Mon. Not. R. Astron. Soc. **282**, 631–640 (1996)

Merger trees and the multiplicity function of haloes

D. D. C. Rodrigues and P. A. Thomas
Astronomy Centre, Dept. of Physics and Astronomy, University of Sussex, Brighton BN1 9QH

Accepted 1996 April 29. Received 1996 March 8; in original form 1995 November 6

ABSTRACT
We present a new method for calculating the merger history of matter haloes in hierarchical clustering cosmologies. The linear density field is smoothed on a range of scales, these are then ordered in terms of decreasing density and a merger tree constructed. The method is similar in many respects to the block model of Cole & Kaiser but has a number of advantages: (i) it retains information about the spatial correlations between haloes, (ii) it uses a series of overlapping grids and is thereby much better at finding rare, high-mass haloes, (iii) it is not limited to haloes whose mass ratios are powers of two, and (iv) it is based on an actual realization of the density field and so can be tested against N-body simulations. The major disadvantages are (i) the minimum halo mass is eight times the unit cell with a corresponding loss of dynamic range, and (ii) occasionally the relative location of haloes in the tree does not reflect the correct ordering of their collapse times, as computed from the mean halo density. We show that our model exhibits the required scaling behaviour when tested on power-law spectra of density perturbations, but that it mimics peaks theory in predicting more massive haloes for flat spectra than does the Press–Schechter formalism.

Key words: galaxies: formation – galaxies: luminosity function, mass function.

1 INTRODUCTION

A basic tenet of modern cosmology is the idea that the present large-scale structure of the Universe originated by the gravitational growth of small matter inhomogeneities. These initial density fluctuations are thought to be imprinted in a universe dominated by collisionless dark matter at very high redshifts. Their distribution of amplitudes with spatial scale depends ultimately both on the nature of this collisionless matter and on the physical processes operating prior to the epoch of recombination. A family of these generic models comprises the moderately successful hierarchical cosmogonies, which suppose that the variance of initial fluctuations decreases with scale. This means that small structures are the first to collapse and that galaxies, groups and clusters are formed by the merging of non-linear objects into larger and larger units. This merging sequence can be visualized as a hierarchical tree with the thickness of its branches reflecting the mass ratio of the objects involved in the merging (Lacey & Cole 1993). If we imagine time running from the top of the tree, the main trunk would represent the final object, while its past merging history would be represented schematically by the ramification of this trunk into small branches, representing accretion of small sub-lumps, and by the splitting into branches of comparable thickness when merging of sub-clumps of comparable size occurs.

The linear growth of the density field is well-understood, but collapsed objects, or 'dark haloes', are highly non-linear gravitational structures whose dynamical evolution is difficult to trace. Some progress can be made by the direct numerical integration of the equations of motion in N-body simulations, but these are limited in dynamic range and are very time-consuming. Theoretical models are usually based on the analytic, top-hat model of Gunn & Gott (1972). Spherical overdensities in a critical-density universe reach a maximum size when their linear overdensity reaches 1.06, then recollapse and virialize at an overdensity of approximately $\delta_c = 1.69$. Unfortunately real haloes are neither uniform nor spherically symmetric so that their collapse times scatter about the predicted value.

In cosmology we are seldom interested in the specific nature of one individual halo, but rather in the statistical properties of the whole population. The analytical approach to this problem was pioneered by Press & Schechter (1974; hereafter PS). To estimate what proportion of the universe is contained in structures of mass M at redshift z, the density field is first smoothed with a top-hat filter of radius R, where $M = 4/3\pi\bar{\rho}R^3$ and $\bar{\rho}$ is the mean density of the universe. $F(M,z)$ is then defined to be the fractional volume where the smoothed density exceeds δ_c. Assuming a Gaussian density field, then

$$F(M,z) = \frac{1}{2}\mathrm{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma(M,z)}\right), \quad (1)$$

where σ is the root-mean-square fluctuation within the top-

hat filter and erfc is the complementary error function. The key step was to realize that fluctuations on different mass-scales are not independent. In fact, to a first approximation PS assumed that high-mass haloes were entirely made up of lower-mass ones with no underdense matter mixed in. Then F must be regarded as a cumulative mass fraction and it can be differentiated to obtain the differential one,

$$f(M,z) = -\frac{\partial F}{\partial M} = -\frac{1}{\sqrt{2\pi}}\frac{\delta_c}{\sigma^2}\frac{\partial \sigma}{\partial M}e^{-\delta_c^2/2\sigma^2}. \quad (2)$$

The main drawback of this approach is that, because of the above assumption of crowding together of low-mass haloes into larger ones, it seems to undercount the number of objects. As $M \mapsto 0$ (and therefore $\sigma \mapsto \infty$) the fraction of the Universe that exceeds the density threshold tends to one half. For this reason Press & Schechter multiplied f by two to reflect the fact that most of the Universe today is contained in collapsed structures. It is this higher normalization which is referred to below as the PS prediction. In Section 3.2 we will refer to the lower normalization given in equation (2) as 'uncorrected PS'.

Extensions of the PS prescription, to calculate explicitly the integrated merger history of haloes, were developed by Bower (1991) and Bond et al. (1991), and tested against N-body experiments by Lacey & Cole (1993, 1994). In the 'excursion set' formalism introduced by Bond et al. (1991), the top-hat smoothing radius about a given point is first set to a very large value and then gradually reduced until the enclosed overdensity exceeds δ_c (in hierarchical cosmologies this will always occur before the radius shrinks to zero). This gives the largest region that will have collapsed around that point. In their paper this is illustrated graphically in fig. 4 where it becomes evident that an underdense region on a certain mass-scale might not be so on larger mass-scales: this is the reason why the uncorrected PS formulae under-estimate the total mass in collapse objects. They show that the higher normalization PS prediction corresponds to the special case of a 'top-hat' filter in k-space. More importantly their diagram suggests that by varying the density threshold one can build up a picture of the collapse and merger history of the haloes: our paper describes a numerical representation of this process.

Despite being very idealized in nature, ignoring both the internal structure and tidal forces, the PS formula provides a surprisingly good fit to the N-body results (Efstathiou et al. 1988; Lacey & Cole 1994). However, we have to regard these successes with some scepticism, since the basic hypothesis of the excursion set theory works very poorly on a object-by-object basis (White 1995), the numerical simulations are still plagued by resolution effects and limited dynamical range, and the halo statistics are sensitive to the scheme chosen for identifying haloes. Moreover, one should always bear in mind that the PS treatment is a linear approach to a problem which is fundamentally non-linear in nature.

The full non-linear evolution of structure is best described by an N-body simulation. Moreover, with the introduction of techniques such as smoothed particle hydrodynamics (SPH), it is possible simultaneously to follow the evolution of a dissipative, continuous intergalactic medium. However, there are several drawbacks to this approach: N-body simulations are very time-consuming, they have a limited dynamical range and they are very inflexible when trying to model the physical processes happening on small scales (with small numbers of particles). For example, it is likely that the interstellar medium in a protogalaxy will contain a mixture of hot and cold gas as well as stars with a variety of ages and dark matter. Simulations that can handle such situations are only just beginning to appear.

