A catalogue of galaxy cluster models

Kampen, E. van; Katgert, P.

Mon. Not. R. Astron. Soc. **289,** 327–348 (1997)

A catalogue of galaxy cluster models

Eelco van Kampen[1,2★†] and Peter Katgert[1]
[1]*Sterrewacht Leiden, Postbus 9513, 2300 RA Leiden, The Netherlands*
[2]*Royal Observatory Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*

Accepted 1997 February 17. Received 1997 February 10; in original form 1996 November 18

ABSTRACT

We present a technique to construct a fair sample of simulated galaxy clusters, and build such a sample for a specific cosmological structure formation scenario. Conventionally one extracts such a sample from a single low-resolution large-scale simulation. Here we simulate the clusters individually at high resolution. This is made possible by the method of constrained random fields, in which one can put *linear* constraints on peaks in the initial smoothed density field.

We assume that these peaks are the progenitors of present-day rich clusters, and select clusters for a catalogue by selecting their initial peak parameters. We find that the final cluster mass can be well approximated by a linear function of both the amplitude and the curvature of the initial density peak. Because the probability distributions of these peak parameters are known, we can construct a model catalogue selected on expected final cluster mass. Such a catalogue will not have a well-defined richness limit, because the relation between richness and mass is fairly broad. However, by applying the appropriate completeness corrections, the results for the mass-selected catalogue can be compared with observations for richness-selected cluster catalogues.

Each cluster model is evolved from its constrained initial conditions by means of an N-body integrator. This includes an algorithm for galaxy formation, so we produce two-component models consisting of dark matter background particles and galaxies. The latter allow us to obtain *directly* the observable properties of the cluster models, and match these to the observed properties to define the present time in the models, and thus derive the amplitude of the initial density fluctuation spectrum, σ_8.

We build a model cluster catalogue for the $\Omega_0 = 1$ cold dark matter (CDM) scenario that is designed to mimic the ENACS sample of rich Abell clusters. We use the distribution of richness, corrected for incompleteness, to fix the present epoch. We find $\sigma_8 = 0.4$–0.5, which is consistent with other determinations. The catalogue is 70 per cent complete for a richness larger than 50, but we do have a *complete* subsample for richnesses larger than 75.

As a first test we compare the cumulative distribution of line-of-sight velocity dispersions to those found for several observational samples, and find that they match best for $\sigma_8 \approx 0.4$. This means that we find consistent values for σ_8 for the CDM $\Omega_0 = 1$ scenario on cluster scales.

Key words: galaxies: clusters: general – galaxies: formation – galaxies: evolution – cosmology: theory – dark matter – large-scale structure of Universe.

1 INTRODUCTION

Clusters of galaxies are the largest bound structures in the Universe, but they are dynamically young and still contain signatures of their initital conditions. Statistical properties of samples of galaxy clusters can therefore provide constraints for models of large-scale structure formation in the Universe. Large catalogues of observed clusters of galaxies have been compiled, of which the Abell, Corwin & Olowin (1990, ACO from here on) catalogue is the best known, but complete X-ray-selected cluster samples have recently become available (e.g. Ebeling et al. 1996). The ACO clusters have been most extensively observed in the optical, as that is where they have been defined. A renewed interest in their observational properties has recently led to a flood of new data,

★Present address: Theoretical Astrophysics Center, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark.
†E-mail: eelco@tac.dk (EVK)

and to a better understanding of their properties. It is therefore desirable to construct a model catalogue that mimics observed ones, thus allowing a meaningful statistical comparison between observations and predictions. In this paper we discuss a technique to construct such a model catalogue, and we build a catalogue for one particular cosmological scenario, namely $\Omega_0 = 1$ standard cold dark matter (CDM).

One approach for obtaining a sample of model clusters is to extract clusters from large-scale *N*-body simulations. This has been carried out by Frenk et al. (1990), and more recently by van Haarlem, Frenk & White (1997) and Eke, Cole & Frenk (1996). The advantage of this approach is that completeness of the resulting catalogue is automatically guaranteed. However, a disadvantage is that the resolution of the extracted cluster models is relatively poor since voids and filaments, which are simulated as well, take up most of the volume. This means that the fraction of particles in clusters is only about 10 per cent. We avoid this problem by running an ensemble of individual, high-resolution cluster simulations. This technique has been used by various authors in order to study the influence of the choice of initial conditions on the final cluster properties (Evrard 1989; Evrard et al. 1993; van Haarlem & van de Weygaert 1993). However, because these authors did not aim to construct a fair sample, a statistical comparison with observed catalogues as performed by Frenk et al. (1990) and others was not feasible. This is exactly the goal we aim for here: to construct a statistically fair catalogue of individual cluster models, each of which is simulated at relatively high resolution. In this context the term 'fair sample' is meant to indicate a sample that is representative of a volume-limited sample.

We add to this a method that includes the formation and merging of galaxies in dissipationless *N*-body simulations (van Kampen 1997). The basic idea is that each group of particles that is roughly in virial equilibrium is replaced by a single particle with mass, position, velocity and softening corresponding to that group. This is performed several times during the evolution. Since groups that form later in the evolution can contain one or more earlier-formed galaxies, both growing and merging of galaxies are included in this simulation method, although somewhat crudely (improvement on this scheme is in progress). One reason for using this method is that groups of particles that can be identified as galaxies need to be protected from destruction within the main cluster by numerical two-body disruption (Carlberg 1994; van Kampen 1995). However, it also is highly desirable to be able to compare *directly* the properties of the modelled galaxy distribution to observations, instead of making assumptions on the relation between galaxies and dark matter.

Both the fluctuations in the galaxy distribution and those in the underlying dark matter distribution grow with time. The relation between the two (usually described by some bias parameter) may change with time. In the simulations, the present epoch is identified as the time when the properties of the simulated galaxy distribution match the observed ones. We use the normalization obtained by van Kampen (1994) from a set of field models run for the same cosmology as adopted here.

In order to construct a sample of individually simulated cluster models, we need to make a few assumptions. We assume that clusters of galaxies, as we observe them at the present epoch, formed from peaks in the initial density field smoothed at the length-scale appropriate for clusters. These initial peaks have several characteristics that can be used to predict whether or not they will evolve into rich Abell clusters. We also assume that the initial density fluctuations are Gaussian distributed. This allows the use of the theory of Gaussian random fields (Rice 1954; Doroshkevich 1970; Adler 1981; Peacock & Heavens 1985; Bardeen et al. 1986, BBKS from here on) which provides the probability distributions for the peak parameters in the early (linear) stages of the evolution of the density field.

The construction of a sample of cluster models involves the generation of a set of initial conditions that produce a set of simulated clusters with the same *statistical* properties as for an observed sample. Such realizations of initial matter distributions that produce a cluster with specific properties can be obtained by means of the Hoffman–Ribak method of constructing constrained random fields (Hoffman & Ribak 1991). We used the implementation of this method by van de Weygaert & Bertschinger (1996). An important limitation of this method is that it can only constrain *linear* functionals of a field. This means that the defining quantity of the catalogue has to be a linear functional as well. Given this limitation on the constrained random field method of generating initial conditions, we argue (in Section 3.2) that the final cluster mass is the best choice for the defining quantity.

In order to construct a catalogue, we assume that $\mathcal{N}$ rich Abell clusters within a given volume originate from those $\mathcal{N}$ initial density peaks that have the largest predicted final mass. We use the amplitude *and* the curvature of the initial peak to predict the final cluster mass (Section 3.3 and Appendix B). This combination appears to give a better prediction of the actual final cluster masses than that based on the initial amplitude alone, which would be sufficient if the evolution would remain linear. Using the probability distributions from Gaussian random field theory we then sample the distribution of expected total masses to construct a fair statistical sample.

In order to build a model catalogue, we need to choose a cosmological scenario, and an observational catalogue to compare to. We adopt the standard CDM $\Omega_0 = 1$ scenario, because it allows us to compare with other studies (Frenk et al. 1990; White, Efstathiou & Frenk 1993; Eke et al. 1996). It also has the property that the amplitude of the dark matter fluctuations need not be fixed at the outset. This is important, since we do not know a priori how the galaxy distribution will evolve with respect to the dark matter.

The observational catalogue we try to mimic is a complete subset of the Abell catalogue for which data was gathered in the context of an ESO Key-programme (den Hartog 1995; Katgert et al. 1996; Mazure et al. 1996). This sample, denoted as ENACS from here on, is the largest well-defined catalogue presently available with extensive information on the internal structure of clusters of galaxies.

Because there is bound to be significant scatter in the relation between richness and the expected final cluster mass, our model catalogue, constructed to be complete in mass, will not be complete in richness. In Section 5 we discuss the cause of this incompleteness, and apply a statistical method to correct for it. This gives us complete distribution functions for cluster properties like richness and line-of-sight velocity dispersion.

In Section 6 we present the model counterpart of the ESO-Kpgm catalogue for the $\Omega_0 = 1$ CDM scenario. We use the distribution of cluster richness in addition to the galaxy autocorrelation function (van Kampen 1996) to fix the normalization (in terms of σ_8). We find that both measures provide a consistent timing for the $\Omega_0 = 1$ CDM scenario.

We can use the catalogue to test the validity of the chosen scenario by means of other measures, like the distribution of cluster shapes or line-of-sight velocity dispersions. We describe several such tests in Section 7. A preliminary comparison to observations

and a more detailed study of the intrinsic properties of the model clusters can be found in van Kampen (1994).

2 RICH ABELL CLUSTERS AND PEAKS IN GAUSSIAN RANDOM FIELDS

Our aim is to build a sample of model clusters that can be compared to an observed catalogue. We therefore need to find a relation between the defining criteria for observed and model clusters. The former are found within the discrete distribution of galaxies as observed in projection on the sky, while the latter are defined in the continuous density field which we sample with the particle distribution in a numerical simulation. In this section we discuss the observational and theoretical ways to define clusters of galaxies, and also the numerical methods to build cluster models.

2.1 Abell's observational definition of a rich cluster

We will compare our models to rich Abell (1958) clusters, so we first list Abell's main criteria for the definition of an observed richness class ≥ 1 cluster. They read as follows.

Richness criterion – 'A rich cluster must contain at least 50 members that are not more than 2 magnitudes fainter than the third brightest member'.

Compactness criterion – 'A rich cluster must be sufficiently compact that its 50 or more members are within a given (projected) radial distance R of its centre'.

Abell's choice $R = 1.5\,h^{-1}$ Mpc is usually referred to as the 'Abell radius'. The richness and compactness criteria jointly set the length- and mass-scale of Abell clusters. We now know that the Abell radius is somewhat small compared with more physical scalelengths like the turnaround radius, so the Abell radius defines just the central region of a cluster. Most observations are restricted to an area within this radius, or to an even smaller region (the cluster catalogue extracted from, e.g., the APM survey uses half the Abell radius: Dalton et al. 1992).

Instead of a richness class we will use the richness measure $C_{\rm ACO}$, which actually gives the number of galaxies satisfying both of Abell's criteria.

2.2 Clusters as peaks in the smoothed density field

Theoretically, clusters are identified in the density contrast field

$$\delta(\boldsymbol{r}) \equiv \frac{\rho(\boldsymbol{r}) - \langle\rho(\boldsymbol{r})\rangle}{\langle\rho(\boldsymbol{r})\rangle} \tag{1}$$

as fluctuations at a certain mass-scale (e.g. Kolb & Turner 1990). This requires smoothing, or 'windowing', of the field at a corresponding length-scale. The smoothed density contrast field is

$$\delta_W(\boldsymbol{r}, R_W) = 4\pi \int \delta(|\boldsymbol{r} - \boldsymbol{r}'|) W(r', R_W) r'^2 \mathrm{d}r', \tag{2}$$

where $W(r, R_W)$ is a normalized spherically symmetric window function. Clusters are identified as the peaks in this smoothed density field, and the properties of these peaks characterize and define them. However, galaxy clusters exhibit highly non-linear evolution. This means that the present-day properties cannot be linked directly to the initial conditions from which they formed: we cannot simulate backwards in time. We are forced to make an educated guess for the initial conditions, simulate the formation of clusters starting from these conditions, and compare the resulting models to observations. This means that model clusters are necessarily defined in the initial density field, preferably at the epoch before clusters 'break away' or 'turn around' from the general expansion of the Universe.

To generate a catalogue of cluster models, we must define models by peak parameters in the initial density field that can be sampled from a known probability distribution. If we choose peak parameters that are *linear* functionals of Gaussian random fields filtered with a Gaussian filter, we can use the results of BBKS. We therefore choose the Gaussian filter

$$W_{\rm G}(r, R_{\rm G}) = (2\pi R_{\rm G}^2)^{-3/2} \mathrm{e}^{-r^2/2R_{\rm G}^2}, \tag{3}$$

and make the implicit assumption that structure forms from small Gaussian density fluctuations in the early Universe. Conveniently, integrals with a Gaussian kernel behave well and can often be found analytically.

Note that the probability distributions derived by BBKS are only valid for the initial, linear regime (i.e. $\delta_W \ll 1$). Non-linear evolution changes these distributions in an unknown manner, which makes the relation between the initial peak parameters and the final cluster model rather uncertain. Fortunately, the smoothing scale needed for cluster-sized peaks is large enough that the evolution of clusters in the smoothed density field is not too non-linear, so that the BBKS distributions can be used.

2.3 Initial peak parameters

2.3.1 *Definitions*

It is convenient to express peak parameters in terms of global parameters for the cosmological scenario in which a cluster is to be modelled. The general density fluctuation field smoothed with a Gaussian with length-scale $R_{\rm G}$ can be characterized by the set of spectral moments

$$\sigma_j^2(R_{\rm G}) \equiv \int \frac{k^2}{2\pi^2} |\delta_k|^2 \mathrm{e}^{-R_{\rm G}^2 k^2} k^{2j} \mathrm{d}k, \tag{4}$$

where δ_k is the Fourier transform of the density field. Two special spectral parameters are usually defined:

$$\gamma(R_{\rm G}) \equiv \frac{\sigma_1^2(R_{\rm G})}{\sigma_0(R_{\rm G})\sigma_2(R_{\rm G})} \quad \text{and} \quad R_*(R_{\rm G}) \equiv \sqrt{3}\frac{\sigma_1(R_{\rm G})}{\sigma_2(R_{\rm G})}. \tag{5}$$

They relate to the width of the spectrum and its correlation length respectively. For this paper we adopt the standard adiabatic CDM scenario for $\Omega_0 = 1$ with $R_{\rm G} = 4\,h^{-1}$ Mpc, which has $\gamma(4\,h^{-1}) = 0.729$ and $R_*(4\,h^{-1}) = 5.16\,h^{-1}$ Mpc.

We define the origin of the coordinate system at the centre of the peak, and define all peak parameters at this position. The primary parameter of a peak is its amplitude $\delta_{\rm G}(0, R_{\rm G})$, expressed in units of the rms fluctuation at the same scale:

$$\delta_{\rm G}(0, R_{\rm G}) \equiv \nu\sigma_0(R_{\rm G}). \tag{6}$$

We can further specify the peak using the first and second spatial derivatives of $\delta_{\rm G}(\boldsymbol{r}, R_{\rm G})$. The gradient $\nabla\delta_{\rm G}(0, R_{\rm G})$ is zero for an extremum. This accounts for three parameters. The second derivative tensor $\nabla_i\nabla_j\delta_{\rm G}$ has six independent elements, resulting in six additional parameters. They characterize the shape (two parameters), the orientation (three parameters) and the size (one parameter) of the peak. The latter is the trace of $\nabla_i\nabla_j\delta_{\rm G}$, and is usually defined in a dimensionless form as the *curvature*:

$$x(\boldsymbol{r}, R_{\rm G}) \equiv -\frac{\nabla^2\delta_{\rm G}(\boldsymbol{r}, R_{\rm G})}{\sigma_2(R_{\rm G})}. \tag{7}$$

This curvature (a measure for the width of the peak) is positive for a peak and negative for a dip. The shape of the peak can be characterized by the two axial ratios of the isodensity contours around the peak, as found from the small-r approximation of the smoothed density field around the peak using a second-order Taylor expansion in spherical coordinates (r, θ, ϕ) :

$$\delta_{\rm G}(r) = \delta_{\rm G}(0) + \frac{r^2}{2}\nabla^2\delta_{\rm G}(0)\mathcal{F}_5(a_{12}, a_{13}, \theta, \phi)$$

$$= \nu\sigma_0(R_{\rm G}) - \frac{r^2}{2}x(0, R_{\rm G})\sigma_2(R_{\rm G})\mathcal{F}_5(a_{12}, a_{13}, \theta, \phi) \quad (8)$$

(BBKS). Here $\mathcal{F}_5(a_{12}, a_{13}, \theta, \phi)$, given by equation (A7) in Appendix A, describes the asphericity, and

$$a_{12} \equiv \frac{a}{b}\ , \quad a_{13} \equiv \frac{a}{c} \quad (9)$$

are the axial ratios, where a is the major axis, b the intermediate axis, and c the minor axis.

