one hand we have shown that a modest degree of biasing is natural (a factor of 2.2 for $n=1$, 1.5 for the effective $n=-3$ cold dark matter case), which would again tend to reduce $\Delta T/T$. On the other, we have demonstrated that $\xi(r)$ for galaxies in such a case does not describe the large-scale density field correctly. Perhaps the safest procedure would be a normalization analogous to the one for large-scale damping – i.e. requiring the typical maxima to reach $\delta=1$ at some redshift which is observationally reasonable for galaxy formation, $z\geqslant2$ say. At present, it is not clear whether the assumption that $W(r)=1$ at $r=10\,h_{50}^{-1}$ Mpc today need be correct, and we believe that constraints on Ω etc. from such calculations should be treated with caution.

5 Conclusions

We have calculated the simplest consequences of assuming that galaxies form from Gaussian primordial density fluctuations, given that individual bound objects will form at the sites of density maxima. Our principal conclusions are as follows:

(i) Maxima generally have a fairly small dispersion in values of $\delta\varrho/\varrho$. One can define an 'epoch of galaxy formation' $\hat{z}_f$ from the median value of $\delta\varrho/\varrho$, but there is a spread (typically $\Delta z\sim1$ if $\hat{z}_f\sim3$) in formation redshifts.

(ii) The values of $\delta\varrho/\varrho$ at a maximum are generally $\gtrsim2\sigma$ (where σ is the rms variation in $\delta\varrho/\varrho$), somewhat dependent on the spectrum. This provides a natural mechanism for a modest degree of 'biasing', without appealing to non-linearities in the efficiency of galaxy formation as a function of density.

(iii) Typical maxima are triaxial, with a tendency to be slightly more prolate than $1:x:x^2$. The asymptotic tendency of high maxima to become spherical does not become important until $>4\sigma$ fluctuations are considered. For maxima of supercluster scale, subsequent evolution (More *et al.* 1986, in preparation) suggests that this may be related to the observed tendency towards filamentary structure in the Universe.

(iv) The two-point correlation function for maxima is generally not the same as the autocorrelation function of the matter distribution on any scale – maxima do not trace mass. This point is enhanced if the ~75 per cent of mass not initially contained in a collapsing proto-object does not subsequently undergo infall.

(v) The above points imply that the expected fluctuations in the Microwave Background have sometimes been overestimated; Universes with low values of Ω may still be allowed.

Acknowledgments

We are grateful to Martin Rees, George Efstathiou and Nick Kaiser for illuminating discussions on this topic.

References

Blumenthal, G. R., Faber, S. M., Primack, J. R. & Rees, M. J., 1984. *Nature,* **311,** 517.
Bond, J. R. & Efstathiou, G., 1984. *Astrophys. J.,* **285,** L45.
Davis, M., Efstathiou, G., Frenk, C. S. & White, S. D. M., 1985. *Astrophys. J.,* **292,** 371.
Doroshkevich, A. G. & Shandarin, S. F., 1978. *Mon. Not. R. astr. Soc.,* **182,** 27.
Efstathiou, G. Fall, S. M. & Hogan, C., 1979. *Mon. Not. R. astr. Soc.,* **189,** 203.
Efstathiou, G. & Silk, J., 1983. *Fund. Cos. Phys.,* **9,** 1.
Heavens, A. F., 1985. *Mon. Not. R. astr. Soc.,* **213,** 143.
Kaiser, N., 1984. *Astrophys. J.,* **284,** L9.
Peebles, P. J. E., 1980. *The Large-Scale Structure of the Universe,* Princeton.
Peebles, P. J. E., 1983. *Astrophys. J.,* **274,** 1.
Politzer, H. D. & Wise, M. B., 1984. *Astrophys. J.,* **285,** L1.
Rees, M., 1984. *J. Astrophys. Astr.,* **5,** 331.
Rees, M., 1985. *Mon. Not. R. astr. Soc.,* **213,** 75P.
Rice, S. O., 1954. *Selected papers on Noise and Stochastic Processes,* p. 133, ed. Wax, N., Dover.
Schaeffer, R. & Silk, J., 1985. *Astrophys. J.,* **292,** 319.
Uson, J. M. & Wilkinson, D. T., 1984. *Astrophys. J.,* **277,** L1.
Vittorio, N. & Silk, J., 1984. *Astrophys. J.,* **285,** L39.
White, S. D. M., Frenk, C. S. & Davis, M., 1983. *Astrophys. J.,* **274,** L1.
Wilson, M. L. & Silk, J., 1981. *Astrophys. J.,* **243,** 14.