Thus it is highly desirable to set up a simple but efficient Monte Carlo procedure which mimics the general features of the hierarchical clustering process and can be used to carry out a large parameter investigation with little time-consumption. The first model to be presented along those lines was the 'block model' of Cole & Kaiser (1988), first used to study the abundance of clusters and subsequently some aspects of galaxy formation (Cole 1991; Cole et al. 1994; Heyl et al. 1995; Baugh, Cole & Frenk 1996). It starts with a large cuboidal block, with sides in the ratio $1 : 2^{1/3} : 2^{2/3}$, and subdivides it into two sub-blocks of the same shape. If the initial block has an overdensity δ (drawn from a Gaussian with variance $\sigma(M)$), then the two sub-blocks will inherit the same overdensity with an extra perturbation, added to one of them and subtracted from the other, drawn from a Gaussian with variance Σ, where $\Sigma^2 = \sigma^2(M) - \sigma^2(M/2)$. This quadratic procedure is applied iteratively to each of the sub-blocks until the imposed mass resolution is achieved. The advantage of the method comes from the fact that the relative position of all sub-blocks is known at all times so that it is simple to follow the merger history of any halo detected at any stage of the simulation.

Kauffmann & White (1993) adopt a different approach which makes use of the conditional merging probabilities derived by Bower (1991) and Bond et al. (1991). Given that a halo has a particular mass at some redshift, then one can work out the probability distribution for the mass of the halo (centred on the same point) at some earlier redshift. By generating a large number of representative haloes, say 100 or more, it is possible to allocate sub-haloes with the correct spectrum of masses. This method gives a better resolved mass spectrum (not restricted to powers of two) but restricts halo formation to occur at specific redshifts and is much more complicated to implement than the block model.

Here we present a new method for following halo evolution which is much closer in spirit to the N-body simulations without compromising the simplicity and speed of the above analytical techniques. It allows a continuous spectrum of halo masses (above a minimum of 8 unit cells) and a variable collapse time. We start with a full realization of the initial linear density field defined on a cubical lattice. (This constitutes part of the initial conditions for a cosmological simulation, which can therefore be used to test our method.) Secondly, we smooth the density field in cubical blocks on a range of scales, using for each scale of refinement a set of eight displaced grids. The blocks are then ordered in decreasing overdensity (i.e. increasing collapse time). We then run down this list, creating a merger tree for haloes. (The decision whether to merge two sub-haloes together into a larger one is crucial for preventing the growth of unphysically large structures.) As a bonus our technique retains spatial information about the relative location of haloes (i.e. a measure of their separation, not just the merging topology).

In the next section we describe our merger algorithm in more detail. Tests on simple power-law spectra of density perturbations are presented in Section 3, and the relative success, benefits and disadvantages of our method are contrasted with others in Section 4.

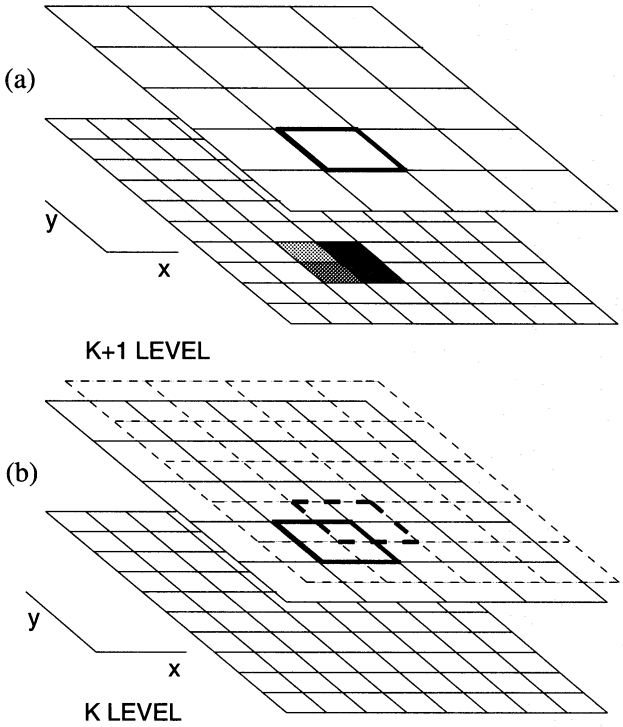

Figure 1. A two-dimensional representation of the blocking scheme. (a) Each block in the upper panel is constructed by averaging the four cells or blocks beneath it. (b) This picture shows two sets of overlapping grids, each of which aligns with the same sub-grid from the previous level of smoothing.

2 THE ALGORITHM

We begin with a realization of the chosen density field in a periodic cubical box of side $L \equiv 2^l$, where l is a positive integer. A standard initial condition generator is used which populates the box with waves of random phase and amplitude drawn from a Gaussian of mean zero and variance equal to the chosen input power spectrum. Neither the fact that L is a power of two, nor the periodic boundary conditions are strictly necessary, but they are used for simplicity.

Next we average the density fluctuations within cubical *blocks* of side 2, 4, ..., L. At each smoothing level we use eight sets of overlapping grids, displaced by half a block-length in each co-ordinate direction relative to one another (see Fig. 1). This ensures that density peaks will always be approximately centred within one of the blocks, and is a major advantage over other methods.

The density fluctuations within blocks and base-cells are now ordered in terms of decreasing density, giving the same order in which they would collapse as the universe ages (under the naïve assumption that they all have the same morphology at all times: we will test the accuracy of this assumption later).

The final step is to build up a merger tree to express the collapse history of blocks. This is a much harder problem than in the simple block model because the blocks that we use are not always nested inside one another but may overlap. Our initial guess was to merge together all collapsed blocks that overlap with one another, but this leads to very elongated structures which can stretch across a large fraction of the box. While these may represent large-scale pancakes or filaments, they are clearly not the kind of simple virialized haloes that we are trying to identify. In practice they would probably break up into smaller objects and so we need to find some way to limit their growth. The procedure we use to do this is as follows.

(1) First some terminology. Regions in the Lagrangian density field that lie above the threshold for collapse are known as *haloes*. Initially these coincide with the cubical blocks but they need not do so at later times once overlapping blocks begin to collapse. The merger tree consists of a list of cells and sub-haloes which constitute each halo. (For simplicity each cell, block or halo also contains a link to its 'parent' halo but these are not strictly required.)

(2) Initially, there are no collapsed haloes. We start at the top of the ordered list of cells and blocks and run down it in order of increasing collapse time.

(3) Each cell that collapses is given a parent halo, provided that it has not already been incorporated into some larger structure (this avoids the cloud-in-cloud problem).

(4) For each block that collapses we first obtain a list of all the haloes with which it overlaps and to what extent. The action to be taken depends upon this degree of overlap.