2.3.2 *Probability distributions*

BBKS give the probability distributions for the peak parameters (the amplitude ν, the curvature x and the shape parameters a_{12} and a_{13}) in the linear field smoothed with a Gaussian window of size $R_{\rm G}$.

In order to make dependences transparent, we use functions $\mathcal{F}$ (adapted from BBKS) which are listed for reference in Appendix A. The differential peak amplitude probability density is given in the form of the expected number of peaks within a volume V with an amplitude in the range $[\nu, \nu + {\rm d}\nu]$:

$$N_{\rm peak}(\nu, R_{\rm G}){\rm d}\nu = V\frac{{\rm e}^{-\nu^2/2}}{(2\pi)^2 R_*^3}\mathcal{F}_1(\gamma, \nu){\rm d}\nu. \quad (10)$$

The other probabilities are conditional; given a peak with height ν, the probability distribution for its curvature to be equal to x is

$$P(x|\nu, R_{\rm G}) = \frac{{\rm e}^{[-(x-\gamma\nu)^2/2(1-\gamma^2)]}}{\sqrt{2\pi(1-\gamma^2)}}\frac{\mathcal{F}_2(x)}{\mathcal{F}_1(\gamma, \nu)}. \quad (11)$$

For a peak with curvature x, the axial ratios a_{12} and a_{13} have a joint probability distribution

$$P(a_{12}, a_{13}|x) = \frac{225}{\sqrt{\pi}}\frac{\mathcal{F}_3(a_{12}, a_{13})x^8}{\mathcal{F}_2(x)}\exp[-\tfrac{5}{2}x^2/\mathcal{F}_4(a_{12}, a_{13})]. \quad (12)$$

This distribution does not depend on ν and $R_{\rm G}$ directly, but only through x. All probabilities do depend on the spectral parameters $\gamma(R_{\rm G})$ and $R_*(R_{\rm G})$.

2.4 Building individual cluster models

2.4.1 *Constrained initial conditions*

It is straightforward to generate initial configurations for an average patch of universe in a given cosmological scenario (e.g. Efstathiou et al. 1985). However, for our purpose we need initial conditions that will produce a cluster defined by certain peak parameters. This is made possible by the method of constrained random fields, pioneered by Bertschinger (1987). We use a code written by van de Weygaert & Bertschinger (1996), based on the Hoffman & Ribak (1991) method, which is very well suited for our purpose. It can sample an unsmoothed density field constrained to form a peak characterized by linear functionals of the Gaussian smoothed density field. Therefore the only restriction of this scheme is that we must use *linear* peak parameters to define clusters in the initial conditions. This is an important limitation on the choice of a quantity that defines a model cluster catalogue (Section 3.2).

2.4.2 *Non-linear evolution and galaxy formation*

The highly non-linear evolution of clusters of galaxies requires the use of N-body methods. However, N-body simulations suffer from a problem that is particularly severe for galaxy clusters: small groups of particles that represent galaxies in such simulations become disrupted by numerical and tidal effects (Carlberg 1994; van Kampen 1995). This problem can be solved by replacing those groups of particles by a single 'galaxy particle' just after they have formed into a virialized system that resembles a galactic halo. This new galaxy particle is softened according to the size of the group it replaces, and adopts the position and momentum of the centre of mass of the original group. In van Kampen (1995) such galaxy particles were only formed at a single epoch. Here an extension of this scheme to 'continuous' galaxy formation, which consists simply of applying the algorithm several times during the evolution, is used (van Kampen 1997). Since groups of particles can contain already-formed galaxies, merging is included as well, although somewhat crudely.

The actual N-body code we use is the Barnes & Hut (1986) treecode, slightly adapted for the modelling of clusters (see van Kampen 1995 for details) and supplemented with the galaxy formation algorithm. For all simulations we generated initial conditions in a 64^3 cube with a size of $32\,h^{-1}$ Mpc, using an implementation of the Hoffman & Ribak (1991) method by van de Weygaert & Bertschinger (1996). Because we use non-periodic boundary conditions (the universe outside the simulation volume is homogeneous), we cut a sphere out of the volume containing 64^3 particles. The radius of this sphere is chosen such that it is sufficiently large to easily include the final turnaround radius, thus containing the full cluster and its region of influence (see Appendix C for a discussion on the reliability of the simulations).

2.4.3 *Model cluster richness*

Abell's richness measure $C_{\rm ACO}$ is defined in a magnitude interval. This means that we should correctly model the *shape* and the *amplitude* of the luminosity function in order to obtain a sensible richness measure from our numerical models. Frenk et al. (1988) found for CDM models that the shape of their mass function differs from that of the observed luminosity function. However, such results strongly depend on how one identifies galactic haloes at the present epoch (Suginohara & Suto 1992). The influence of the details of the modelling of galaxy formation is also present in our simulations. We expect to obtain a more realistic luminosity function since we identify and form galaxies *during* the simulation, rather than only at the end. If we assume a constant mass-to-light ratio, which should be reasonable for the factor of 30 in mass range, we find that the resulting luminosity function fits the observed one quite well. This means that we can obtain a realistic richness measure for our cluster models.

The assumed constant M/L implies that the mass interval corresponding to Abell's richness criterion ranges from the mass of the third most massive galaxy to a mass 6.31 times smaller. There is one property of our galaxies that plays a role here: the galaxy particle masses are multiples of the N-body particle mass, which is 7×10^{10} $M_\odot$ for our simulations. Before we determine the richness, we add

or subtract a random fraction of half the N-body particle mass. This removes the artifical discreteness sufficiently well.

Note that we do not have sufficient foreground and background galaxies in our simulations to mimic the ACO richness measure. For this reason we will use a different richness measure, which is statistically equivalent to Abell's (see Section 5.2).

3 DEFINITION OF A MODEL CATALOGUE

3.1 From initial density peaks to rich Abell clusters

In order to build a catalogue of individual cluster models, we need to assume a one-to-one mapping of the peak parameters ν and x in the initial smoothed density to properties of clusters in the final unsmoothed density field. However, this mapping might be complicated by initial peaks merging into a single final cluster, a worry that comes in two flavours.

3.1.1 *Mergers of rich cluster peaks*

In a hierarchical structure formation scenario, like CDM, merging of clumps into larger clumps is an ongoing process. Because we assume a one-to-one mapping of peaks in the initial density field to overdensities in the final density distribution, we need to check if merging is still important on cluster scales. If evolution would remain linear, peaks that were separate in the initial smoothed density field would remain separate for the same comoving smoothing scale. They can only merge due to non-linear evolution of the density field. We estimate the fraction of clusters that can merge from the statistics of the smoothed density field, i.e. the correlation function of Abell cluster peaks, and their peculiar velocities. The average peculiar velocity for a cluster-sized peak is about $300\,\mathrm{km\,s^{-1}}$, with a maximum of about $600\,\mathrm{km\,s^{-1}}$, for an $\Omega_0 = 1$ CDM cosmology with $\sigma_8 = 0.6$ (de Theije, van Kampen & Slijkhuis 1997). In this scenario, a typical cluster is therefore expected to travel about $2\,h^{-1}$ Mpc up to the present epoch.

The average nearest-neighbour distance $r_{\rm nn}$ is given by

$$r_{\rm nn}^3 = \frac{3}{4\pi} < n_{\rm c} >^{-1} -3 \int_0^{r_{\rm nn}} \xi_{\rm cc}(r) r^2 {\rm d}r, \qquad (13)$$

where $n_{\rm c}$ is the cluster number density and $\xi_{\rm cc}(r)$ the cluster–cluster correlation function, which is observed to be a power law,

$$\xi_{\rm cc} = (r/r_{\rm cc})^{-1.8}, \qquad (14)$$

where the cluster–cluster correlation length $r_{\rm cc}$ is around $18\,h^{-1}$ Mpc (e.g. Peebles 1993). This means that for a mean cluster density of $8.6 \times 10^{-6}\,h^3\,\mathrm{Mpc}^{-3}$ (Mazure et al. 1996) the average nearest-neighbour distance is $20\,h^{-1}$. Therefore one does not expect two Abell-sized peaks to merge within a Hubble time. This is confirmed from an estimate of the probability of finding, at the present epoch, a cluster–cluster separation of less than $2\,h^{-1}$ Mpc (the typical distance travelled up to the present epoch), using the observed $\xi_{\rm cc}$. We derive a probability of 2 per cent, i.e. about two clusters in the catalogue. This means that we expect one close pair that could merge if moving towards each other. Note that we derived this probability for the *present epoch*, and that it must have been smaller at earlier times.

3.1.2 *Late cluster formation from subcluster mergers*

A cluster can also form from two or more subclusters, which are more abundant than rich clusters themselves, through a merger at late times. One needs somewhat special initial conditions for clusters to form this way, because the separation of such subclusters in the initial conditions cannot be too large or they would remain distinct, despite non-linear clustering. It is more likely that late merging involves unequal mass clumps, where the main clump would not have been massive enough to become a rich cluster were it not for the 'secondary infall' of nearby subclumps.

Such cases of cluster formation are unavoidable in hierarchical formation scenarios. The direct way to access their significance is to look at a large-scale simulation of volume equal to that of the whole catalogue (see e.g. van Haarlem, Frenk & White 1997). We will discuss this question in Section 5.

3.2 The catalogue defining quantity

3.2.1 *Total cluster mass*

As a single cluster model is defined by its local initial conditions, a model catalogue of clusters is defined by the statistics of peaks in the initial cosmological density field. We will construct a model catalogue under the assumption that a certain subset of peaks in the initial smoothed density field will each produce a cluster at the present epoch. In order to build a catalogue that mimics a sample of rich Abell clusters, we need to establish a relation between peak characteristics in the initial density field and the criteria used by Abell at the present epoch. Unfortunately, his criteria are not easily translated into the formalism of Gaussian random fields. Therefore we are left with a choice for a sensible physical or observable quantity that does provide a relation between initial and final conditions.

Physical parameters are very easily obtained for model clusters since the data are clean and unprojected. Far less choice is available in the observations, which provide projected, biased and generally less clean data. We will focus mainly on galaxies, not on dark matter, although it is probably the dominant mass component. The dark matter distribution can only be obtained through indirect measurements like the X-ray brightness distribution (e.g. Sarazin 1986; Forman & Jones 1990) or the effects of gravitational lensing (e.g. Fort & Mellier 1994).

Which physically significant quantities can we derive from the galaxy distribution? The quantities that are best observable are the two spatial coordinates on the sky, which can be obtained very accurately. Nevertheless, the position of the *centre* of the projected galaxy distribution, required for most of the derived properties, is relatively difficult to obtain. Global properties derived from the projected distribution are the total number of galaxies within a certain radius, which was used by Abell for his richness criterion, the shape (ellipticity) and orientation of the distribution (e.g. Binggeli 1982; Rhee, van Haarlem & Katgert 1989; de Theije, Katgert & van Kampen 1995), measures of substructure (reviewed by Beers 1992), the azimuthally averaged surface density profile (e.g. Beers & Tonry 1986; Rhee et al. 1989), and the projected mean harmonic radius, which is often used for mass estimates.

With regard to velocity information, only line-of-sight components are readily observable. Apart from the distance to the cluster, the most useful global quantity derived from these is the central line-of-sight velocity dispersion. This would be a good candidate for the catalogue defining quantity, except that it is a non-linear quantity, and can therefore not be used in the the Hoffman–Ribak method of constrained random fields. The total velocity dispersion is correlated with the total mass of the cluster, which is also one of its most basic physical properties. We demonstrate below that the final cluster mass can be estimated fairly reliably from the initial

peak parameters, and we therefore choose *the expected final mass* to be the defining quantity of a catalogue of cluster models.

Although in principle mass estimates can be obtained from the combination of the projected galaxy distribution and the central line-of-sight velocity dispersion, reliable total masses are not known for most Abell clusters. All we can do is to hope that richness correlates sufficiently well with mass, and that those peaks in the Gaussian smoothed initial density field that have the largest expected final mass are the progenitors of the rich Abell clusters.

3.2.2 *Smoothing scales*

Before we discuss the relation between the final mass of a cluster and the parameters of the initial density peak, we need to choose the smoothing scale that is appropriate for clusters. We base our choice on an argument involving the cluster mass. In general, the total mass for a given spherical window function $W(r, R_W)$ is given by

$$M_W = V_W \bar{\rho}[1 + \delta_W(0, R_W)], \tag{15}$$

where the volume $V_W(R_W)$ of the window function is

$$V_W(R_W) = 4\pi \int W(r, R_W) r^2 \mathrm{d}r. \tag{16}$$

The (observational) compactness criterion implies that Abell clusters are defined using a top-hat window

$$W_{\mathrm{TH}}(r, R_{\mathrm{TH}}) = \begin{cases} 1 & \text{for } r \le R_{\mathrm{TH}} \\ 0 & \text{for } r > R_{\mathrm{TH}}, \end{cases} \tag{17}$$

with $R_{\mathrm{TH}} = 1.5\,h^{-1}$ Mpc, the Abell radius. The mass enclosed by a sphere of this size is about $4 \times 10^{12}\,h^{-1}\mathrm{M}_\odot$ for the background universe (for $\Omega_0 = 1$). A 'typical' cluster has a mass of about $3 \times 10^{14}\Omega_0\,h^{-1}\mathrm{M}_\odot$ (e.g. Peebles 1993), so the cluster mass originates from a volume 70 times larger. This corresponds to a top-hat filter of about $6\,h^{-1}$ Mpc in the initial density field. The corresponding Gaussian smoothing radius R_{G} is found by setting the enclosed volumes equal for both windows (BBKS). This gives $R_{\mathrm{G}} \approx 4\,h^{-1}$ Mpc, as the ratio between the two radii is $(4/18\pi)^{1/6} \approx 2/3$.

3.3 Cluster mass threshold

In this section we estimate the total mass of a cluster given the properties of the peak in the initial smoothed density field from which it originates.

3.3.1 *Peak amplitude threshold: 'linear' mass*

In the linear approximation the final total mass is proportional to the peak amplitude in the initial smoothed density field. We define model clusters using the Gaussian window function with radius $R_{\mathrm{G}} = 4\,h^{-1}$ Mpc. If we take the same value for the present density field as for the initial field, and assume the field to evolve linearly, the total mass is determined exactly by the peak amplitude ν alone and given by

$$M_{\mathrm{G}}(\nu) = (2\pi R_{\mathrm{G}}^2)^{3/2} \bar{\rho}[1 + \nu\sigma_0], \tag{18}$$

because the relation $\delta_{\mathrm{G}} = \nu\sigma_0$ remains valid throughout the evolution. In that case the linear relation $\sigma_0(t) = a\sigma_0(t_0)$ describes the time dependence, where t_0 is the present epoch and a is the cosmological expansion factor [with $a(t_0) \equiv 1$] and ν remains constant.

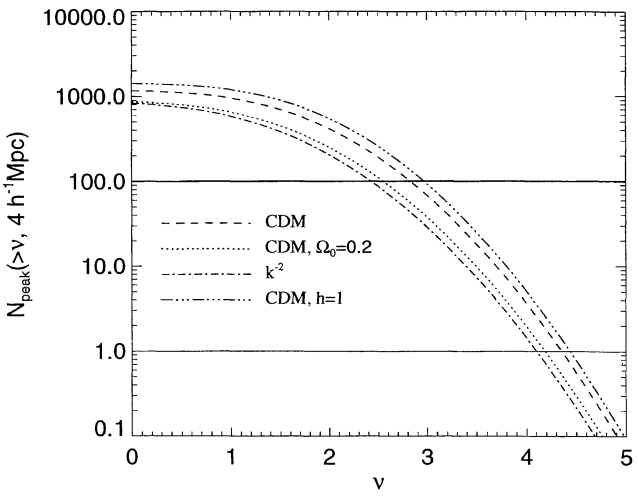

Figure 1. Expected range in ν for 100 clusters in a volume of $10^7\,h^{-3}\mathrm{Mpc}^3$. All scenarios are for $\Omega_0 = 1$ and $h = 1/2$ unless stated otherwise.

We could then define a model catalogue by setting a threshold on ν, given by $N_{\mathrm{peak}}(> \nu, R_{\mathrm{G}}) = \mathcal{N}$, where $\mathcal{N}$ is again the total number of clusters in the catalogue with volume V. The threshold ν_{min}, which is formally an expectation value, should then correspond to the threshold in the defining quantity of the observational catalogue. In the case of rich Abell cluster catalogues this is the value of 50 for the Abell richness measure.

We can calculate the expected maximum peak amplitude ν_{max} in that volume from $N_{\mathrm{peak}}(> \nu, R_{\mathrm{G}}) = 1$. We illustrate this in Fig. 1 for several cosmological parameters and fluctuation spectra, smoothed on the cluster scale of $4\,h^{-1}$ Mpc. The heavy horizontal line determines the threshold ν_{min} for the choices $\mathcal{N} = 100$ and $V = 10^7\,h^{-3}\,\mathrm{Mpc}^3$, which are typical values for a cluster catalogue. The thin line indicates the expected ν_{max}, which is supposed to represent the largest cluster in V. Of course, larger peak amplitudes will occur in larger volumes.