Appendix: Evaluation of $\rho(\delta)$

We first need to consider the covariance matrix $M_{ij}=\langle V_i V_j\rangle$ where $\mathbf{V}=(\delta, \delta_i', \delta_{11}'', \delta_{22}'', \delta_{33}'', \delta_{12}'', \delta_{13}'', \delta_{23}'')$. The non-zero components of this may be written as follows:

$M_{11}=\sigma_0^2$ (referred to as simply σ^2 in the text)

$M_{22}=M_{33}=M_{44}=\sigma_1^2/3$

$M_{55}=M_{66}=M_{77}=\sigma_2^2/5$

$M_{15}=M_{16}=M_{17}=-\sigma_1^2/3$

$M_{88}=M_{99}=M_{10\,10}=M_{56}$, etc. $=\sigma_2^2/15$

where

$$\sigma_m^2=\frac{4\pi V_u}{(2\pi)^3}\int_0^\infty |\delta_{\mathbf{k}}|^2 k^{2m+2}\,dk.$$

The multivariate Gaussian distribution $f(\mathbf{V})$ contains the quadratic form $\tilde{\mathbf{V}}\cdot\mathbf{M}^{-1}\cdot\mathbf{V}$. To calculate this, we must invert the 10×10 matrix $\mathbf{M}$, which is a reasonably straightforward exercise. In order

to investigate the asymptotic limits $\delta \gg 1$ we re-order the quadratic form, which contains terms proportional to δ^2, $x_i\delta$, and x_ix_j, where $\mathbf{x}$ is a six-component vector containing the six independent second derivatives (i.e. V_5–V_{10}). We write

$$\tilde{\mathbf{V}} \cdot \mathbf{M}^{-1} \cdot \mathbf{V} = Q + (\tilde{\mathbf{x}} - \tilde{\boldsymbol{\mu}}) \cdot \mathbf{N} \cdot (\mathbf{x} - \boldsymbol{\mu})$$

where $Q \propto \delta^2$ and the 'means' $\mu_i \propto \delta$. Q turns out to be simply δ^2/σ_0^2, and μ_i either $-(\sigma_1^2/3\sigma_0^2)\delta$ (for δ''_{11}, δ''_{22}, δ''_{33}) or zero (δ''_{12} etc.). The matrix $\mathbf{N}$ is simply the part of $\mathbf{M}^{-1}$ which mixes second derivatives in the original quadratic form. The meaning of this new quadratic form becomes clear if we fix δ and set all the $x_i = \mu_i$ except one, x_j. As we vary x_j, the distribution peaks at a value μ_j, and is normally distributed with a standard deviation given by components of $\mathbf{N}$. As δ increases, the means of δ''_{11}, δ''_{22}, δ''_{33} become more negative, in proportion to δ, but the spread of values about the mean remains constant; it is fixed by $\mathbf{N}$, which is independent of δ. So, for asymptotically high maxima ($\delta \gg \sigma_0$), it is a good approximation to set $\delta''_{11} = \delta''_{22} = \delta''_{33} = -(\sigma_1^2/3\sigma_0^2)\delta$ (i.e. principal axes become equal), and the integration to find $\eta(\delta)$ can be done analytically. The result is

$$\eta(\delta) = \frac{1}{(2\pi)^2} \left(\frac{\sigma_1^2}{3\sigma_0^2}\right)^{3/2} \delta^3 \exp(-\delta^2/2\sigma_0^2) \qquad \delta \gg \sigma_0.$$

This is similar to the expression quoted by Doroshkevich & Shandarin (1978), except that their constant of proportionality is independent of spectrum. The above authors give no derivation of their result, but we believe it to be incorrect as the above form agrees both with numerical integration and simulation.