Any uncollapsed cells are added to the new halo. This represents accretion of intergalactic material.

If a halo is discovered whose mass is less than that of the block and at least half of which is contained within the block, then it is merged as part of the new structure. This would represent accretion of existing collapsed objects.

If the collapsing block has half or more of its mass contained in *exactly one* pre-existing halo then merge them together as part of the new structure. This would represent accretion of the block by a larger collapsed object. The restriction to exactly one pre-existing halo prevents the linking together of adjacent haloes without the collapse of any new matter (see Fig. 2a). It is this condition which prevents the growth of long filamentary structures.

Initially the method produces haloes of mass 1 and 8 cell units, but as blocks begin to merge they produce haloes of a wide variety of shapes and a continuous spectrum of masses. The two most common methods of sudden change in halo mass are creation by the merger of several sub-units (Fig. 2b) and accretion of a new block of approximately equal mass which overlaps with the halo (Fig. 2c). These produce approximately cubic structures, or triaxial with axial ratios ranging from 3:2 to 1:1 (see Fig. 7, later). By contrast, in the block model halo masses always increase by a factor of two at each merger event and the axial ratio is fixed.

3 RESULTS

3.1 Self-similarity

We have tested our algorithm on power-law density fluctuation spectra, which should give self-similar scaling on scales much smaller than the box-size. We take a power-law spectrum $P(k) \propto k^n$, where $n = -2$ or 0 to span the range of solutions expected in the real Universe. In an infinite box these would translate to a root-mean-square density fluctuation spectrum $\sigma(m) \propto m^{-\alpha}$ where $\alpha = (3+n)/6$. However, in practice we are missing a lot of power outside the box and so the decline is steeper than this at high masses, especially for $n = -2$. This is illustrated in Fig. 3(a) where the spectrum is clearly not a

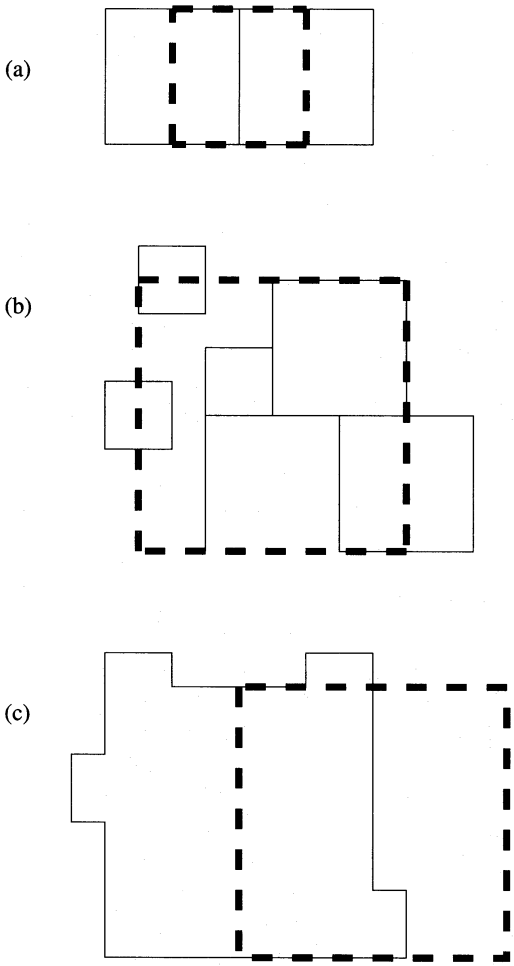

Figure 2. (a) We wish to avoid mergers such as that shown in this diagram where two pre-existing haloes (solid boxes) are linked together by the collapse of a third block. This is the reason for the restriction on merging discussed in the text. (b) A typical example of the formation of a new halo by merger of many smaller sub-units. (c) The growth of a halo by accretion of another block of almost equal size.

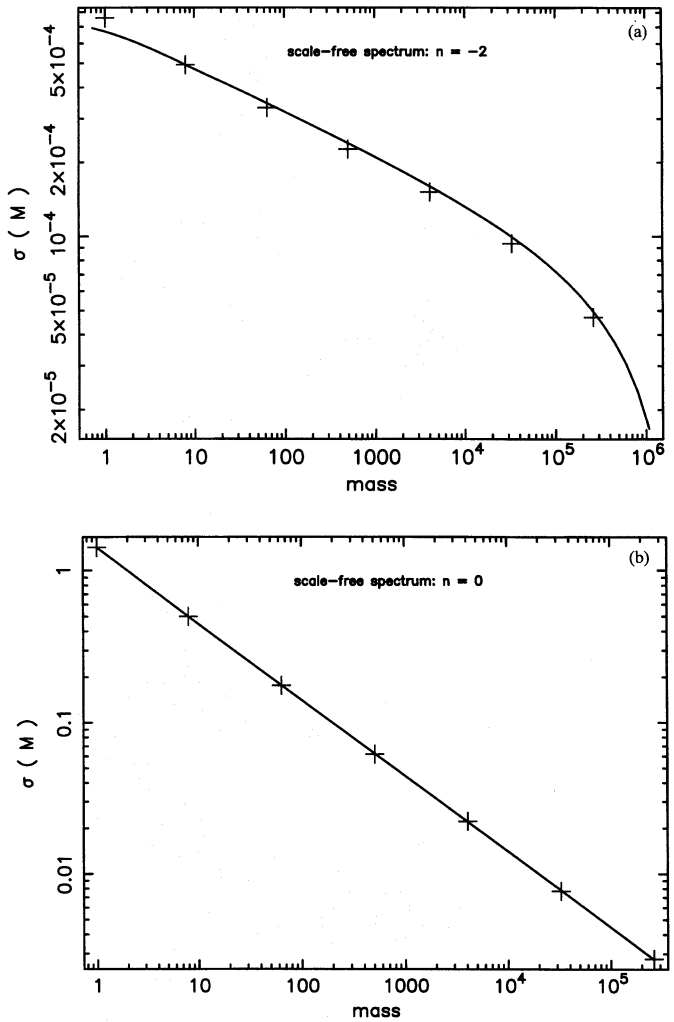

Figure 3. The measured root-mean-square mass fluctuation on various mass-scales for the 128^3 box: (a) $n = -2$, (b) $n = 0$. For $n = -2$ the solid line is calculated by direct summation of waves inside the box with a cubical window function. The corresponding curve is normalized to the second point of the data. The plots are on a logarithmic scale.

power law, but is well-fitted by the solid line which shows $\sigma(m)$ calculated by direct summation of waves inside the box with a window function associated with a cubical filter. For $n = 0$ the effect is not so severe, so we fit the data with the functional form of $\sigma(m)$ for an infinite box. Note that, because we are using a cubical filter, the normalization is different from what it would be for a spherical top-hat. This difference is irrelevant for the purposes of this paper because the normalization that we use is arbitrary; however, it could be important if we were to compare our predictions with the results of N-body simulations.