Note that it is not sufficient just to threshold the 'linear' mass in the initial smoothed density field in order to guarantee the presence of a rich Abell cluster in the final density field. Because richness and mass are not likely to be perfectly correlated, there are peaks with a small 'linear' mass that produce rich clusters (and vice versa), making the catalogue selected by mass incomplete in richness. Since we cannot use the constrained random field method to constrain Abell richness, we have to tighten the correlation of initial peak parameters and final richness as much as possible. One way of doing that is to go from 'linear' mass, as given by the peak parameter ν, to a 'non-linear' mass.

Several non-linear effects come into play that determine the final mass of a galaxy cluster. The collapse time of a cluster depends not only on the peak amplitude of the initial overdensity, but also on its curvature x (van Haarlem & van de Weygaert 1993), on tidal effects from the environment, on the presence of substructure (Cavaliere et al. 1986) which can result in pre-virialization (Peebles 1990) and thus a slower collapse, and on the amount of shear in the initial velocity field (Bertschinger & Jain 1994). If the non-linear evolution of these properties is significant, the peak in the final density field may not coincide with the peak in the initial density field, not even for the field smoothed on the scale of $4\,h^{-1}$ Mpc.

We show below that the effect of the peak parameter x, which is related to the slope of the initial density profile, can be taken into

account, and can be used for a better approximation of the expected final cluster mass.

3.3.2 *Expected 'non-linear' cluster mass*

Cluster peaks in the density field smoothed at $4\,h^{-1}$ become non-linear. We find that δ_G can grow up to about 10 for the structure formation scenario that we have adopted, namely standard CDM for $\Omega_0 = 1$. Therefore ν becomes a function of time, and $\sigma_0(t)$, the rms fluctuation on scale R_G, will become quasi-non-linear. Note that $\nu(t)$ is a *locally* evolving quantity whereas $\sigma_0(t)$ evolves *globally*, so the latter will grow less rapidly than the former which describes the evolution of an individual peak. Furthermore, the effect of the initial curvature of the peak, as set by x, will become important. Van Haarlem & van de Weygaert (1993) found that the final cluster mass is larger for smaller x, i.e. for more extended initial configurations. We can try to incorporate these effects by looking at the possible evolution of the initial smoothed radial density constrast profile around a peak with given ν and x. The initial density profile is given by BBKS:

$$\delta_G(r) = \frac{\nu - \gamma x}{\sigma_0(1-\gamma^2)}\xi_G(r) + \frac{\nu - x/\gamma}{3\sigma_0(1-\gamma^2)}\nabla^2\xi_G(r), \qquad (19)$$

where $\xi_G(r)$ is the smoothed density autocorrelation function. Remember that all quantities depend on the smoothing radius R_G. In the linear regime $\delta_G(0) = \nu\sigma_0$, and $\xi_G(0) = \sigma_0^2$ by definition, so we can derive that initially

$$\nabla^2\xi_G(0) = -3\gamma^2\sigma_0^2. \qquad (20)$$

For small r both terms in equation (19) initially contribute roughly equally to the density profile for most cosmological scenarios. For large r the $\nabla^2\xi_G(r)$ term quickly vanishes and the profile is well approximated by just the first term. For small r the profile is well approximated by a second-order Taylor expansion around the peak, as in equation (8). Here we will just consider the spherically averaged profile, which means that the contribution from $\mathcal{F}_5$ cancels.

We now substitute non-linear quantities in equation (19) in order to obtain an evolution equation for the (smoothed) density profile. If the evolution would remain linear, ν, x and γ would remain constant, and σ_0 would evolve linearly with the cosmological expansion factor a. The autocorrelation function scales with σ_0^2, and so do its derivatives because the slope of $\xi_G(r,t)$ remains constant, so that the $x(t)$ terms in equation (19) cancel. How does non-linear evolution change this behaviour? The rms fluctuation σ_0 at $4\,h^{-1}$ Mpc Gaussian smoothing is about unity at the present epoch for most cosmological scenarios, and its time evolution therefore becomes non-linear. Yet, ν and x become non-linear more quickly because they are defined for *peaks*. Now, if $\xi_G(0,t)$ would grow at the same rate as $\nabla^2\xi_G(0,t)$, both $x(t)$ terms in equation (19) still cancel and we would only see non-linear behaviour in $\delta_G(0,t)$ due to $\nu(t)\sigma_0(t)$.

However, non-linear evolution causes the slope of $\xi_G(t)$ to become steeper, which changes the evolution of $\nabla^2\xi_G(0,t)$ with respect to $\xi_G(0,t)$. So the second term of equation (19) evolves somewhat faster than the first because of the increasing slope of $\xi_G(t)$. This might seem unimportant, but the x terms, which grow relatively fast, no longer cancel. This is demonstrated in more detail in Appendix B. If we use a first-order approximation for $\nu(t)$ and $x(t)$, we find a relation of the form

$$M_G(\nu_i, x_i, t_0) = (2\pi R_G^2)^{3/2}\bar{\rho}\left[1 + (c_0 + c_1\nu_i + c_2 x_i)\sigma_0\right], \qquad (21)$$

where the c_j are constants at epoch t_0, and ν_i and x_i are the initial peak parameters. If the field and its derivatives just evolved linearly, $c_0 = 0$, $c_1 = 1$ and $c_2 = 0$. Non-linear evolution results in a non-vanishing c_0, where the sign depends on the exact balance of the constants in the linear approximations for $\nu(t)$ and $x(t)$, $c_1 > 1$ and $c_2 < 0$. Again, see Appendix B for more details.

3.4 Sampling constraints above the expected final mass threshold

We have defined a model catalogue of $\mathcal{N}$ clusters as follows. We first draw a large number of (ν, x) sets for a sufficiently small ν_{min} found from (10), and then apply the $M(\nu, x)$ threshold so that $\mathcal{N}$ models remain. However, the values for the parameters c_j in (21) are unknown a priori because no model has actually been built yet. So one first needs to choose a few representative models from the $\mathcal{N}$ models selected for small deviations of the linear values $c_0 = 0$, $c_1 = 1$ and $c_2 = 0$, for example $c_0 = 0$, $c_1 = 1.5$ and $c_2 = -0.5$ (see Appendix B). After simulating these models, one can fit the c_j to the measured total mass in these simulations within a comoving sphere of radius $6\,h^{-1}$ Mpc around the centre. This provides a new total mass threshold, which can then be used to adjust the selection of the $\mathcal{N}$ models from the large ensemble. If the initial sets of constraints are chosen sensibly, along the lines described in Appendix B, and not too close to the initial threshold, most of the trial models will remain in the catalogue. This iterative procedure is repeated until all $\mathcal{N}$ models are run and the c_j have converged to values that define the final catalogue.

Each model is defined not only by ν and x, but also by its shape parameters. These do not form part of the predicted mass relation that defines the catalogue, but are still chosen so that we can also study the dependence of the final cluster models on initial shape. Given ν and x, the parameters a_{12} and a_{13} are drawn simultaneously from the probability distribution in equation (12). The orientation is drawn from an isotropic distribution, as the phases of the Gaussian density field are random. The actual orientation is important only to determine the projected shape of the object on the sky.

Deviates are generated using the rejection method (e.g. Press et al. 1988). To draw ν, a convenient comparison function is found by substituting ν^3 for $\mathcal{F}_1(\gamma,\nu)$ in (10). The comparison function needed for drawing x is obtained by replacing $\mathcal{F}_2(x)$ in (11) with x^3 (see also Appendix A). For the joint sampling of the shape parameters the comparison function is simply a constant that is larger than the maximum of (12). Deviates of a_{12} and a_{13} are (uniformly) generated between 1 (by definition) and some maximum for which the probability is really small already, for which we took 10. This may not be the most efficient algorithm, but it works sufficiently well for our purposes given that (12) is rather complicated.

4 CONSTRUCTION OF A MODEL CATALOGUE

4.1 Choice of the cosmological scenario

We build a catalogue of model clusters within the standard CDM scenario, i.e. we adopt the $\Omega_0 = 1$ adiabatic CDM initial fluctuation spectrum (Davis et al. 1985). This scenario has the advantage that the amplitude of the initial density fluctuation spectrum, denoted by σ_8, can be fixed *after* having run the models, by comparing the properties of the galaxy distribution to observational data. Fig. 2 shows the unbiased standard Ω_0=1 adiabatic CDM spectrum as

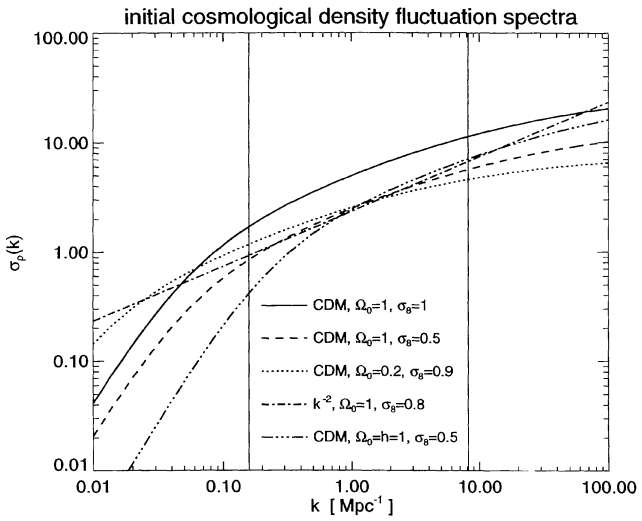

Figure 2. Selected cosmological density fluctuation spectra, plotted in the form of $\sigma_\rho(k) \equiv (4\pi k^3)^{1/2}|\delta_k|$.

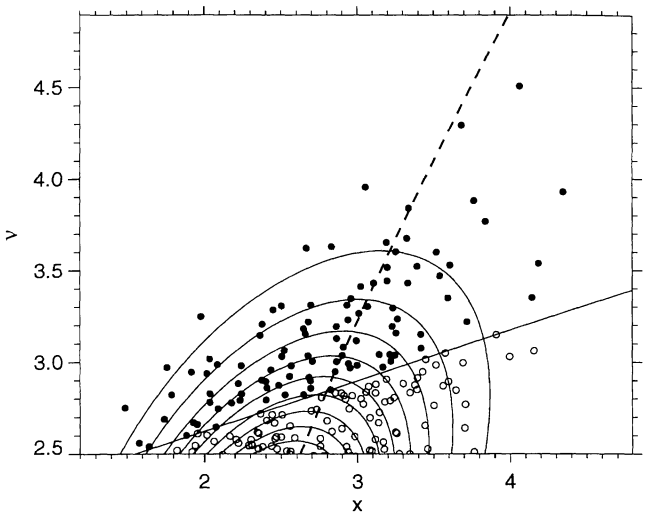

Figure 3. Selection of models on expected total mass within $4\,h^{-1}$ Mpc at $\sigma_8 = 0.46$ from 200 generated (ν, x) sets. The thick solid line selects the 99 cluster models that have been simulated (filled symbols) and form the model catalogue. The open symbols represent (ν, x) sets that have not been simulated. The dashed line denotes the expectation value of x for a given ν (equation A10). The curved solid lines show probability density contours of $P(x|\nu)$ (see equation 11 in Section 2.3, and Section 4.4.3).

given by Davis et al. (1985) in the form of $\sigma_\rho(k) \equiv (4\pi k^3)^{1/2}|\delta_k|$, which is the rms contribution to the density fluctuation field. If σ_8 is less than the observed value of unity, this decreases the amplitude of the spectrum by a factor of σ_8 for all wavenumbers. The vertical lines indicate the range of wavenumbers included in our simulations. The dashed line shows the $\Omega_0 = 1$ CDM scenario, which we adopted here, for $\sigma_8 = 0.5$, a value that should not be far off from the one that best fits the observational data. We also show a few other scenarios, with values for σ_8 such that the spectrum is not too different from the Ω_0=1 CDM spectrum within the range of wavenumbers present in the simulations. Of course they will be significantly different at larger and/or smaller wavenumbers: e.g. the featureless power-law spectrum with index -2 (dashed-dotted line) has substantially more power on both larger and smaller scales.

Another important reason for choosing the $\Omega_0 = 1$ CDM scenario is that it allows comparison to published results, notably those of Frenk et al. (1990) and White et al. (1993). Furthermore, CDM does not have many more problems than most of the competing scenarios (e.g. Davis et al. 1992; Ostriker 1993; Bertschinger 1993), and explains many observations, although not in its original unbiased form.

4.2 The target observational catalogue

As we want our sample of model clusters to be fair, i.e. statistically complete for a specific volume, its observational counterpart also needs to be as complete as possible. A sample that was constructed with this requirement in mind is one for which data have been obtained in the context of an ESO Key-programme. This sample of rich Abell clusters, the ENACS sample for short, is based on a subsample of 128 $R \geq 1$ clusters selected from both the Abell and the ACO catalogues in a cone of about 2.6 sr out to $z = 0.1$ (Katgert et al. 1996). It occupies a single closed volume that is well away from the Zone of Avoidance.

We choose a $z < 0.08$ subsample of the ENACS sample (which goes out to $z = 0.10$) as the observational catalogue to be mimicked, because it is the most complete subsample selectable that is large enough for statistical purposes (Mazure et al. 1996). It contains 76 clusters in a volume of $9.2 \times 10^6\,h^{-3}$ Mpc^3. After corrections for completeness, Mazure et al. (1996) found that there should have been 79 in that volume, implying a comoving number density of $8.6 \pm 0.6 \times 10^{-6}\,h^3$ Mpc^{-3} for $R \geq 1$ Abell clusters.

4.3 The $\Omega_0 = 1$ CDM model catalogue

We now define a catalogue of cluster models that mimics the ENACS sample. For the volume of the catalogue we take $10^7\,h^{-3}$ Mpc^3. This means that according to the number density found by Mazure et al. (1996) there should be 86 rich Abell clusters in this volume, of which 83 define an ACO-like observational sample. To come close to the target of 83 cluster models, we built a simulation set that consists of 99 cluster models. This redundancy somewhat alleviates completeness problems associated with thresholding on the predicted final mass, which arises from scatter in the relation of richness versus final mass (Sections 5 and 6).

We generated the simulation set for the CDM $\Omega_0 = 1$ scenario using the technique described in Section 3.4. We used 200 sets of (ν, x) from which the desired 99 will be selected. The 200 sets are shown in Fig. 3 in the form of a scatter plot of ν versus x. The solid line depicts the fitted $M(\nu, x)$ threshold found after the last iteration that separates the 200 sets in 99 candidate catalogue entries (filled circles), and 101 models that are below the expected final mass threshold (open circles).

The present epoch for the adopted scenario, as obtained from comparison to the galaxy autocorrelation function, is most likely to be $\sigma_8 = 0.4$–0.5 (Davis et al. 1985; van Kampen 1994). Note that the *COBE* measurements require a rather different range of values, as is discussed in Section 8. Our simulations have data saved for $\sigma_8 = \{0.25, 0.31, 0.36, 0.41, 0.46, 0.51, 0.56, 0.61, ...\}$, so we took $\sigma_8 = 0.46$ for the purpose of fitting the c_j in equation (21). In Fig. 4 we show two fits: one for the linear mass estimate based on a linear function of ν alone, and one for $M_G(\nu, x)$, given by (21). The first clearly has a larger spread than $M_G(\nu, x)$; the variance of the residuals of the fit, $\sigma^2_{\rm res}$, is twice as large. We found $c_0 = -4.3$, $c_1 = 4.9 \pm 0.3$ and $c_2 = -1.3 \pm 0.2$, with no correlation between the residual scatter in the $M_G(\nu, x)$ relation and the shape parameters

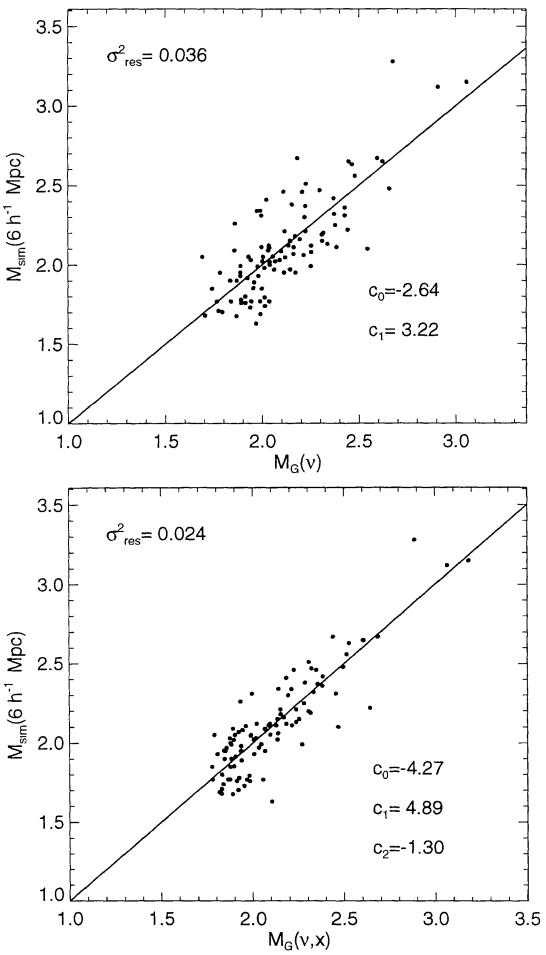

Figure 4. Fits for the expected masses $M_G(\nu)$ (top panel) and $M_G(\nu,x)$ (bottom panel) to the total mass found within $6\,h^{-1}$ Mpc in the simulations for $\sigma_8 = 0.46$.

a_{12} and a_{13}. We suspect that the remaining scatter is mostly due to the non-linear evolution and the difference between Gaussian and top-hat smoothing, as well as to differences in the initial *velocity* field, notably the shear (Bertschinger & Jain 1994; van de Weygaert & Babul 1994). This latter effect will be investigated in another paper.