The results presented here were mostly obtained using boxes of side $L = 128$. We tried a range of box-sizes, from $L = 32$ to 256, to test the effect of variable resolution on our results. The code needs about $2L^3$ words of memory so $L = 256$ is the largest practical size on a workstation. If the merger tree is to be used as the basis of galaxy formation models, however, then much more storage is required and $L = 128$ would be the largest simulation we can allow for.

Fig. 4 shows the cumulative mass function, $F(M, z)$, for $L = 128$, averaged over 16 realizations. For both spectra the output is shown for seven redshifts corresponding to fractions of the box contained in collapsed regions stepping by multiples of 1/8 (these choices were made simply to get well-spaced curves in the figure: we can reconstruct the curves at any intermediate time). The dashed lines show the PS prediction where $\sigma(m)$ is obtained from fits to the points shown in Fig. 3. In both cases the evolution is approximately self-similar. This can be seen more clearly in Fig. 5 which shows a differential plot, $\partial F/\partial \ln \nu$, where $\nu = \delta_c/\sigma(M, z)$ is the ordinate ($\nu = (M/M_*)^{1/2}$ for $n = 0$). Also shown are the PS and peaks theory predictions,

$$-\frac{\partial F}{\partial \ln \nu} = \frac{2\nu}{\sqrt{2\pi}} e^{-\frac{1}{2}\nu^2}. \tag{3}$$

When expressed in this way the functional form of the mass distribution is absolutely universal, i.e. it does not depend on any parameter of the simulation.

Consider first the $n = 0$ case. Here the differential mass curves seem to have the same shape as the PS prediction, but with a higher normalization at large ν (alternatively one could say that δ_c should be reduced slightly so as to shift the predicted curve to the right). The reason for this is discussed in Section 3.3 below. There is no evidence of a departure from the PS curve at a mass of 64, corresponding to the size of smoothing blocks of side 4 (this is in contrast to the $n = -2$

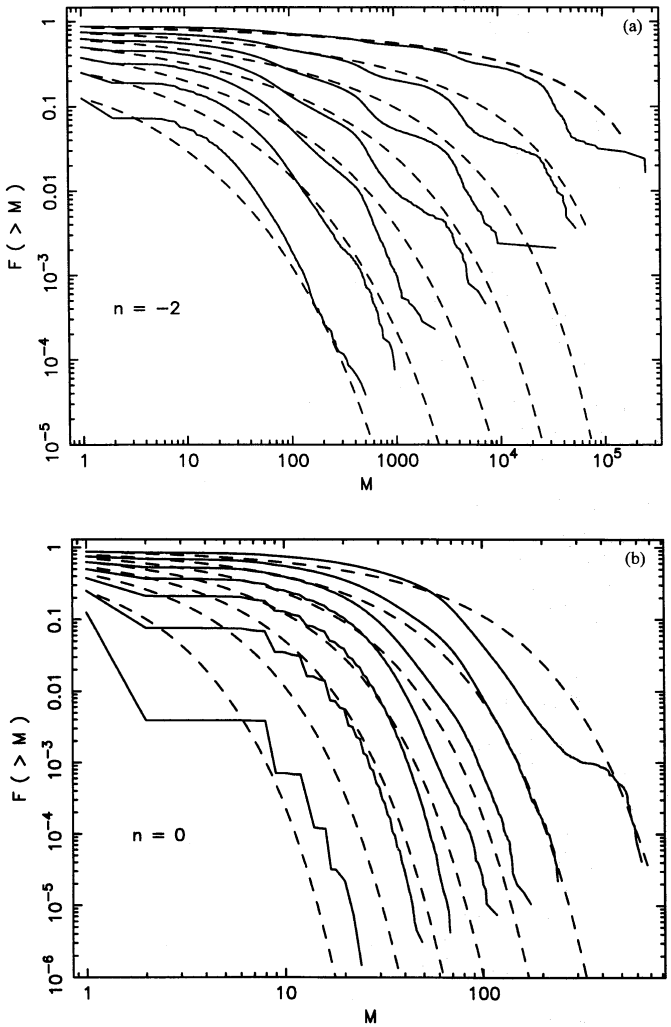

Figure 4. The average cumulative mass function for the sixteen $L = 128$ boxes at seven different output times (starting at the bottom curve when 1/8 of the box is in collapsed regions, and then increasing by multiples of 1/8) for (a) $n = -2$ (with $M_* = 0.18, 0.79, 2.90, 11.22, 54.31, 437.44, 12\,516$ in cell-units) and (b) $n = 0$ (with $M_* = 0.74, 1.58, 2.63, 4.20, 7.09, 13.91, 39.67$ in cell-units). The dashed lines show the corresponding Press–Schechter predictions (with an extra factor of two).

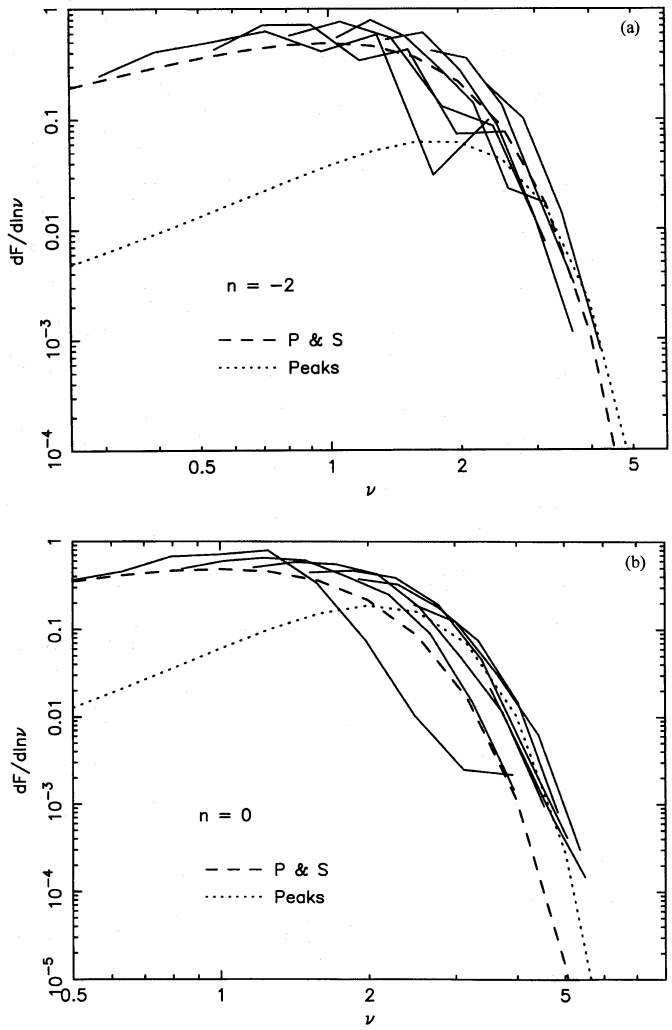

Figure 5. The differential mass function, $\partial F/\partial \ln \nu$, for the $L = 128$ box at the same output times as those from the previous figure: (a) $n = -2$, (b) $n = 0$. The dashed and dotted lines show the corresponding Press–Schechter and peaks theory predictions.

case). The self-similarity breaks down at the final output time, when $\frac{7}{8}$ of the mass is in haloes. The maximum mass of collapsed haloes at earlier times is quite small, less than 300 even for the largest box, $L = 256$. Given that the smallest haloes to collapse in our model (apart from isolated cells) have mass 8, then this gives a very small dynamic range.

The curves for the steeper spectrum, $n = -2$, extend to much higher masses because the spectrum has more power on large scales than for $n = 0$. Here we do see evidence of kinks at the blocking masses of 64, 512, 4096 and 32 768: there is an excess of haloes of slightly higher mass and a deficit at intermediate masses. Apart from the scatter associated with this blocking, the evolution is once again self-similar.

If we transcribe the plots in Fig. 5 on to the plots of fig. 1 from Lacey & Cole (1994) we conclude that, to bring our mass functions in close agreement with the N-body data (which fit reasonably well the PS prediction), the value of δ_c has to be increased above that of the top-hat model (the $n = -2$ mass functions do not need much adjustment). This is in agreement with the fact that our model assigns, by definition, overdensities to the resulting haloes which are systematically higher than their measured values, and one way of eliminating this offset would be raising the value of δ_c, as discussed in Section 3.2 below.

3.2 Properties of haloes

Fig. 6 shows a projection of the largest haloes in one $L = 128$ box of each spectral type at a time when half the mass has collapsed into haloes. Many of the irregular shapes that are visible are due to projection effects.

Our haloes exhibit a wide variety of shapes (Note, however, that these refer to the initial, Lagrangian particle distribution and do not directly correspond to the shape of the haloes where they end up at some later time.) They typically start with axial ratios of 1:1 (for collapse of isolated blocks as in Fig. 2b) or 1:1.5 (for the collapse of overlapping blocks as in Fig. 2c), developing rapidly to more complex structures with a great variety of shapes. Fig. 7 shows the distribution of axial ratios for all haloes of mass greater than or equal to 8 (for both spectra) in the Lagrangian distribution: there is little difference between the two. The haloes show a wide range of triaxiality ranging from prolate to oblate (while in the block

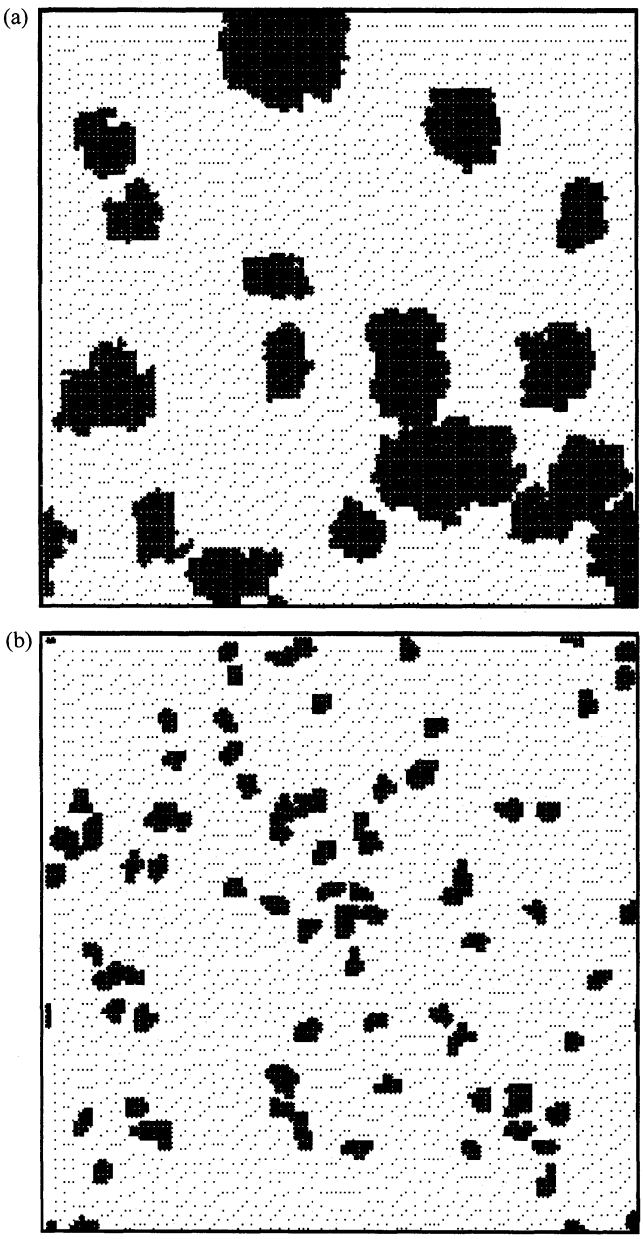

Figure 6. Projections of the distribution of haloes taken at the time when half the box is contained in collapsed structures: (a) $n = -2$, mass greater than 700; (b) $n = 0$, mass greater than 50.

model they are always 'maximally triaxial', i.e. $a_2^2 = a_1 a_3$ with $a_1 > a_2 > a_3$ being the semi-axes).

A drawback of our method comes from the fact that the overdensity of a collapsing block, δ_b, is not necessarily equal to the mean overdensity of the resulting halo, δ_h. It is the former value which we must associate with the halo if the topology of the merger tree is to be preserved (or at least we must maintain the same ordering of densities for haloes as for their parent blocks). The differences can be quantified in terms of the ratio $\chi = (\delta_b - \delta_h)/\delta_b$ which is plotted in Fig. 8 at a time when half the mass is in collapsed structures: we show the mean value plus 1σ error bars.

Note first that haloes of fewer than eight cells have overdensities that are much less than the assigned one. These structures are, however, leftovers of the merging process (the smallest blocks have a mass of 8 units) and so they should not be considered as collapsed haloes, but rather as clouds of intergalactic material to be accreted later by a neighbouring halo.

For haloes of mass 8 or larger the agreement is much better, but nevertheless the true overdensity of a halo remains systematically lower than the assigned one. The effect is largest for $n = 0$ where the mean value of χ is about 0.15. For $n = -2$, it varies from approximately zero in the largest haloes to 0.1 in the low-mass ones. The reason for the offset is that high-density cells can contribute to the overdensity of more than one block. Referring again to Fig. 2(c), if the region of overlap between the two blocks were of higher density than its surroundings then the density of the whole halo would be lower than that of either block from which it is constructed. If desired the assigned halo densities could be systematically reduced to bring them into agreement with the measured ones; equivalently one could raise the value of the critical density, δ_c, required for collapse above that of the top-hat model.