4.4 Testing the construction technique

With the aid of a conventional large-scale, low-resolution simulation we have tested the assumptions made in the method we use to construct a catalogue of galaxy cluster models. What we test here is whether the final galaxy clusters (i.e. the final peaks) indeed have a one-to-one correspondence to initial density peaks, and how well the expected mass estimator M_G performs.

4.4.1 Description of the large-scale simulation

We evolved a patch of universe in a cubic volume of $(256\,h^{-1}\,\mathrm{Mpc})^3$ using a P^{3}M code (Bertschinger & Gelb 1991), and saved particle positions and velocities at the same epochs as for the individual cluster models. A more detailed description of this simulation can be found in de Theije et al. (1997). Given a cluster density for rich Abell clusters of $8.6\times10^{-6}\,h^3\,\mathrm{Mpc}^{-3}$, we expect 144 of such clusters in the chosen volume. We employed 128^3 particles, so that each particle has a mass of $4.4\times10^{12}\,\mathrm{M}_\odot$. This provides sufficient resolution to identify rich clusters and trace them back to the initial density field. Note that the resolution is significantly lower than for the individual simulations, so that properties obtained for the clusters in the single large-scale simulation will be much noisier than for their individually simulated counterparts.

After running the simulation, the $4\,h^{-1}$ Mpc Gaussian smoothed density field and its second derivative were obtained on a 128^3 lattice for each output time. Peaks were identified on these lattices by searching for those cells that have a larger density than their 26 neighbouring cells. We then use second-order polynomial fitting to find the off-grid positions of the peaks, and the actual peak parameters ν and x at those positions. At the final epoch these peaks correspond to N-body particle groups which should be the final-epoch rich galaxy clusters.

4.4.2 Tracing cluster peaks

We first look at the link between initial and final density peaks by linking the former to the latter using our single large-scale simulation. We link the peaks at different epochs by finding the smallest spatial separation between the respective positions in consecutive time-slices.

We found that none of the initial peaks has merged during the evolution up to $\sigma_8 = 0.6$, well above the value of $\sigma_8 = 0.4 - 0.5$ found previously (see Section 4.3). This justifies our use of initial peaks to construct a cluster catalogue, and confirms the estimate in Section 3.1.1.

4.4.3 Distribution and evolution of ν and x

We examined the (ν,x) distribution found in the single large-scale simulation, which is complete by definition, in order to test the generation of deviates for the peak parameters that define the catalogue of high-resolution models. We indeed find a similar distribution for (ν,x) to that for the models making up the model catalogue (see Fig. 3), so the sampling of the latter is representative (see also van Kampen 1994).

For the catalogue of high-resolution models we show, in Fig. 5(a), for each simulation, the relation between the initial peak parameters, ν_i and x_i, and their final values, ν_0 and x_0, at the present epoch, t_0. Filled symbols are for ν, whereas open ones are for x. The solid line indicates linear evolution, for which the peak parameters remain constant (i.e. $\nu_0 = \nu_i$ and $x_0 = x_i$). Whereas the evolution of ν is mildly non-linear, the curvature x evolves well into the non-linear regime, but with a large scatter.

In Fig. 5(b) we show the results of three simulations by van Haarlem & van de Weygaert (1993) (open triangles), who give initial–final relations for x for peaks with $\nu_i = 3$ (note that they have their final epoch at $\sigma_8 = 1$, whereas we have $\sigma_8 = 0.46$). The dotted line is a fit to those points. We also include those models in Fig. 5(a) with $|\nu_i - 3| < 0.05$, as those are closest to the $\nu_i = 3$ peaks of van Haarlem & van de Weygaert (1993). A fit to their data is shown as a dashed line, and it can be seen that the slope of the initial–final relationship for x at a fixed ν_i evolves from unity for the initial epoch (solid line), through zero near $\sigma_8 \approx 0.4$, to a steepening negative slope for larger σ_8.

The evolution of the peak parameters of the high-resolution models might not be similar to that in the low-resolution simulation if the size of the high-resolution simulation spheres was not chosen large enough, as that would artificially reduce large-scale tidal

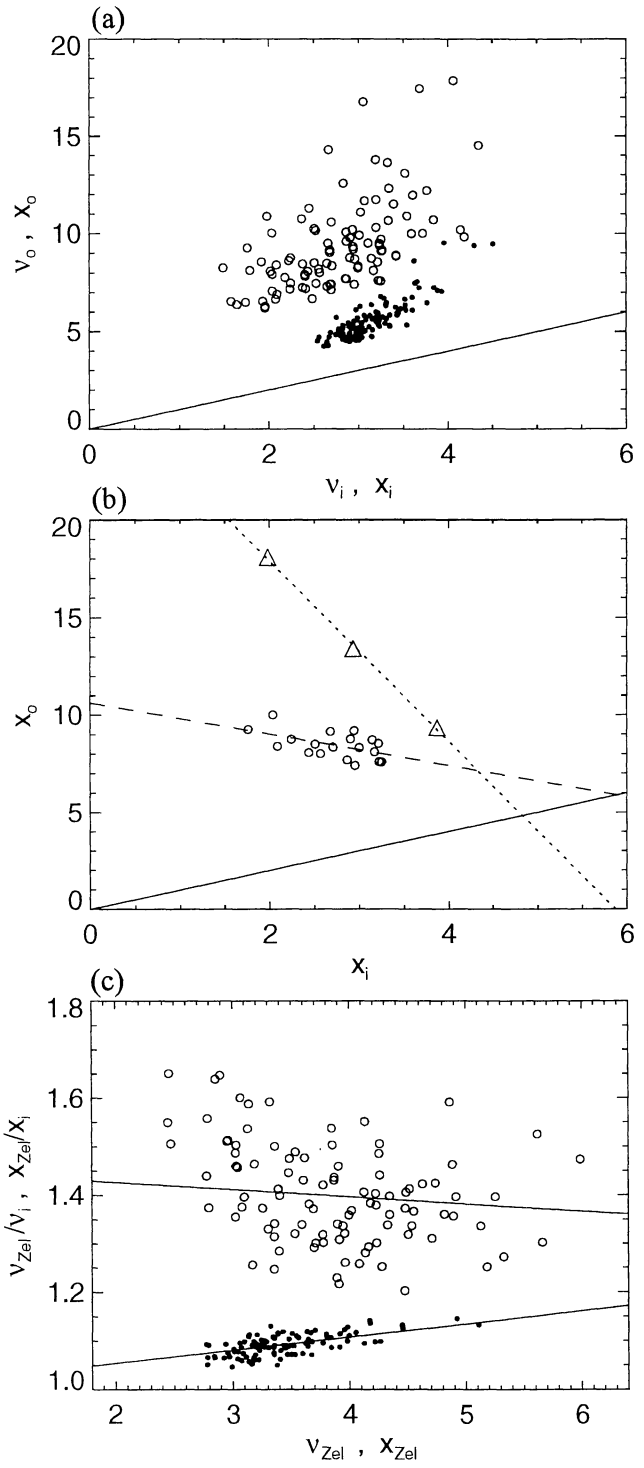

Figure 5. (a) Relations between linear and final values of ν (closed symbols) and x (open symbols) from the model catalogue, for $\sigma_8 = 0.46$. (b) Same as (a), but for those (x_i, x_0) that have $|\nu_i - 3| < 0.05$. The dashed line is a fit to these points. Open triangles are for simulations by van Haarlem & van de Weygaert (1993), for $\sigma_8 = 1.0$. A fit to these three points, which all have $\nu_i = 3$, is shown as a dotted line. Notice the evolution from solid line ($\sigma_8 = 0$), via dashed line ($\sigma_8 = 0.46$), to dotted line ($\sigma_8 = 1.0$). (c) ν_{Zel} and x_{Zel} versus the ratios ν_{Zel}/ν_i and x_{Zel}/x_i, with solid lines showing fits to these, which were used for Figs 6(b) and 7. See text for more details.

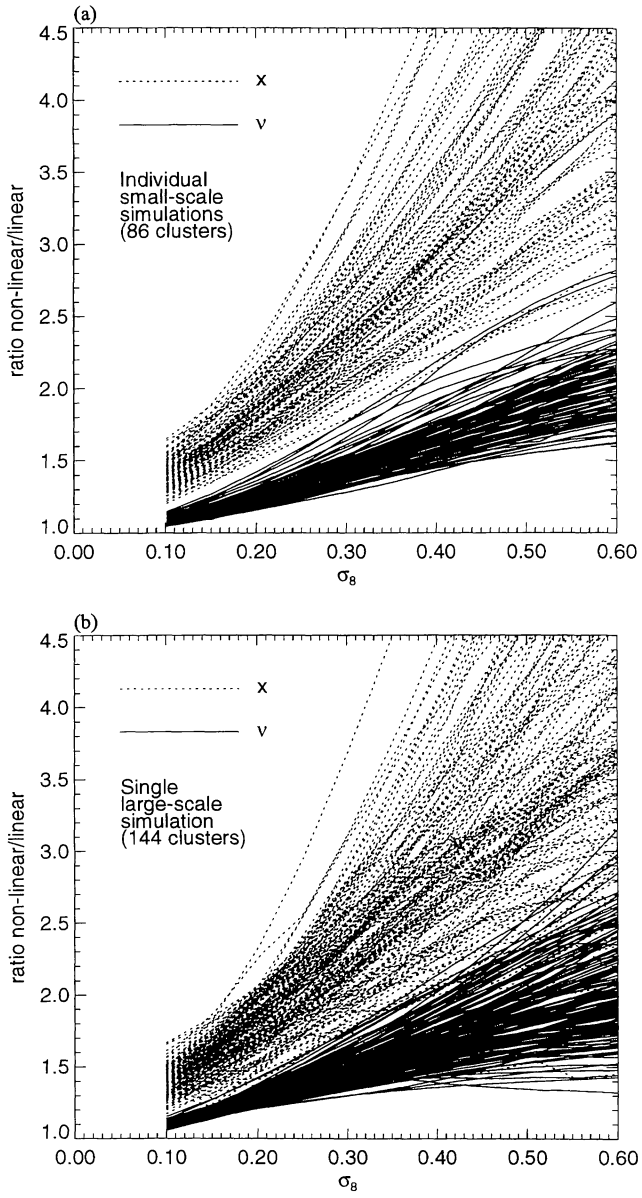

Figure 6. (a) Evolution of the non-linear to linear ratio for the peak amplitude, ν, and the peak curvature, x, for all catalogue models. (b) Same as (a), but for the peaks found in the single large-scale simulation.

fields. We show the evolution with σ_8 of the ratio of final to initial values of ν and x for each high-resolution cluster model in Fig. 6(a).

A similar plot for the low-resolution simulation cannot be readily made, as we do not know the initial values of the peak parameters, but only ν_{Zel} and x_{Zel}, the values at the end of the Zel'dovich approximation which was used to calculate the linear evolution and the first part of the non-linear evolution. In Fig. 5(c) we show the relation between the values of the peak parameters at the start of the N-body integration (ν_{Zel} and x_{Zel}, and the ratios ν_{Zel}/ν_i and x_{Zel}/x_i). This figure shows that at the start of the N-body integration the amplitude ν of the peak is, on average, 1.1 times larger than the initial value, while the curvature is on average 1.4 times larger than its initial counterpart. Using Fig. 5(c), we have produced Fig. 6(b), which shows the evolution of the peak parameters for the low-resolution simulation. The initial scatter in Fig. 6(b) was put in by

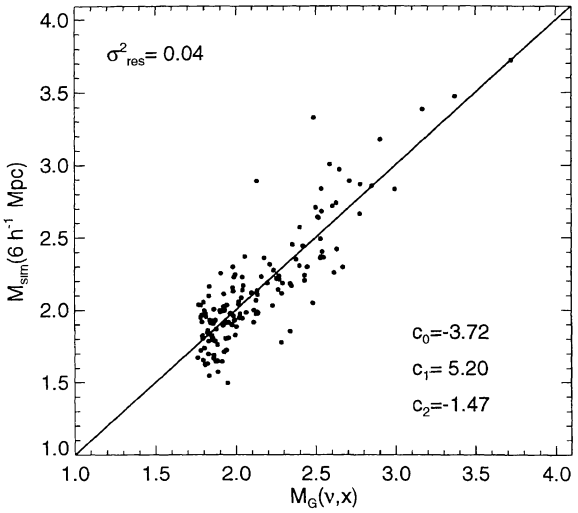

Figure 7. Fit for the expected mass M_G of peaks in the single large-scale simulation to the total mass found within $6\,h^{-1}$ Mpc around these peaks.

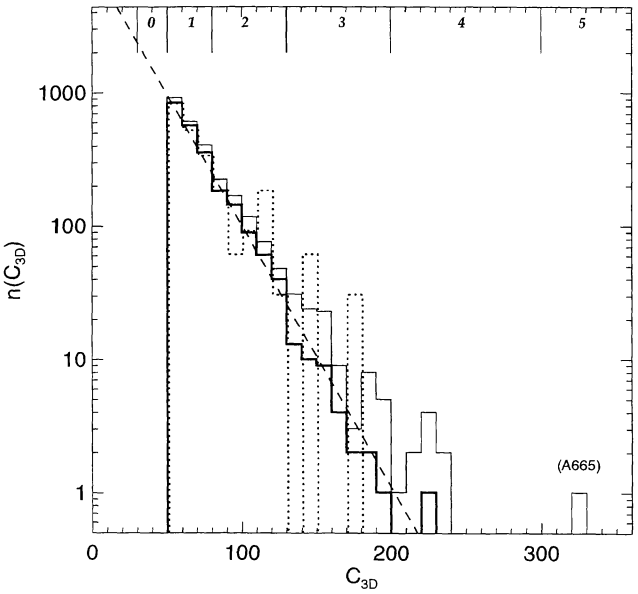

Figure 8. Richness distribution for the ACO catalogue (thin line), its $z < 0.08$ subsample (thick line), and the rescaled ENACS sample (dotted line), also for $z < 0.08$. Numbers at the top indicate the traditional Abell richness classes.

hand; it reflects the scatter in the fits from Fig. 5(c) and was sampled from a Gaussian distribution. It appears that the individual simulations perform well; the peak parameters from the single large-scale simulation show a very similar evolution with time to those for the collection of high-resolution models, although they are somewhat noisier.

4.4.4 *Fairness of the expected mass relation*

Using the centre of mass of the actual N-body particle group near the density peaks in the low-resolution simulation, we obtain the mass M_{sim} within a radius of 6.0 h^{-1} Mpc (i.e. the top-hat mass-scale that was used to define the catalogue). We expect 144 rich Abell clusters in this volume (see above), so we select a threshold for M_G that produces that number.

To test the $M_G - M_{sim}$ catalogue defining relation, we trace back clusters found at the present epoch to peaks in the smoothed initial density field. In Fig. 7 we plotted the final mass M_{sim} against the predicted final mass M_G that is calculated from the initial peak parameters according to equation (21), where we again have used the fits of Fig. 5(b) to transform from the 'Zel'dovich' peak parameters to the initial ones.

The relation in Fig. 7 compares quite well to the relation for the single, high-resolution, cluster models, as shown in Fig. 4(b). Although the constants c_j are slightly different, the ratio c_1/c_2, which determines the slope of the final mass threshold, is almost the same (3.54 as compared to 3.76).

5 COMPLETENESS

5.1 Completeness of the ENACS sample

The ENACS sample was constructed to be complete in richness (Katgert et al. 1996) by selecting a subsample from the ACO catalogue within a particular volume (see Section 4.3). However, from comparison to the Edinburgh–Durham Cluster Catalogue (EDCC from here on), Mazure et al. (1996) estimated that the full $R \geq 1$ ($C_{ACO} \geq 50$) ACO catalogue is 94 per cent complete out to $z = 0.08$. We have to accept this incompleteness, since the ACO catalogue is still the largest observational sample available. In the following we therefore assess completeness with respect to the full ACO cluster catalogue, keeping in mind that this catalogue is only 94 per cent complete.