More serious is the variance of χ, approximately 0.2, which means that two haloes with the same assigned density can have quite disparate true overdensities. The model supposes that they collapse at the same time, whereas the full non-linear evolution would presumably show otherwise. We occasionally find some high-mass haloes (mass greater than 8) with large χ, which contribute significantly to the enlargement of the error bars at those scales. These are effectively leftovers of the merging process and should not be treated as collapsed haloes in subsequent applications of the method (that is, in a realistic galaxy formation modelling they should be considered as sources of material to be accreted at a later stage of the hierarchy).

The variance in χ is unwelcome but is only one contribution to the dispersion between overdensity and collapse time. We note that N-body simulations show for each particle a poor correspondence between the expected mass of its parent halo (predicted from the initial conditions) and the true value measured from a numerical simulation, evolved from the same initial conditions (White 1995; Bond et al. 1991). Moreover, gravitational collapse is clearly not as simple as the spherical model assumes. However, if a simple semi-analytical model of the gravitational clustering is desired, then the simple relation between collapse redshift and initial overdensity given by the spherical model seems the most obvious choice.

3.3 The number density of high-mass haloes

Our method passes the test for self-similarity, yet for $n = 0$ it predicts more high-mass haloes than Press–Schechter or other methods based on similar ideas, such as the block model. This is an expected outcome of our use of overlapping grids. In this sense our method is closer in spirit to the peaks theory of Bardeen et al. (1986), as we attempt to explain below.

Press–Schechter does not count the number of haloes of a given mass. Rather, it counts the fraction of the Universe where, if one were to put down a top-hat filter of the appropriate mass, the overdensity would exceed a certain critical threshold. Regions that just poke above this threshold for a single position of their centres contribute nothing to the mass function. This becomes increasingly likely as one moves to rarer and rarer objects (of higher and higher mass). For these it is much better simply to count the number of peaks that exceed the threshold density after filtering on the appropriate

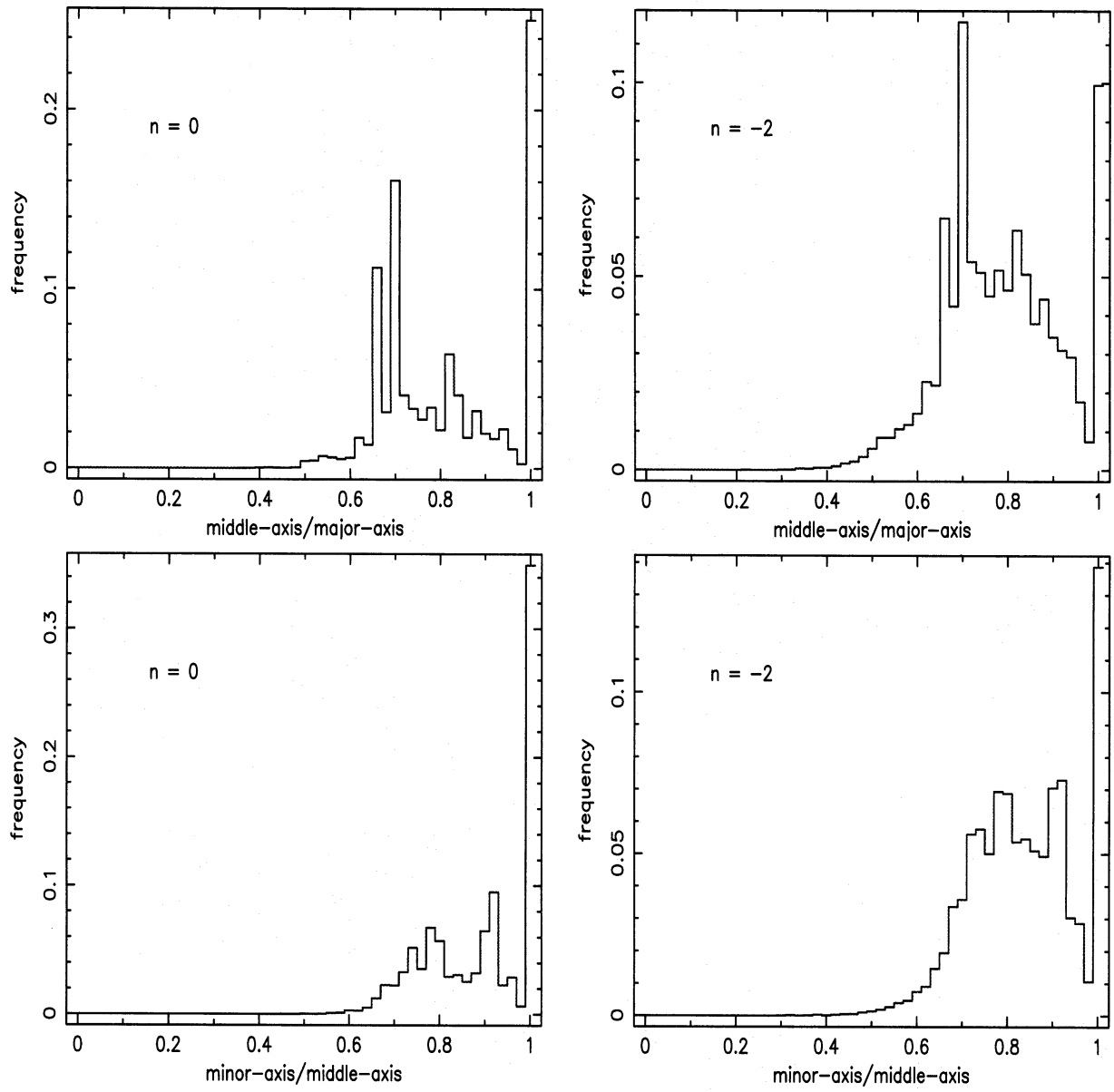

Figure 7. Distribution of axial ratios for haloes taken from one $L = 128$ realization when half of the box mass is contained in collapsed structures: $n = 0$ and $n = -2$, mass greater than or equal to 8.

scale. The necessary theory has been exhaustively analysed by Bardeen et al. (1986) who showed that uncorrected PS (without the extra factor of two) underestimates the number of high-mass haloes by a factor $\alpha^{3/2}\nu^3$, where $\delta_c = \nu\sigma(m)$ (this result is for a Gaussian filter but similar results will hold for all filters with just a small difference in scaling).

A direct demonstration of the above difference between Press–Schechter and the actual number of high-mass peaks in the density field is shown by the numbers in Table 1. Columns 2–9 show the measured number of blocks that exceed the density threshold given in the first column in each of the eight sub-grids of mass 512. These agree with the PS prediction, as indeed they should by construction. When we combine the various grids, however, an interesting thing happens. Column 10 shows the number of separate (i.e. non-overlapping) overdense blocks in the combined grid. At high overdensity all the haloes we have identified are distinct (they exceed the threshold for just one position of the smoothing grid). The total number of haloes is therefore greatly in excess of the PS prediction and far closer to that given by peaks theory. For $n = -2$ the excess is a factor of three which brings them into approximate agreement once the PS prediction has the extra factor of two applied. For $n = 0$, however, the difference is much larger and the number of peaks is a factor of 3–4 larger than even the corrected PS estimate. This goes some way to explaining the difference between the PS prediction and the measured cumulative mass function in Fig. 5.