We see from Fig. 8 that the ENACS sample, as a subsample of the ACO catalogue, is 100 per cent complete out to $z = 0.08$. Plotted is the distribution of richness C_{ACO} of all ACO clusters (thin solid line), ACO clusters out to $z = 0.08$ (thick solid line), and ENACS clusters out to the same redshift (dotted line), where the latter is rescaled to the ACO volume. The distributions of the latter two are clearly statistically the same, indicating that the ENACS catalogue is a representative subsample of the ACO catalogue. An exponential function of the form $n(R) = a\ e^{-bR}$ has been fitted to the richness distribution of the $z < 0.08$ ACO clusters for use in Section 6, and is shown in Fig. 8 as a dashed line.

5.1.1 *ACO and 3D richness*

Both Abell (1958) and ACO (1990) have corrected their richness determinations for foreground and background galaxies, by subtracting a 'field correction' term from the raw number count. Since we model clusters individually, our simulation volume does not contain any obvious foreground or background galaxies. Common observational practice is to use a 'gapper' criterion to remove such galaxies, i.e. eliminate all galaxies that are on the non-cluster side of a gap in the line-of-sight velocity distribution for all galaxies observed towards the cluster. Note that we will remove 'interlopers', non-cluster members near the turn-around radius that are not eliminated by the 'gapper' criterion, when we calculate line-of-sight velocity dispersions in Section 7.1.

For our model clusters, with their limited simulation volume, we did not find any gaps of 1000 km s^{-1}, which is the value used by Katgert et al. (1996) for the ENACS sample. So we indeed do model just what an observer would call the 'main' system, and we can simply apply Abell's (1958) criteria to each of our simulations,

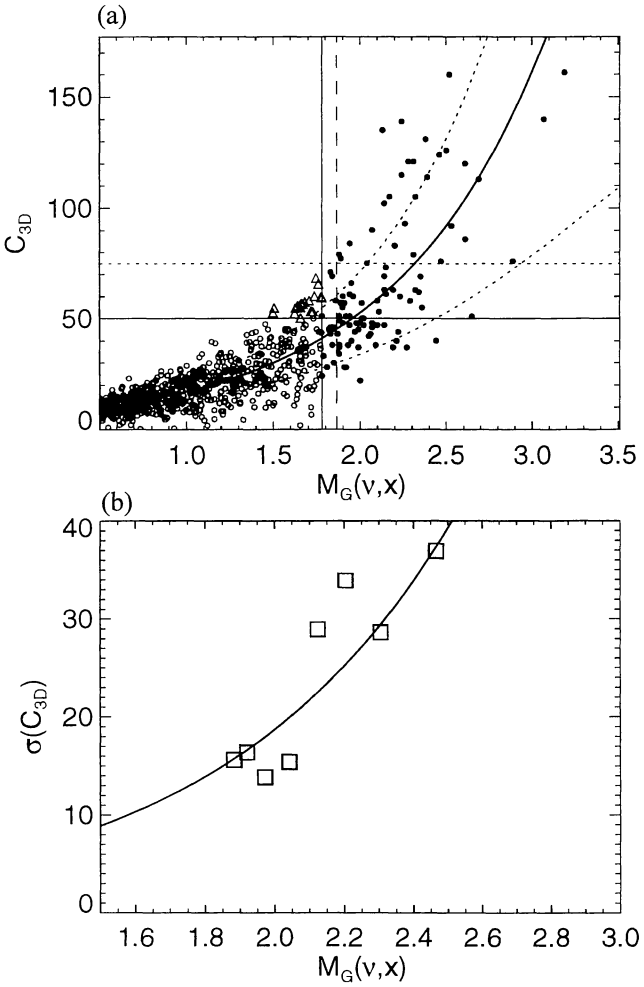

Figure 9. (a) 3D richness C_{3D} versus predicted final cluster mass M_G for our simulated clusters at $\sigma_8 = 0.46$ (filled symbols), and peaks sampled below the M_G threshold for 99 clusters (open symbols and triangles). The dashed line indicates the M_G threshold for 86 clusters. The horizontal lines indicate the $C_{3D} = 50$ and $C_{3D} = 75$ richness limits. The solid curve is an exponential fit to the points above the M_G threshold (the vertical solid line), with the dotted lines showing the 1σ deviation around that fit, as obtained from (b). (b) Variance in richness as a function of 3D richness, as obtained from the *simulated* models, i.e. the filled symbols displayed in (a). The solid curve is again an exponential fit to these points, and was used to plot the dotted lines in (a).

without making any corrections. In fact, this means that we can only obtain the *3D richness* of a cluster.

We define the 3D richness measure C_{3D} as the number of galaxies within the magnitude interval $[m_3, m_3 + 2]$ within a cylinder of radius $1.5\,h^{-1}$ Mpc (Abell's compactness criterion) and depth $10\,h^{-1}$ Mpc, which is the typical cluster turn-around radius for our catalogue. We chose a constant depth because the radius of the simulation sphere was adjusted according to the expected final cluster mass in order to include in the modelling a sufficient volume around the cluster turn-around radius.

Mazure et al. (1996) found that *on average* C_{ACO} is equal to C_{3D}, but there exists a significant scatter in their relation, because for C_{ACO} a constant field correction term was applied while for C_{3D} the field correction term was estimated for each individual cluster. We used their data as plotted in their fig. 1(c) to find that scatter. A linear fit gives $\sigma_{C_{ACO}-C_{3D}}(C_{3D}) = 0.19C_{3D}$.

5.2 Completeness of the model catalogue

The main purpose of our catalogue of galaxy cluster models is the comparison of its statistical properties to those of the ENACS catalogue, so we need both samples to be complete in the same (observable) quantity. The ENACS sample is complete in the ACO richness measure C_{ACO}, so we desire that the model catalogue is complete in C_{ACO} as well. However, we can only obtain C_{3D} for our models, so we can only strive for completeness in C_{3D}. However, as C_{ACO} is statistically equivalent to C_{3D} (Mazure et al. 1996), we can still compare the statistical properties of both catalogues.

Unfortunately, completeness in C_{3D} is not readily achieved. As discussed in Section 3, we had to construct the catalogue using the final mass estimator M_G. Only if there exists a noiseless relationship between M_G and C_{3D} will completeness in 3D richness immediately be guaranteed. However, richness is determined from the projected galaxy distribution, while M_G is related to the total matter distribution within a sphere. For this reason, and others that are discussed in Section 5.3 below, there is a rather large scatter in the correlation between M_G and C_{3D}, so near the mass threshold some clusters drop out of the richness-selected catalogue, while others can enter from below the expected mass threshold. In other words, low-mass clusters with a relatively high richness are, by construction, not included in our catalogue, and a catalogue that is complete in mass will not be complete in richness. As this comprises a significant fraction of all clusters, we need to correct for this.

5.3 Correcting for incompleteness

We correct for incompleteness in richness *directly* from the relation between richness C_{3D} and predicted final mass $M_G(\nu, x)$ (the catalogue defining quantity). The method consists of fitting an appropriate function to the (M_G, C_{3D}) pairs obtained from the simulations, calculating the scatter of the residuals for several bins in M_G, and fitting a function to those as well. In order to find a richness for the M_G that we did not select for simulation, we sample these richnesses from a normal distributed with mean and variance given by the two fits.

In Fig. 9(a) we show the relation between C_{3D} and M_G. The filled symbols indicate the *simulated* clusters at $\sigma_8 = 0.46$. The vertical dashed line shows the M_G threshold that selects 86 clusters, the number appropriate to the chosen volume (see Section 4.3), and the vertical solid line the M_G threshold that selects the 99 cluster models that we have actually simulated. The horizontal line shows the richness limit used to compile observational cluster catalogues (on average, $C_{3D} \geq 50$ gives the same number of clusters as $C_{ACO} \geq 50$).

To estimate the distribution of rich Abell clusters with masses below the $M_G(\nu, x)$ threshold, we assume a functional form for the mass–richness relation, and measure its total scatter. This functional form should follow the general trend of the points, and stay positive for all M_G. We find that an exponential of the form $C_{3D} = a\ e^{bM_G}$ achieves all this. A fit of this form is shown in Fig. 9(a) as a solid line. We found $a = 5.6$ and $b = 1.12$. The scatter around this fit as a function of M_G is shown in Fig. 9(b), along with an exponential fit to it: $\sigma(C_{3D}) = 0.96\ e^{1.48M_G}$. This fit provides the 1σ deviations from the M_G–C_{3D} relation which are plotted as dotted lines in Fig. 9(a). Combining both fits we derive that $\sigma(C_{3D}) = 0.35C_{3D}$.

We now assume that low-mass/high-richness clusters still correspond to initial density peaks, and should therefore be found amongst peaks *below* the M_G threshold. For the standard CDM

scenario chosen here, the adopted volume of $10^7\,h^{-3}$ Mpc is expected to contain $N_{\rm peak}(>0, R_{\rm G}) = 1164$ peaks (see also Section 3.3.1). We then use both fits to estimate $C_{\rm 3D}$ for each $M_{\rm G}$ of the remaining 1078 peaks *below* the $M_{\rm G}$ threshold. This is done by sampling them from a normal distribution with mean $5.6\,{\rm e}^{1.12M_{\rm G}}$ and variance $0.35C_{\rm 3D}$, i.e. the fits from Fig. 9. The resulting data points are plotted as triangles and open circles in Fig. 9(a). All points below the $M_{\rm G}$ threshold that have $C_{\rm 3D} > 50$ should be the clusters that we miss in our catalogue; these points are plotted as triangles. Note that most of them are near the $M_{\rm G}$ threshold, so it did make sense to run 13 more cluster models below that threshold. In fact, in turns out that four of those show up as rich clusters (i.e. $C_{\rm 3D} \geq 50$) at $\sigma_8 = 0.46$, as can be seen in Fig. 9(a). According to the exponential fit, and its scatter, this is a plausible number.

The distribution of all points plotted in Fig. 9(a) is the *complete* richness distribution (in a statistical sense). Since, on average, $C_{\rm 3D}$ is equal to $C_{\rm ACO}$, we found the complete distribution function for $C_{\rm ACO}$ as well. Note that the number of $C_{\rm 3D} \geq 50$ rich Abell clusters we find from this method is not necessarily the amount we need (i.e. 86 in our case). This will depend on the value of σ_8, and we actually use this to constrain its value in Section 6.1.2.

For the choice of $\sigma_8 = 0.46$ that we used to illustrate the method, we find that we have 78 clusters with $C_{\rm 3D} \geq 50$, of which 60 were actually simulated. For the other 18 clusters we just sampled the richness from an estimated probability distribution, i.e. we did not simulate them, and therefore any other properties of these models are unknown. In Section 7.1.1 we use a similar method to obtain line-of-sight velocity dispersions.

5.4 A complete subsample

If we increase the richness threshold used to define a sample of galaxy clusters, we can obviously make the sample complete since the richest clusters are all simulated. Indeed Mazure et al. (1996) found that they could construct a truly *complete* subsample of ACO clusters for a $C_{\rm 3D}$ richness limit of 75, in a cone out to $z = 0.1$. After inspection of Fig. 9(a) we find that our simulation set at σ_8=0.46 is also complete for $C_{\rm 3D} \geq 75$ (i.e. above the horizontal dotted line in the figure). This allows a direct comparison of statistical properties of these complete observational and model samples, but the sample size is obviously rather small: 33 and 29 clusters respectively (for slightly different volumes, and for $\sigma_8 = 0.46$).

5.5 Sources of incompleteness

In this section we estimate the *intrinsic* noise of the richness measure that is due to its very definition and due to projection effects. This allows us to find the scatter in the $M_{\rm G}$–$C_{\rm 3D}$ relation, as the total scatter found in Section 5.3 is the combination of all these.

5.5.1 *Evolution of the third brightest galaxy*

Recall that richness is defined as the number of galaxies within the magnitude range $[m_3, m_3 + 2]$. Because the luminosity function is not very well sampled at the high-end, m_3 can vary quite significantly from cluster to cluster for equal total galaxy numbers and identical luminosity functions. In other words, the upper limit of the magnitude range, the magnitude of the third brightest galaxy, is far more important for the richness measure than is the lower limit. A consequence of this is that the richness of a cluster can change significantly with time if m_3 changes due to, for example, a galaxy merger event.

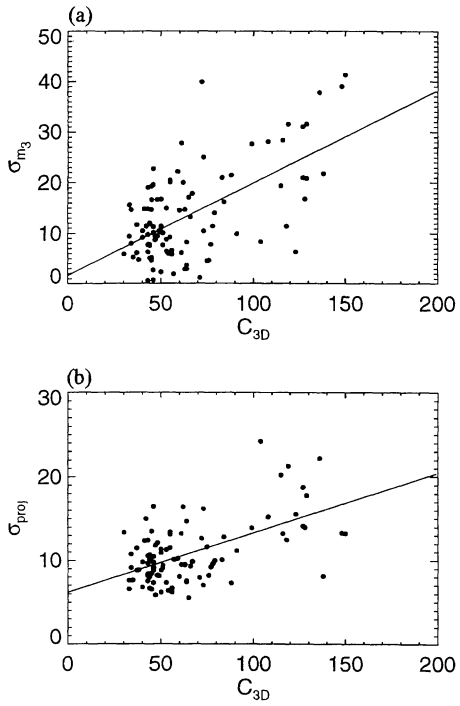

Figure 10. Noise in the 3D richness measure. (a) Scatter in 3D richness due to m_3. (b) Scatter in 3D richness due to projection.

We estimate the rms scatter due to the evolution of m_3 by taking m_i for $i = \{1, 2, 3, 4, 5\}$ and calculating the standard deviation of the five corresponding richness measures. This gives us a measure for the 'jump' in richness due to a change of the third brightest galaxy. We find a spread of around 18 per cent on average, with a minimum of 1 per cent (model 4), and a maximum of 52 per cent (model 21). More quantitatively, the scatter due to m_3 as a function of $C_{\rm 3D}$ is plotted in Fig. 10(a), and a linear fit gives $\sigma_{m_3}(C_{\rm 3D}) = 2 + 0.18C_{\rm 3D}$.

5.5.2 *Projection effects*

Since richness is a 'projected' quantity, measured within a cylinder, it is strongly dependent on the orientation of the simulation volume towards the observer. This projection effect is strongest for hierarchical formation scenarios like CDM, which produce highly irregular, clumpy galaxy distributions. We calculated the variation of richness due to orientation by observing each model cluster from 200 random directions. We find that the spread in R due to projection effects is on average 13 per cent, with a maximum of 39 per cent for model 97 (a poor cluster), and a minimum of 7 per cent for model 76 (a fairly compact, $C_{\rm 3D} = 75$ cluster). We find from Fig. 10(b) that, on average, $\sigma_{\rm proj}(C_{\rm 3D}) = 5.7 + 0.072C_{\rm 3D}$.

5.5.3 *Scatter in the $M_{\rm G}$–$C_{\rm 3D}$ relation*

The noisiness in m_3 and the strong dependence on projection make richness a measure with a rather large intrinsic spread. Quadratically adding these two noise sources, we estimate the total observational scatter in $C_{\rm 3D}$ to be $\sigma_{\rm int}(C_{\rm 3D}) = 6 + 0.2C_{\rm 3D}$. For $C_{\rm ACO}$ we need to add the scatter in the $C_{\rm ACO}$–$C_{\rm 3D}$ relation as well, yielding $\sigma_{\rm int}(C_{\rm ACO}) = 6 + 0.27C_{\rm ACO}$.

Note that there are further sources of noise in richness. For example, one might easily miss a few galaxies when observing a

rich cluster in projection due to overcrowding, especially if there is a large cD galaxy present. However, we consider these noise sources as insignificant as compared with those discussed, as only 25 per cent of all clusters have a cD galaxy (Leir & van den Bergh 1977).

With $\sigma(C_{3D})$ known from the fit shown in Fig. 9(b), we can quadratically subtract $\sigma_{int}(C_{3D})$ from that to find a scatter of about $0.25C_{3D}$ in the M_G–C_{3D} relation.

6 THE FINAL CLUSTER CATALOGUE

6.1 Timing of the models

The normalization of the standard $\Omega_0 = 1$ CDM density fluctuation is a free parameter that can be changed without changing the characteristic scales of the CDM scenario. Thus, a numerical simulation adopting such a scenario can be run to an arbitrary time and then renormalized to fix the present epoch. Usually one assumes a linear bias between the galaxy and the dark matter fluctuation spectra, and the observation that at present the rms of the $8\,h^{-1}$ Mpc top-hat smoothed galaxy distribution is about unity (Davis & Peebles 1983). A major advantage of actually forming galaxies during the simulation is that it allows us to make a direct comparison with observations without having to assume a value for the bias.

Initial conditions for the cluster models were set up such that σ_8, the rms of the $8\,h^{-1}$ Mpc top-hat smoothed total density field, is unity at the present epoch, i.e. there is no linear bias. For the standard CDM scenario σ_8 was found to be significantly smaller than unity in most earlier work (e.g. Davis et al. 1985; Frenk et al. 1990; Bertschinger & Gelb 1991). We ran all models up to $\sigma_8 = 0.65$, and a subset to $\sigma_8 = 1$, the unbiased normalization.

6.1.1 *Galaxy autocorrelation function*

A traditional measure for timing numerical simulations of the large-scale structure of the Universe is the galaxy autocorrelation function. Interestingly, this function is claimed to be mainly determined by galaxies within $10\,h^{-1}$ Mpc of cluster centres (Peebles 1974; McGill 1991). The galaxy autocorrelation function obtained from our cluster models might therefore be representative of the global one. However, quite a few problems might spoil this idea in practice. First of all, the dynamical evolution of both the galaxy and the dark matter component inside the cluster is simulated only approximately, because several physical processes are modelled only crudely, or not at all. Secondly, we did not transform possible cD galaxies into single galaxy particles. Since about 25 per cent of all observed clusters contain one or more cDs (Leir & van den Bergh 1977), we need to recognize these cases and separate cluster dark matter particles and those belonging to the cD.

We therefore believe a determination of the galaxy autocorrelation function from our cluster model will not be sufficiently unbiased. For this reason we decided to use the normalization found by van Kampen (1994) from several unconstrained simulations, i.e. average patches of universe for the same scenario. These 'field models' provide better estimates for the large-scale autocorrelation function corresponding to the cosmology chosen. The same galaxy formation parameters were used as for the cluster models presented in this paper. The galaxy autocorrelation function was calculated for the simulated galaxy distribution, and compared to the observed one. It was found that both its slope and the amplitude argue for a σ_8 in the range 0.46 to 0.56 (van Kampen 1994), although its slow evolution does not rule out values somewhat outside this range.

Once we have chosen σ_8, i.e. defined the present epoch in the simulations, we can test statistical properties of the catalogue. Alternatively, we can use such statistics as an additional and independent means of normalizing the cosmological power spectrum. Besides the autocorrelation function we use the distribution of *richness* as an independent measure to constrain σ_8 by matching it to that observed for Abell clusters. If we have selected the correct scenario and if our modelling is sufficiently realistic, both measures should yield consistent timing, certainly if these measures are on roughly the same length-scale. Note that the results of *COBE* demand $\sigma_8 \approx 1.3$ for this scenario (Bunn, Scott & White 1995), i.e. a linear *anti*bias, but remember that *COBE* measures power on a scale well outside the dynamical range of our simulations (indicated by the two vertical lines in Fig. 2). This will be discussed further in Section 8.

6.1.2 *Richness distribution*

We now combine all richness determinations for the individual models to obtain a statistical distribution function. Fig. 11 shows histograms of the richness measure C_{3D} found from our models (thin solid lines) for four values of σ_8. Only the 86 clusters above the M_G threshold are used for each histogram. We applied the method of Section 5.3 to correct for incompleteness. The corrected distributions are plotted in Fig. 11 as thick solid lines.

The dotted lines denote the distributions of richness measure C_{ACO} for the ENACS sample, which is complete (see Section 5.1). To guide the eye, we also included, as a dashed line, the fit to the richness distribution for the complete ACO catalogue (see Fig. 8), rescaled to the ENACS volume. As the distributions for C_{ACO} and C_{3D} are statistically equivalent, we can directly compare our corrected C_{3D} distribution to the C_{ACO} distribution of the ENACS sample and the rescaled ACO catalogue.

We clearly observe a monotonic evolution of our corrected C_{3D} distribution. Those for $\sigma_8 = 0.46$ or 0.51 fit the observed distribution best, although the values of 0.41 and 0.56 are not strongly ruled out. However, even smaller or higher values for σ_8 are excluded (we did not plot these in Fig. 11). We can summarize the evolution by plotting the density of rich clusters as a function of σ_8. We do this for both the corrected $C_{3D} \geq 50$ sample and the complete $C_{3D} \geq 75$ subsample. These are shown in Fig. 12, along with exponential fits to them. This functional form probably fits best for the range of σ_8 we considered here, although one expects the curve to stop rising exponentially for large σ_8.

From the exponential fits we find that $\sigma_8 = 0.46$ matches best. White et al. (1993) find a value of $\approx$0.52 from a comparison of cluster abundances calculated using the Press–Schechter (1974) formalism and standard N-body modelling (i.e. no modelling of galaxy formation) to X-ray data. A similar comparison to the optical cluster abundance obtained from the median velocity dispersion of $C_{ACO} \geq 50$ clusters gives $\sigma_8 \approx 0.62$. Eke et al. (1996) find 0.50 ± 0.04 from similar simulations (but at a higher resolution) as White et al. (1993), when compared to just the X-ray data, where they used the same X-ray data as White et al. but with a slightly different way of inferring the X-ray cluster abundance. Unfortunately it is not straightforward to compare these inferred values for σ_8 to the one we infer from the richness distribution, as there may be systematic differences between the richness distribution and the X-ray temperature function.

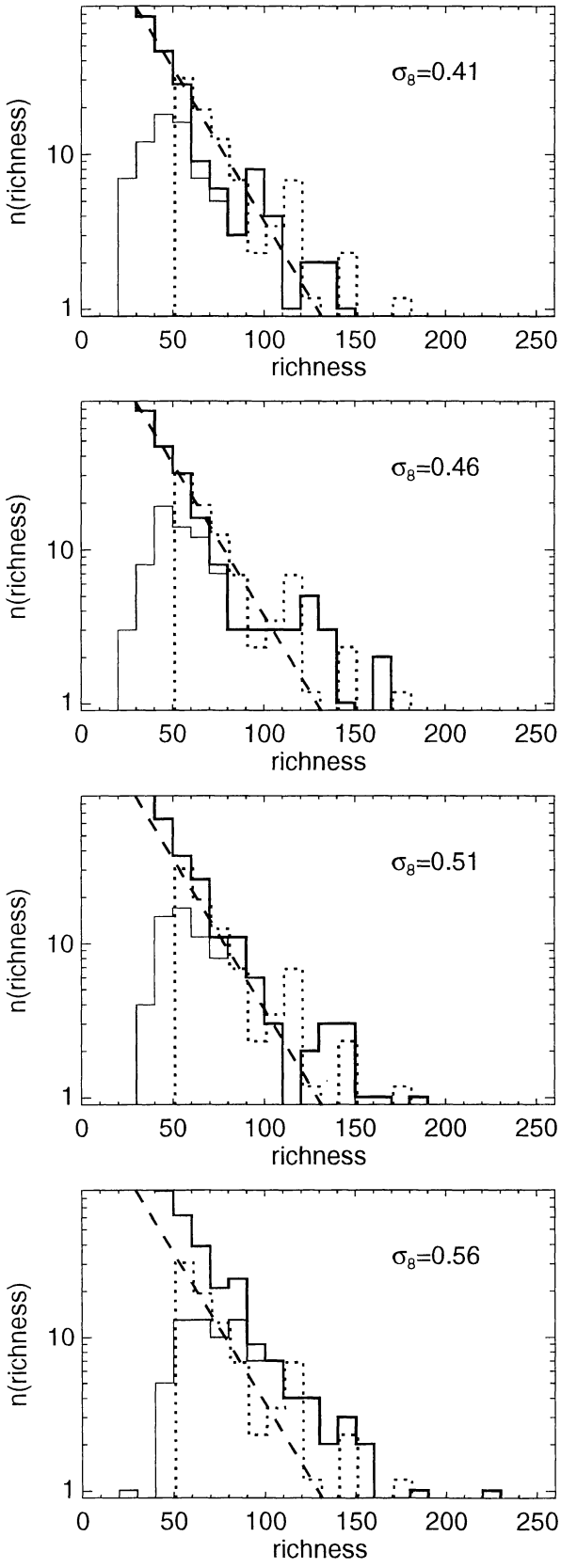

Figure 11. Uncorrected distributions of the richness measure C_{3D} obtained for our model clusters (thin solid lines), and those corrected for incompleteness (thick solid lines), for four different choices of σ_8. The dotted lines give the observed distributions of C_{ACO} for the ENACS sample, the dashed lines a fit to the distribution for the ACO sample, normalized to the ENACS volume.

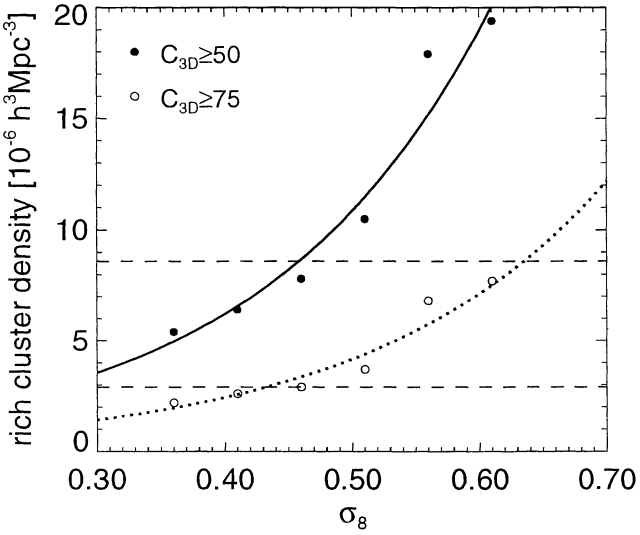

Figure 12. Number density of rich Abell clusters as a function of σ_8 for the standard CDM scenario. Closed symbols are for the corrected $C_{3D} \geq 50$ sample, open ones are for the complete $C_{3D} \geq 75$ subsample. The horizontal dashed lines indicate the observed densities for both samples, whereas the curves show exponential fits to the plotted points. The preferred value for σ_8 is 0.46.

6.1.3 Choice for σ_8

The galaxy autocorrelation function and the richness distribution for the models leave a reasonably narrow range of allowed values for σ_8. In defining the catalogue using the expected final mass M_G, we have adopted $\sigma_8 = 0.46$ as the final epoch, as it is the preferred value for the richness distribution, which is the most restrictive of the two measures.

By adopting $\sigma_8 = 0.46$ we have a single catalogue that can be used for subsequent study. This should include a study of the distribution of intrinsic properties of clusters, and a more detailed comparison to the ENACS and other observed samples.

6.2 Selection of catalogue entries

Now that we have 'calibrated' σ_8 to be 0.46, we select those cluster models from the set of 99 simulated that have $C_{3D} \geq 50$ in order to construct a richness-selected catalogue. We find that we have simulated 60 rich clusters out of the 86 desired, i.e. we are 70 per cent complete in richness C_{3D}. However, the density of rich clusters we find for $\sigma_8 = 0.46$ is only $7.8 \times 10^{-6}\, h^3\,\mathrm{Mpc}^{-3}$, while that should be 1.1 times larger according to the fit of Fig. 12 *and* the density found for the ENACS sample (Mazure et al. 1996). So we expect to find 66 rich clusters in our M_G-selected catalogue, i.e. we expect to be 77 per cent complete, and to miss 20 rich clusters that will form from peaks below the M_G threshold for the whole set of 99 models. The fraction of peaks that form a rich cluster at the present epoch is quite small there, so it is not straightforward to model those clusters (see also Fig. 9a).

Fortunately, as discussed in Section 5.3, we do have a complete subsample of $C_{3D} \geq 75$ clusters in our simulation set, i.e. a sample for which no statistical corrections are needed and for which we have each entry numerically modelled.

The resulting catalogue is presented in Table 1, which lists the defining constraints as well as some global physical properties

Table 1. Parameters and global properties of the $\Omega_0 = 1$ CDM model cluster catalogue for $\sigma_8 = 0.46$, with $H_0 = 50\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Please refer to Section 6.3 for a detailed description of each column.

no. (a)	ν	x	a_{12}	a_{13}	$M_G(\nu,x)$	M_{sim}	M_A	C_{3D}	$\overline{C_{3D}}$	σ_{dm}	σ_{gal}	σ_{los}	$\overline{\sigma_{los}}$	m_3	m_{10}	r_{turn}
					$[10^{15}M_\odot]$					$[\mathrm{km\,s^{-1}}]$				$[10^{12}M_\odot]$		[Mpc]
1	3.957	3.054	1.163	1.250	2.89	3.28	1.10	76	76	1265	1276	1189	1201	9.53	5.60	12.73
2	2.986	2.085	1.206	1.706	2.14	2.02	0.41	63	59	974	914	648	611	5.49	3.61	9.32
3	2.781	2.037	1.355	1.608	1.94	2.26	0.31	51	43	532	507	518	540	7.23	5.26	8.61
4	3.842	3.338	1.206	1.706	2.69	2.67	0.78	113	123	1235	1279	1351	1124	3.87	3.10	11.38
5	3.306	2.502	1.876	2.119	2.36	2.37	0.50	55	57	778	686	643	730	6.40	3.90	10.31
6	2.874	2.484	2.358	3.300	1.91	1.91	0.36	28	33	653	683	688	675	4.76	2.80	7.89
7	3.517	3.195	1.227	1.427	2.39	2.42	0.77	114	129	896	835	746	837	4.03	2.99	11.39
8	3.600	3.518	1.332	1.504	2.38	2.36	1.03	131	128	1049	1030	883	985	4.74	3.52	11.95
9	3.881	3.763	1.208	1.460	2.61	2.65	0.82	120	127	1164	1127	1095	962	3.61	2.78	10.97
10	3.470	3.540	1.672	1.898	2.24	2.13	0.83	139	138	1048	982	879	903	3.86	3.05	10.33
11	3.528	3.606	1.302	1.389	2.28	2.25	0.90	121	118	1159	1114	1084	1063	4.56	3.21	11.25
12	3.601	3.252	2.294	2.710	2.46	2.31	0.58	124	116	1141	1101	1028	973	4.20	3.47	10.03
13	4.507	4.061	1.120	1.302	3.19	3.15	1.33	161	150	1232	1240	1149	1299	4.66	3.44	12.91
14	3.206	2.376	2.451	2.857	2.29	2.38	0.45	58	51	851	797	701	668	5.93	4.14	9.94
15	3.675	3.325	1.346	1.770	2.52	2.56	0.83	160	136	1256	1279	1274	1192	3.58	2.68	11.14
16	3.284	2.448	1.553	2.841	2.35	2.46	0.56	69	75	741	770	763	819	5.41	4.23	10.94
17	2.560	1.580	1.595	3.322	1.83	1.68	0.36	46	40	814	843	771	690	7.63	4.08	8.67
18	3.653	3.190	1.203	1.580	2.53	2.63	0.82	92	104	1013	1104	1054	1149	4.80	3.36	11.59
19	3.218	2.677	1.527	2.058	2.22	2.11	0.49	44	44	683	573	636	640	9.58	4.73	10.03
20	3.233	3.262	1.477	1.715	2.07	1.95	0.61	46	46	995	993	939	880	8.69	4.94	9.85
21	2.828	2.036	3.610	4.762	1.99	1.76	0.40	37	45	810	939	848	728	6.36	3.01	9.00
22	2.671	1.930	1.513	1.988	1.85	1.95	0.28	30	41	645	677	857	655	6.48	3.16	6.99
23	3.629	2.829	1.083	1.359	2.61	2.65	0.63	86	84	981	970	876	828	6.21	4.57	11.64
24	3.064	2.520	1.560	1.709	2.10	2.05	0.41	47	49	900	910	844	764	6.68	3.45	9.81
25	3.004	2.704	1.621	1.957	1.98	1.79	0.50	48	46	780	704	651	704	6.34	2.85	7.96
26	3.538	4.184	1.110	1.272	2.13	2.11	0.69	135	127	1035	1042	992	909	3.63	2.69	9.79
27	3.081	2.907	1.264	1.497	2.01	2.02	0.52	50	55	832	729	697	677	5.60	2.74	10.18
28	3.521	3.392	1.242	1.704	2.34	2.32	0.65	62	61	913	822	656	696	6.97	2.93	11.03
29	3.309	2.933	2.597	3.175	2.24	2.21	0.52	115	108	825	738	646	684	3.15	2.24	9.26
30	3.768	3.838	1.052	1.368	2.47	2.10	0.92	76	72	1076	1015	1005	1085	6.22	3.44	11.16
31	2.895	2.400	1.252	2.000	1.95	1.73	0.30	40	46	692	616	622	602	6.17	4.30	8.67
32	3.929	4.345	1.312	1.570	2.50	2.48	1.20	126	119	1403	1353	1294	1245	4.54	2.83	11.93
33	3.412	3.022	1.387	1.953	2.32	2.47	0.48	105	99	878	792	846	829	4.86	3.45	9.24
34	2.751	1.489	2.193	2.933	2.06	1.77	0.25	43	44	672	634	515	586	8.89	5.24	8.85
35	2.689	1.741	1.689	2.062	1.92	1.70	0.27	51	48	606	583	556	547	4.38	3.22	6.54
36	2.938	2.016	1.364	1.845	2.11	1.63	0.56	53	56	764	746	723	739	7.66	4.63	8.89
37	3.193	3.229	1.266	1.590	2.04	1.97	0.65	75	77	1055	1016	1003	966	6.39	4.05	9.94
38	2.902	2.373	1.342	1.859	1.97	1.77	0.37	44	45	789	938	885	795	5.14	2.94	9.44
39	2.792	2.232	2.203	2.506	1.89	1.68	0.46	77	71	745	814	727	669	5.17	3.32	8.81
40	2.827	2.241	1.073	1.927	1.93	1.78	0.36	38	44	696	698	689	659	5.35	2.93	8.24
41	4.294	3.682	1.198	1.328	3.07	3.12	1.27	140	148	1258	1268	1201	1186	5.10	3.63	12.91
42	3.442	3.195	1.085	1.479	2.31	2.20	0.65	121	129	849	821	753	765	4.20	2.51	10.38
43	3.302	3.067	1.276	1.376	2.20	2.30	0.85	83	73	1207	1181	1164	1074	7.14	4.49	11.03
44	2.860	2.407	1.736	2.114	1.91	1.76	0.44	57	54	756	669	627	624	7.48	4.57	8.69
45	3.345	2.957	2.506	3.096	2.27	1.99	0.43	37	46	694	701	484	579	5.16	2.49	8.95
46	3.311	2.694	1.404	1.733	2.31	2.51	0.61	79	66	1213	1170	934	914	8.16	5.20	10.27
47	3.429	3.107	1.610	1.869	2.32	2.19	0.47	63	73	770	676	706	729	4.43	2.40	9.42
48	3.431	3.331	1.145	1.250	2.26	2.15	0.87	93	88	1247	1171	1147	1058	5.41	3.04	11.14
49	2.778	2.177	1.828	2.037	1.89	2.09	0.36	57	63	786	724	730	629	4.45	3.07	8.74
50	2.946	1.916	2.755	3.636	2.14	2.34	0.30	47	43	653	572	622	595	5.58	4.14	9.00
51	3.265	3.010	1.473	1.613	2.17	2.16	0.68	105	91	1163	1121	995	956	5.77	3.82	10.68
52	3.229	2.938	1.845	2.141	2.15	2.18	0.47	61	64	828	789	676	782	5.67	3.25	8.93
53	3.193	2.862	1.166	1.550	2.14	2.15	0.51	102	115	965	965	845	824	4.22	2.59	9.52
54	2.980	2.239	1.055	1.490	2.09	2.11	0.38	38	43	636	658	679	616	8.18	4.02	9.43
55	3.294	3.233	1.124	1.391	2.14	2.06	0.60	69	79	838	803	701	720	5.83	2.87	10.42
56	2.923	2.554	1.490	1.789	1.94	1.89	0.39	47	45	763	727	673	689	6.08	4.11	9.31
57	3.128	2.863	1.353	1.460	2.07	2.08	0.75	90	83	1182	1259	1220	1045	5.37	3.62	10.03
58	3.145	2.362	1.389	1.637	2.23	2.46	0.63	60	64	956	960	915	892	7.65	5.76	10.99
59	3.247	1.978	1.739	2.020	2.44	2.67	0.49	40	35	839	800	863	757	8.50	4.28	10.40
60	2.980	2.564	2.558	3.030	2.00	2.31	0.36	22	34	632	682	594	614	7.14	2.24	9.02
61	3.035	2.676	1.309	1.515	2.02	2.12	0.49	50	50	830	879	710	716	7.70	5.88	9.81
62	2.822	1.788	1.946	2.193	2.05	1.99	0.29	42	46	584	590	619	577	4.59	2.18	8.68