At lower overdensities the disagreement in the predicted number of haloes in the peaks and PS models is much less severe. The peaks methodology breaks down once neighbouring haloes begin to overlap: this can be seen in the reduced number of 1σ peaks found in our $n = -2$ realizations. The predicted mass function from peaks theory is shown in Fig. 5 as a dotted line and can be seen to agree well with our model at high overdensity.

Our use of overlapping grids is therefore crucial. They ensure that all haloes are approximately centred within one of the grid cells. Other methods, such as the block model,

Table 1. Number of blocks of mass 512 ($L = 128$) above the density thresholds 3σ, 2σ, σ: (a) $n = -2$, (b) $n = 0$. Columns 2–9 correspond to each of the eight sub-grids used for that level of refinement, Column 10 shows the number of isolated peaks and Column 11 shows the PS prediction of the expected number of haloes of this mass.

(a) over-density	single grids 1	2	3	4	5	6	7	8	combined grids	PS expected number
$\geqslant 3\sigma$	5	7	4	1	6	7	6	8	18	5.5±2.3
$\geqslant 2\sigma$	88	87	92	77	92	87	79	88	109	93.2±9.7
$\geqslant 1\sigma$	628	654	630	662	628	648	639	647	281	650±26

(b) over-density	single grids 1	2	3	4	5	6	7	8	combined grids	PS expected number
$\geqslant 3\sigma$	6	7	4	9	4	7	8	7	38	5.5±2.3
$\geqslant 2\sigma$	99	99	96	82	90	86	99	90	324	93.2±9.7
$\geqslant 1\sigma$	630	641	649	656	652	617	670	677	671	650±26

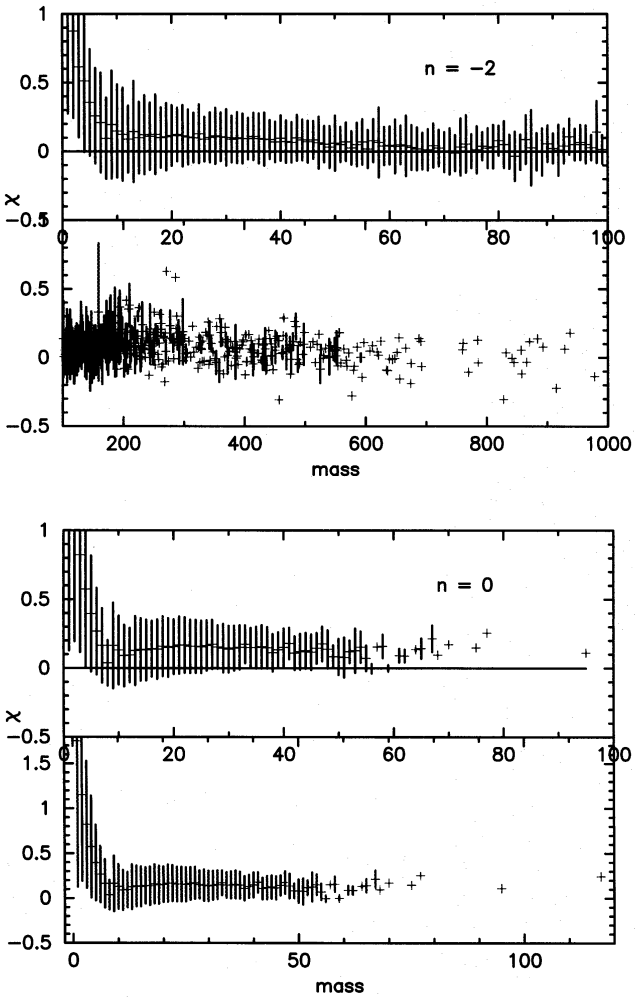

Figure 8. The relative difference between the assigned and the true overdensities of haloes, for the $L = 128$ box at the time when half of its mass is contained in collapsed regions.

which have fixed borders between mass cells, have difficulty in detecting structures that cross cell boundaries and are, by construction, forced to agree with Press–Schechter. This can lead to a gross underestimate of the number of rare, high-mass peaks, especially for steep spectra. This does not imply that our method necessarily gives a better description of the growth of structure in the Universe because all these theoretical models are highly idealized. Substructure may lengthen collapse times and tidal field may need to be taken into account.

4 DISCUSSION

We have presented a new method of constructing a hierarchical merger tree based on actual realizations of the linear density field. We smooth on a set of interlaced, cubical grids on a variety of mass-scales, then order in terms of decreasing density. We run down the resulting list, merging together overlapping blocks to form collapsed haloes. The main properties of our model are as follows.

(1) The model exhibits the scaling behaviour expected from power-law spectra.

(2) For a flat, white-noise spectrum, $n = 0$, the mass function is well-fitted by the PS model, and comes into agreement with N-body results of Lacey & Cole (1994), provided that we raise the normalization at high overdensity by a factor of 3–4 or alternatively increase the value of δ_c. This difference arises because of the failure of the PS method to count the correct number of massive, rare (high-ν) peaks. The dynamic range for the masses of haloes for this spectrum is quite small – at a time when three quarters of the box is in collapsed structures, the mass of the largest halo is just 300 cells, even for the box of side $L = 256$.

(3) The mass function for a steeper spectrum, $n = -2$, lies much closer to the PS prediction, but shows small kinks at masses corresponding to those of the smoothing blocks. The dynamic range is much larger than for the $n = 0$ case.

(4) The collapsed haloes tend to be triaxial with a wide mixture of prolateness and oblateness.

(5) The mean overdensity of haloes tends to be lower than that assigned to them by our scheme, and to have a scatter of about 20 per cent about the mean value.

A shortcoming of our method is that the minimum halo mass is eight cells with a corresponding loss of dynamical range. We see in Fig. 4 that for $n = -2$ we achieve a useful mass range of approximately four decades whereas for $n = 0$ it is only one and a half decades. This is because $n = -2$ corresponds to a flat spectrum, $\sigma \propto M^{-1/6}$, with almost all scales collapsing simultaneously, while for $n = 0$ the spectrum