Table 1 – *continued*

no. (a)	ν	x	a_{12}	a_{13}	$M_G(\nu,x)$	M_{sim}	M_A	C_{3D}	$\overline{C_{3D}}$	σ_{dm}	σ_{gal}	σ_{los}	$\overline{\sigma_{los}}$	m_3	m_{10}	r_{turn}
					[$10^{15}M_\odot$]					[km s^{-1}]				[$10^{12}M_\odot$]		[Mpc]
63	3.348	3.592	1.151	1.323	2.10	2.12	0.70	58	67	933	987	878	950	5.13	2.38	10.92
64	2.971	1.758	1.350	1.497	2.21	2.34	0.28	40	37	605	540	485	551	10.45	4.57	9.94
65	2.746	2.089	1.550	1.862	1.88	1.90	0.37	79	53	631	651	634	545	4.96	3.56	8.50
66	3.180	2.646	1.272	1.733	2.19	2.12	0.32	37	43	649	757	589	605	6.06	2.79	9.87
67	3.031	2.504	1.161	1.502	2.07	2.09	0.45	48	47	722	634	753	705	7.11	4.07	9.59
68	3.116	2.993	1.684	2.469	2.02	2.03	0.44	47	50	772	767	749	725	4.85	2.57	9.48
69	2.957	2.432	1.453	1.718	2.01	1.93	0.40	57	53	919	979	900	792	5.71	4.65	9.55
70	3.018	2.034	1.196	1.783	2.19	2.41	0.34	63	47	735	663	690	687	7.52	5.38	9.39
71	2.883	2.221	1.447	2.137	1.99	2.05	0.31	48	50	648	695	542	561	6.05	4.33	8.48
72	3.621	2.663	2.494	3.195	2.65	2.22	0.49	51	55	817	688	693	728	6.19	3.42	8.50
73	3.153	2.657	1.148	1.560	2.15	2.21	0.49	73	63	751	695	597	653	6.91	5.34	10.24
74	3.157	3.253	1.403	1.616	1.99	2.05	0.49	45	40	735	749	682	737	6.68	4.61	9.75
75	3.037	2.901	1.290	1.706	1.96	2.11	0.49	60	48	875	815	795	764	5.78	4.02	9.83
76	3.350	4.140	1.050	1.323	1.95	2.08	0.82	66	65	999	932	794	876	7.60	4.83	10.42
77	3.218	3.718	1.220	1.497	1.92	2.07	0.60	28	62	914	903	860	917	10.88	3.50	10.27
78	3.074	3.417	1.441	1.656	1.86	1.97	0.50	58	60	899	855	895	817	5.41	3.15	9.57
79	3.149	3.416	1.337	1.842	1.94	1.95	0.58	84	78	722	676	773	833	4.18	3.34	9.55
80	3.040	3.143	1.185	1.433	1.90	2.02	0.61	55	56	810	802	812	809	6.66	3.71	10.11
81	3.041	3.213	1.252	1.475	1.88	2.00	0.56	38	38	846	802	727	834	8.29	4.09	9.59
82	3.037	3.247	1.927	2.604	1.87	1.77	0.52	43	45	924	1045	983	883	6.77	3.79	9.44
83	3.005	3.219	1.534	2.110	1.84	1.74	0.48	38	43	808	773	692	761	7.53	3.43	9.17
84	2.972	3.166	2.041	2.278	1.82	1.69	0.55	45	51	736	662	667	757	6.46	4.13	8.85
85	2.949	2.853	1.582	2.242	1.88	1.99	0.26	34	42	610	607	655	614	5.69	3.12	8.63
86	2.969	2.951	1.621	1.799	1.88	1.77	0.40	49	55	890	858	812	817	7.39	4.32	8.32
87	2.983	2.998	1.202	1.370	1.88	1.85	0.45	37	34	746	781	657	646	7.25	4.27	9.07
88	2.992	2.941	1.513	1.653	1.90	2.05	0.64	60	64	1044	1078	1092	932	6.68	3.81	9.85
89	2.918	2.681	1.661	2.506	1.90	1.85	0.34	48	50	642	667	633	667	5.62	2.62	8.56
90	3.003	2.863	1.838	2.028	1.94	1.98	0.49	61	53	829	728	725	736	7.49	5.32	9.59
91	2.899	2.691	1.364	1.931	1.88	2.03	0.27	51	37	528	508	494	503	8.69	6.33	7.06
92	2.859	2.692	1.499	1.980	1.83	1.80	0.49	71	44	851	839	809	731	8.34	5.33	8.37
93	2.823	2.645	1.608	1.776	1.81	1.93	0.39	28	45	752	736	629	680	5.94	2.34	8.65
94	2.823	2.523	1.754	2.066	1.84	1.95	0.32	69	61	640	676	505	553	4.89	3.65	7.26
95	2.794	2.401	1.267	1.715	1.85	1.90	0.41	46	46	882	866	773	722	6.24	3.96	8.93
96	2.660	1.955	1.416	2.747	1.83	1.71	0.18	43	43	498	463	473	474	4.93	2.52	6.18
97	2.646	2.076	1.475	1.912	1.78	1.77	0.15	24	30	561	556	639	545	7.67	3.45	9.13
98	2.611	1.957	2.532	2.933	1.78	1.85	0.22	51	44	474	447	422	499	8.06	4.92	6.66
99	2.540	1.643	1.634	1.832	1.79	2.05	0.23	33	33	645	682	630	553	5.63	2.13	8.02

Notes: (a) bold face: $C_{3D} \geq 75$; italics: $C_{3D} < 50$.

(discussed below). The cluster models that are not rich enough are listed in italics, and those that are in the complete $C_{3D} \geq 75$ are shown in bold face. Note that the selection of models that are part of the catalogue is specific for the choice of the present epoch. An extreme example of this is model no. 60. Its mass within $6\,h^{-1}$ Mpc almost doubles when going from $\sigma_8 = 0.46$ to 0.51 because another clump enters the top-hat window. However, it is also the least rich model of all (because its line of sight is perpendicular to the merging direction), and therefore is well below the richness limit.

6.3 Global properties of the cluster models

The properties of all models are listed in Table 1. Columns 2 to 5 give the peak parameters used to define the cluster, which also determine the expected final cluster mass $M_G(\nu,x)$ listed in column 6. The mass within $6\,h^{-1}$ Mpc found in the simulations, used to fit (21), is given in column 7. The Abell mass M_A, i.e. the mass within a sphere of $1.5\,h^{-1}$ Mpc, can be obtained for most observed clusters (although usually only approximately), and is therefore listed in column 8. The 3D richness measure C_{3D}, which is *on average* equal to the ACO richness measure that usually defines an observational catalogue, is listed in column 9. Since its value can be quite dependent on the chosen line of sight, especially for low-richness clusters, we show in column 10 the same quantity averaged over 200 lines of sight.

The next four columns list various velocity dispersions, starting with the intrinsic velocity dispersions σ_{dm} and σ_{gal}, for the dark matter and the galaxies respectively, which are obtained within the 3D Abell radius and divided by $\sqrt{3}$ in order to compare them to the line-of-sight velocity dispersion. Column 12 gives the line-of-sight velocity dispersion $\sigma_{l.o.s.}$ obtained from the projected galaxy distribution. Column 13 provides, for comparison, the same quantity averaged over 200 lines of sight. The remaining three columns provide a few more useful cluster properties: the masses of the third- and the tenth-ranked galaxies, and the cluster turn-around radius.

7 SOME TESTABLE PROPERTIES OF THE $\Omega_0 = 1$ CDM CATALOGUE

7.1 Cumulative distribution of line-of-sight velocity dispersions

The velocity dispersion of a galaxy cluster generally increases with time for most cosmological scenarios. This makes it a useful quantity for testing the timing of our CDM $\Omega = 1$ model catalogue. We compare to the ENACS data (Mazure et al. 1996) and observations by Collins et al. (1995), taken in the context of the Edinburgh–Milano Cluster Redshift Survey (EDSGC for short). Both these samples are relatively complete and should provide the best comparison, especially the ENACS sample because we used exactly the same data analysis procedures. We also consider the samples of Zabludoff, Huchra & Geller (1990, ZHG from here on) and Girardi et al. (1993, GBGMM from here on), although both of these are not drawn from a complete sample. The ZHG sample might be complete but is rather small (25 clusters), which makes completeness hard to establish, while the GBGMM sample is biased towards richer clusters (although an attempt has been made to correct for this).

Clearly we have to use the line-of-sight velocity dispersion, denoted by $\sigma_{\rm l.o.s.}$, for a comparison to observations. There are quite a few methods for obtaining line-of-sight velocity dispersions from observational data, all dealing with the removal of foreground and background galaxies, and 'interlopers', unbound galaxies near the cluster turn-around radius. With these galaxies removed, the velocity dispersions found are generally smaller than obtained from the contaminated velocity distribution (e.g. Lucey 1983; Frenk et al. 1990; den Hartog & Katgert 1996; Mazure et al. 1996).

However, as discussed in Section 5.1.1, in our models we do not have any foreground or background galaxies due to the limited simulation volume. With respect to removing interlopers, for the most conservative schemes, like '3σ-clipping', we find very few interlopers in our models. For more strict (but physically motivated) interloper removers, such as those proposed by den Hartog & Katgert (1996), we indeed find quite a few more, and, more importantly, in roughly the same numbers as for observed clusters. Furthermore, the interlopers that are removed are indeed outside the turn-around radius, justifying the method of den Hartog & Katgert (1996).

Because EDSGC, ZHG and GBGMM use fairly conservative interloper removers, we do not need to remove any galaxies from our models for a comparison to these samples. For a match to the ENACS clusters, we apply the interloper remover of den Hartog & Katgert (1996), and also use the smaller aperture of $1.0\,h^{-1}$ Mpc instead of the Abell radius used for the other samples. This smaller aperture results in slightly larger line-of-sight velocity dispersions, while the interloper remover used for the ENACS sample produces significantly smaller dispersions.

7.1.1 *Correction for incompleteness*

Before we can compare our line-of-sight velocity dispersions to those observed, we again need to correct for incompleteness in $C_{\rm 3D}$. We use a similar method to that used to correct the richness distributions (see Section 5.3), but now using the $\sigma_{\rm l.o.s.}$–$M_{\rm G}$ relation. We fit an exponential function to the ($M_{\rm G}$,$\sigma_{\rm l.o.s.}$) pairs obtained from the simulations, and calculate the scatter of the residuals for several bins in $M_{\rm G}$. We then sample the velocity dispersions from a normal distributed with mean and variance given by these two fits. As we use the same $M_{\rm G}$s as in the correction procedure for richness $C_{\rm 3D}$, we can link them and obtain ($C_{\rm 3D}$,$\sigma_{\rm l.o.s.}$) pairs. Adding these to the ones taken directly from the simulations gives us a complete distribution over line-of-sight velocity dispersions.

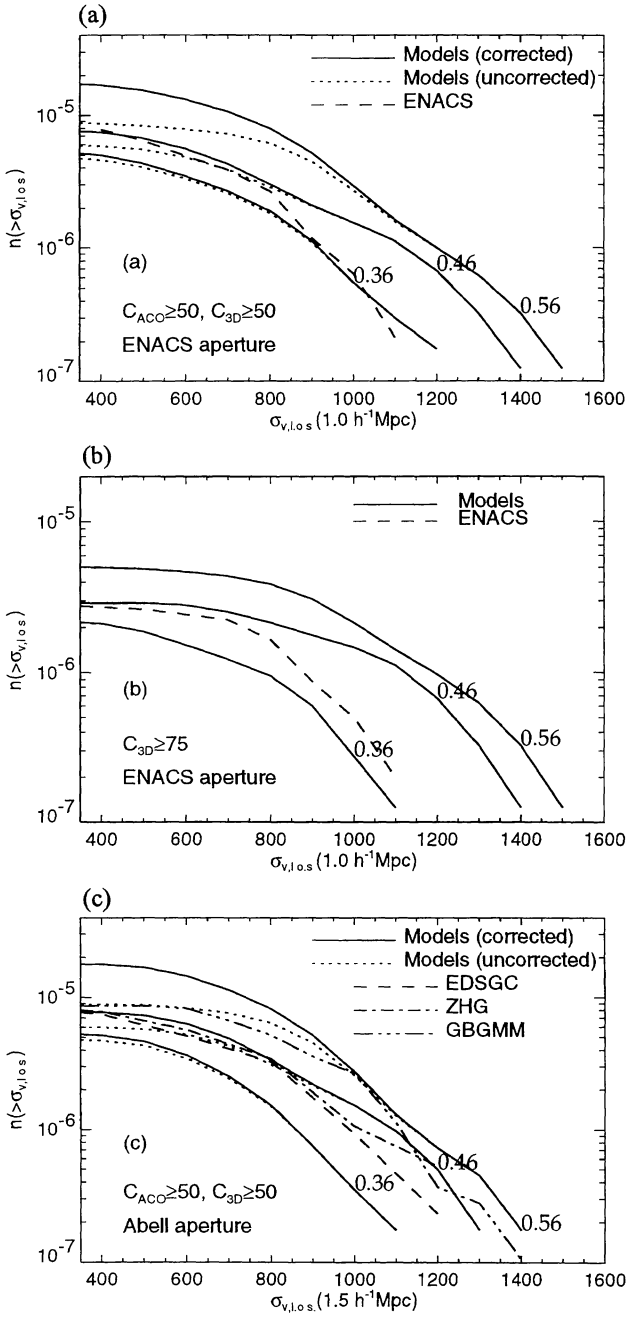

Figure 13. (a) Cumulative distributions of the line-of-sight velocity dispersion within 1.0 h^{-1} Mpc, normalized to the average cluster density. The dashed line denotes the observed distribution for the ENACS sample. Numbers denote the value of σ_8 for each model curve. (b) Same as (a), but for dispersion within 1.5 h^{-1} Mpc. The non-solid linestyles represent the distribution for various observed samples. (c) Same as (a), but for the complete subsample of $C_{\rm 3D} \geq 75$ clusters.