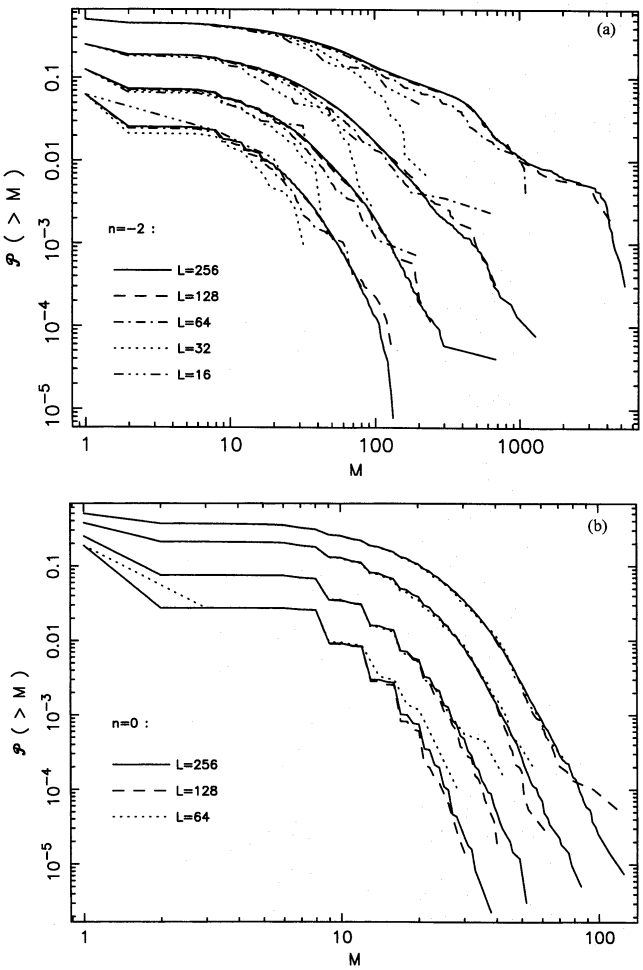

Figure 9. The cumulative mass function for boxes of variable resolution, as indicated: (a) $n = -2$, (b) $n = 0$. The collapsed fractions are indicated in the text.

decreases more rapidly, $\sigma \propto M^{-1/2}$, and the peaks are much more isolated. Fortunately, the physically motivated cold dark matter spectrum can be fitted by an $n = -2$ spectrum over a significant mass range, and so the loss of dynamical range is not so important in a realistic application of the method.

If we wish to use the merger tree for galaxy formation studies then the maximum useful box-size on a workstation is $L = 128$. Then if we simulate a large portion of the Universe, of mass say $10^{16}\,M_\odot$, a block of eight cells corresponds to $3.8 \times 10^{10}\,M_\odot$, which could represent at best a dwarf galactic halo. If the box represents a large galactic halo of mass $10^{13}\,M_\odot$, then we can resolve down almost to globular cluster scales.

We have carried out simulations with a range of box-lengths, $L = 16$, 32, 64, 128, and 256, in order to show the effect of variable resolution (we were not able to perform the $L = 16, 32$ simulations for $n = 0$ due to the lack of dynamical range). The base-cells in each case correspond to one set of blocks of side $256/L$ in the $L = 256$ simulation. The results are presented in Fig. 9 which shows the cumulative mass functions at a range of output times (1/16, 1/8, 1/4, 1/2 for $n = -2$ and 3/16, 1/4, 3/8, 1/2 for $n = 0$). The shape of the mass spectrum is similar but with a slight increase in dynamic range as one moves from $L = 64$ to $L = 256$.

Whether our method provides a better description of the formation of structure than other methods remains to be seen. The block model in particular seems to give good agreement with N-body simulations (Lacey & Cole 1994) which themselves are approximately fitted by Press–Schechter models. However, there is a limited dynamical range in the simulations, the results are sensitive to the precise model for identifying collapsed haloes and the critical overdensity for collapse is usually taken as a free parameter. Given these caveats it is hard to tell whether the fits are mediocre, adequate or good.

One advantage of our scheme is that it is based on an actual realization of a density field which can be used as the starting point for an N-body simulation. Thus we will not be limited to a statistical analysis, but will be able directly to compare individual structures identified in the linear density field with non-linear haloes that form in the simulation. Initial results from other studies (Bond et al. 1991; Thomas & Couchman 1992) suggest that the correspondence is approximate at best, and we may be forced to consider the effect of tidal fields on halo evolution. We have not yet carried out the necessary N-body simulations because we have not up to now had access to the necessary super-computing facilities to evolve (and analyse!) a 256^3 box of particles. Such datasets will soon become available as part of the Virgo Consortium project on the UK's Cray T3D facility, and we intend to report the results in a subsequent paper.

Nevertheless, even in the absence of the numerical tests, we feel that our method is a viable alternative to other methods of calculating the merging history of galactic haloes. It passes the test of self-similarity yet predicts more high-mass haloes than other methods. It has the disadvantage of losing a factor of eight in resolution at low masses, but above this it has a smooth mass spectrum and is not restricted to masses that are a power of two. We intend to contrast the predictions of the block model and this current method in models of galaxy formation such as those discussed by Kauffmann, White & Guiderdoni (1993) and Cole et al. (1994), and explore the role of pre-galactic cooling flows (Nulsen & Fabian 1995).

ACKNOWLEDGMENTS

DDCR would like to acknowledge support from JNICT (Portugal) through programme PRAXIS XXI (grant number BD/2802/93-RM). Part of this paper was written while PAT was at the Institute for Theoretical Physics at Santa Barbara and as such was supported in part by the National Science Foundation under Grant Number PHY89-04035. The paper was completed while PAT was holding a Nuffield Foundation Science Research Lectureship. We would like to thank Shaun Cole for detailed comments which helped to improve the presentation of the paper and for providing us with a copy of the block model program. The production of this paper was aided by use of the Starlink Minor Node at Sussex.

REFERENCES

Bardeen J.M., Bond J.R., Kaiser N., Szalay A.S., 1986, ApJ, 304, 15
Baugh C.M., Cole S., Frenk C.S., 1996, MNRAS, Preprint astro-ph/9602085
Bond J.R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bower R.J., 1991, MNRAS, 248, 332
Cole S., 1991, ApJ, 367, 45
Cole S., Kaiser N., 1988, MNRAS, 233, 637

Cole S., Aragon-Salamanca A., Frenk C.S., Navarro J.F., Zepf S.E., 1994, MNRAS, 271, 781
Efstathiou G., Frenk C.S., White S.D.M., Davis M., 1988, MNRAS, 235, 715
Gunn J.E., Gott J.R., 1972, ApJ, 176, 1
Heyl J.S., Cole S., Frenk C.S., Navarro J., 1995, ApJ, 274, 755
Kauffmann G., White S.D.M., 1993, MNRAS, 261, 921
Kauffmann G., White S.D.M., Guiderdoni B., 1993, MNRAS, 264, 201
Lacey C.G., Cole S., 1993, MNRAS, 262, 627
Lacey C.G., Cole S., 1994, MNRAS, 271, 676
Nulsen P.E.J., Fabian A.C., 1995, MNRAS, 277, 561
Press W.H., Schechter P.G., 1974, ApJ, 187, 425 (PS)
Thomas P.A., Couchman H.M.P., 1992, MNRAS, 257, 11
White S.D.M., 1995, Formation and Evolution of Galaxies: Lectures given at Les Houches, 1993 August

This paper has been produced using the Royal Astronomical Society/Blackwell Science LaTeX style file.