7.1.2 *Comparison to the ENACS catalogue*

The corrected $\sigma_{\rm l.o.s.}$ distributions are plotted as solid lines in Fig. 13(a) for several σ_8s, along with the uncorrected ones (dotted

lines), and the observed distribution found for the ENACS sample (dashed line). We find that the observed distribution matches the modelled one for $\sigma_8 \approx 0.35$–0.45, but has a slightly different slope.

The second comparison we make is that of the complete subsample of $C_{3D} \geq 75$ clusters to a similar complete subsample of the ENACS sample. Note that we do not have to apply any incompleteness corrections here, as both samples are complete for this richness limit. The model distribution for $\sigma_8 \approx 0.4$ matches the observed one best, as shown in Fig. 13(b). The shapes of the curves compare better than for the $C_{3D} \geq 50$ sample.

7.1.3 *Comparison to other observed samples*

We can perform the same comparison for the EDSGC, GBGMM and ZHG samples, although only the first of these is really complete in richness. The line-of-sight velocity dispersions within the Abell radius for several values for σ_8 are plotted in Fig. 13(c), along with the observed distributions. Please note that the interloper removal method of den Hartog & Katgert (1996) was *not* used here, as it was not used for any of these observed distributions either, and also note that the Abell aperture was adopted.

For the EDSGC sample we find that the model curve fits best for $\sigma_8 \approx 0.4$. The other two samples prefer much higher values, but are not complete and are biased to the richest clusters, and therefore to higher velocity dispersions.

7.2 Other observables

To test the chosen scenario one can use many other observables besides the one considered above. For example, the line-of-sight velocity dispersion *profiles* might prove useful, although unfortunately at present these can only be obtained for a few bins. However, a classification derived from this binned profile can be very instructive (den Hartog & Katgert 1996).

Another distribution function is that of the projected shapes of clusters. In a preliminary comparison presented elsewhere (de Theije et al. 1995) it was found that the predicted distribution of shape parameters of model clusters differs significantly from the observed one. Although the observed sample used in the study of de Theije et al. (1996) is not complete, the deviations are unlikely to be due to incompleteness. For a subset of our model catalogue that was rerun for the $\Omega = 0.2$ CDM scenario (in future work we will build complete catalogues for this and other scenarios) there seems to be better agreement between models and data.

8 SUMMARY AND DISCUSSION

The aim of this paper was to present a technique that produces a fair catalogue of individually simulated high-resolution cluster models. We used this technique to build one specific model catalogue that we compared to observed catalogues.

Under the assumption that rich clusters of galaxies originate from peaks in the initial density field Gaussian smoothed at $4\,h^{-1}$ Mpc, we select a catalogue on the basis of characteristic parameters of these peaks. Given the fact that we need to use the constrained random field method to generate initial conditions for the individual numerical models, we need to select the catalogue on the basis of linear functionals of the initial density field. With this in mind, we argued that cluster mass is the best catalogue-defining quantity, because it can be predicted reasonably well from a linear function of the amplitude *and* of the curvature of the initial density peak which is the progenitor of the cluster. Besides this practical argument, total cluster mass *is* a basic property of a galaxy cluster.

For the individual models a numerical code was used that contains a prescription for forming galaxies, and also allows galaxies already formed to grow (i.e. accrete particles) and merge with other galaxies. Not only does this allow us to compare directly to optical catalogues of galaxy clusters, but the properties of the galaxy population also directly set the present epoch for the cosmological scenario in which the model catalogue is embedded, i.e. they set the amplitude of the initial cosmological fluctuation spectrum. Not all cosmological scenarios allow this, but the scenario we chose, namely $\Omega_0 = 1$ CDM, does permit this a posteriori normalization and also makes it possible to compare our results to published numerical work.

We used a traditional, single large-scale simulation in order to test several issues related to the construction of the catalogue and the reliability of the individual models. We traced peaks in these simulations, and found that they do not merge during the entire run. So we indeed have a one-to-one mapping of initial peaks to final rich Abell clusters. Note that this may not be so in other cosmological scenarios. We also found that both the distribution and the evolution of the peak parameters in the single large-scale low-resolution simulation are similar to those of the set of individual high-resolution simulations. Finally, we found that the expected mass relation, a linear function of peak amplitude and peak curvature, is almost equal to that obtained from the clusters extracted from the single simulation.

Having built a catalogue selected on expected final cluster mass, we found that it was 70 per cent complete for $C_{3D} \geq 50$. This incompleteness mostly concerns the low-mass end and is mainly due to the noise still present in the relation between final cluster mass and initial peak parameters. Note, however, that the richness measure itself has a rather large intrinsic noise originating from its very definition. To allow a comparison of the statistical distribution functions of model cluster properties to those observed we devised a method to correct for this incompleteness. However, we do have a complete subsample for $C_{3D} \geq 75$, which permits a fully unbiased comparison to observations.

We use the richness distribution of both samples to set the present epoch in the models, i.e. fix σ_8. A value of 0.46 for σ_8 matches best; this is also in the range of allowed values derived from a match of the galaxy autocorrelation function for a set of field models, similar to the cluster models discussed here, to the observed autocorrelation function (van Kampen 1994). The elimination of the free parameter σ_8 of the $\Omega_0 = 1$ CDM scenario allows us to test it with other measures.

We find that the cumulative distribution of line-of-sight velocity dispersions obtained for the ENACS sample (Katgert et al. 1996), and for other samples as well, fits the model distribution best for $\sigma_8 \approx 0.4$. This is fairly consistent with both the autocorrelation function and the richness distribution. We can obtain a slightly larger σ_8 if we invoke a *velocity* bias, but this is not seen in our models if we compare the intrinsic velocity dispersions of the dark matter to those of the galaxies (van Kampen 1994). Other studies, like those of Cen & Ostriker (1992) and Katz, Hernquist & Weinberg (1992), do not find a significant velocity bias either.

However, there are also problems with the standard CDM scenario. We mention two major ones here (see Ostriker 1993 for a more extensive discussion). De Theije et al. (1995) compared the distribution of cluster shapes for our catalogue to an observed sample, and found that these differ significantly. The results from *COBE* imply that $\sigma_8 \approx 1.3$ for our chosen scenario (Bunn et al.

1995). Note, however, that *COBE* traces length-scales well outside the dynamical range of the individual models. We will need to investigate whether these inconsistencies remain for other tests before we completely discard the $\Omega_0 = 1$ CDM scenario *on the scales we studied here*. It is still useful to test more of the statistical properties of this scenario, and use successes and failures of such tests as a guide for a new choice for the most likely cosmological scenario. Of the scenarios plotted in Fig. 2, we should take a closer look at the k^{-2} spectrum, or low-Ω_0 CDM, as both of these have more power on larger scales and therefore are in better agreement with the *COBE* results.

ACKNOWLEDGMENTS

We would like to thank Roland den Hartog, Rien van de Weygaert and Tim de Zeeuw for many useful suggestions and discussion. Joshua Barnes and Piet Hut are gratefully acknowledged for allowing use of their treecode, Edmund Bertschinger and Rien van de Weygaert for providing their constrained random field code and Edmund Bertschinger for his P^3M code, Erik Deul for allowing us to use the computer systems that are part of the DENIS project, and the ENACS collaboration for allowing us to use their data before publication. EvK acknowledges EelcoSoft Software Services for partial financial support, and a European Community Research Fellowship as part of the HCM programme.

REFERENCES

Abell G. O., 1958, ApJS, 3, 211
Abell G. O., Corwin H. G., Olowin R. P., 1990, ApJS, 70, 1 (ACO)
Adler R. J., 1981, The Geometry of Random Fields. Wiley, Chichester
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15 (BBKS)
Barnes J. E., Hut P., 1986, Nat, 324, 446
Beers T. C., 1992, in Feigelson E. D., Babu G. J., eds, Statistical Challenges In Modern Cosmology. Springer-Verlag, New York, p. 111
Beers T. C., Tonry J. L., 1986, ApJ, 300, 557
Bertschinger E., 1987, ApJ, 323, L103
Bertschinger E., 1993, in Akerlof C. W., Srednicki M. A., eds, Texas/PASCOS 92: Relativistic Astrophysics and Particle Cosmology. The New York Academy of Sciences, New York, p. 297
Bertschinger E., Jain B., 1994, ApJ, 431, 495
Bertschinger E., Gelb J. M., 1991, Comput. Phys., 5, 164
Binggeli B., 1982, A&A, 107, 338
Bunn E. F., Scott D., White M., 1995, ApJ, 441, L9
Carlberg R. G., 1994, ApJ, 433, 468
Cavaliere A., Santangelo P., Tarquini G., Vittorio N., 1986, ApJ, 305, 651
Cen R., Ostriker J. P., 1992, ApJ, 399, L113
Collins C. A., Guzzo L., Nichol R. C., Lumsden S. L., 1995, MNRAS, 274, 1071
Dalton G. B., Efstathiou G., Maddox S. J., Sutherland W. J., 1992, ApJ, 390, L1
Davis M., Peebles P. J. E., 1983, ApJ, 267, 465
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1992, Nat, 356, 489
den Hartog R., 1995, PhD thesis, Rijksuniversiteit Leiden
den Hartog R., Katgert P., 1996, MNRAS, 279, 349
de Theije P. A. M., Katgert P., van Kampen E., 1995, MNRAS, 273, 30
de Theije P. A. M., van Kampen E., Slijkhuis R., 1997, MNRAS, submitted
Doroshkevich A. G., 1970, Astrophys., 6, 320
Ebeling H., Voges W., Böhringer H., Edge A. C., Huchra J. P., Briel U. G., 1996, MNRAS, 281, 799
Efstathiou G., 1990, in Peacock J. A., Heavens A. F., Davies A. T., eds, Physics of the Early Universe. SUSSP Publications, Edinburgh, p. 361
Efstathiou G., Davis M., Frenk C. S., White S. D. M., 1985, ApJS, 57, 241
Eke V. R., Cole S., Frenk C. S., 1996, MNRAS, 282, 263
Evrard A. E., 1989, ApJ, 341, L71
Evrard A. E., Mohr J. J., Fabricant D. G., Geller M. J., 1993, ApJ, 419, L9
Forman W., Jones C., 1990, in Oegerle W. R., Fitchett M. J., Danly L., eds, Clusters of Galaxies. Cambridge Univ. Press, Cambridge, p. 257
Fort B., Mellier Y., 1994, A&AR, 5, 239
Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, ApJ, 327, 507
Frenk C. S., White S. D. M., Efstathiou G., Davis M., 1990, ApJ, 351, 10
Girardi M., Biviano A., Guiricin G., Mardirossian F., Mezzetti M., 1993, ApJ, 404, 38 (GBGMM)
Hoffman Y., Ribak E., 1991, ApJ, 380, L5
Katgert P. et al., 1996, A&A, 310, 8
Katz N., Hernquist L., Weinberg D. W., 1992, ApJ, 399, L109
Kolb E. W., Turner M. S., 1990, The Early Universe. Addison-Wesley, Reading, CA
Leir A. A., van den Bergh S., 1977, ApJS, 34, 381
Lucey J. R., 1983, MNRAS, 204, 33
Mazure A. et al., 1996, A&A, 310, 31
McGill C., 1991, MNRAS, 250, 340
Ostriker J. P., 1993, ARA&A, 31, 689
Peacock J. A., Heavens A. F., 1985, MNRAS, 217, 805
Peebles P. J. E., 1974, A&A, 32, 197
Peebles P. J. E., 1990, ApJ, 365, 27
Peebles P. J. E., 1993, Principles of Physical Cosmology. Princeton Univ. Press, Princeton
Press W. H., Schechter P., 1974, ApJ, 187, 425
Press W. H., Flannery B. P., Teukolsky S. A., Vetterling W. T., 1988, Numerical Recipes: The Art of Scientific Computing. Cambridge Univ. Press, Cambridge
Rhee G. F. R. N., van Haarlem M. P., Katgert P., 1989, A&AS, 91, 513
Rice S. O., 1954, in Wax N., ed., Selected papers on Noise and Stochastic Processes. Dover, New York
Sarazin C. L., 1986, Rev. Mod. Phys., 58, 1
Spitzer L., Jr, 1969, ApJ, 158, L139
Suginohara T., Suto Y., 1992, ApJ, 396, 395
van de Weygaert R., Babul A., 1994, ApJ, 425, 59
van de Weygaert R., Bertschinger E., 1996, MNRAS, 281, 84
van Haarlem M. P., van de Weygaert R., 1993, ApJ, 418, 544
van Haarlem M. P., Frenk C. S., White S. D. M., 1997, MNRAS, 287, 817
van Kampen E., 1994, PhD thesis, Rijksuniversiteit Leiden
van Kampen E., 1995, MNRAS, 273, 295
van Kampen E., 1997, in Clarke D. A., West M. J., eds, Proc. 12th Kingston Meeting on Theoretical Astrophysics: Computational Astrophysics, ASP Conf. Ser. Vol. 123. Astron. Soc. Pac., San Francisco, p. 231
White S. D. M., Efstathiou G., Frenk C. S., 1993, MNRAS, 262, 1023
Zabludoff A. I., Huchra J. P., Geller M. J., 1990, ApJS, 74, 1 (ZHG)

APPENDIX A: EXPRESSIONS FROM THE THEORY OF GAUSSIAN RANDOM FIELDS

In this appendix we summarize some functions taken from BBKS, slightly rewritten for our purposes, as used in Section 2. The first function is a factor in the probability distribution for the peak height ν, which depends on the spectral parameters γ and R_G. It depends on γ and ν only:

$$\mathcal{F}_1(\gamma,\nu) = \frac{\gamma^3\nu^3 - 3\gamma^3\nu + [B(\gamma)\nu^2 + C_1(\gamma)]\mathrm{e}^{-A(\gamma)\nu^2}}{1 + C_2(\gamma)\mathrm{e}^{-C_3(\gamma)\nu}} \tag{A1}$$

(from equation 4.4 of BBKS), with

$$A(\gamma) = \frac{5\gamma^2}{(18 - 10\gamma^2)}\ , \quad B(\gamma) = \frac{432\gamma^2}{\sqrt{10\pi}(9 - 5\gamma^2)^{5/2}} \tag{A2}$$

and the fits

$$C_1(\gamma) = 1.84 + 1.13(1-\gamma^2)^{5.72}$$
$$C_2(\gamma) = 8.91 + 1.27\mathrm{e}^{6.51\gamma^2}$$
$$C_3(\gamma) = 2.58\gamma\mathrm{e}^{1.05\gamma^2}. \quad \text{(A3)}$$

The second function is used for the curvature distribution function:

$$\mathcal{F}_2(x) = \frac{(x^3-3x)}{2}\left[\mathrm{erf}\left(\sqrt{\frac{5}{2}}\,x\right) + \mathrm{erf}\left(\sqrt{\frac{5}{8}}\,x\right)\right] + \sqrt{\frac{2}{5\pi}}\left[\left(\frac{31x^2}{4}+\frac{8}{5}\right)\mathrm{e}^{-5x^2/8} + \left(\frac{x^2}{2}-\frac{8}{5}\right)\mathrm{e}^{-5x^2/2}\right]. \quad \text{(A4)}$$

The next two functions are used for the asphericity probabilities:

$$\mathcal{F}_3(a_{12},a_{13}) = \frac{27(a_{12}^2-1)(a_{13}^2-1)(a_{13}^2-a_{12}^2)}{2a_{12}^{-2}a_{13}^{-2}(1+a_{12}^2+a_{13}^2)^6} \quad \text{(A5)}$$

and

$$\mathcal{F}_4(a_{12},a_{13}) = \frac{a_{13}^4+a_{12}^4-a_{12}^2-a_{13}^2-a_{12}^2a_{13}^2+1}{(1+a_{12}^2+a_{13}^2)^2}. \quad \text{(A6)}$$

Finally, we list the function used to describe the asphericity of peaks in the second-order Taylor expansion, which is a function of the spherical coordinate angles θ and ϕ and axial ratios a_{12} and a_{13}:

$$\mathcal{F}_5(a_{12},a_{13},\theta,\phi) = 2(2a_{13}^2-1) - (2a_{12}^2-1) + 3\sin^2\theta[\sin^2\phi - (2a_{13}^2-1) + (2a_{12}^2-1)\cos^2\phi]. \quad \text{(A7)}$$

This is an adaptation from equation (7.4) of BBKS. It is equal to unity for spherical peaks.

We draw from the probability distributions (10) and (11) by using the rejection method. The comparison function needed is found by substituting ν^3 for $\mathcal{F}_1(\gamma,\nu)$ in (10). Its integral is

$$\frac{V}{(2\pi)^2R_*^3}\left[2-(2+\nu^2)\mathrm{e}^{-\nu^2/2}\right]. \quad \text{(A8)}$$

The comparison function for drawing x, $f_c(x,\mu,\sigma)$, is obtained by replacing $\mathcal{F}_2(x)$ with x^3, where we have defined $\mu=\gamma\nu$ and $\sigma=\sqrt{1-\gamma^2}$. The integral $F_c(x,\mu,\sigma)$ of the comparison function is

$$F_c(x,\mu,\sigma) = -\sigma^2(\mu^2+2\sigma^2+\mu x+x^2)\mathrm{e}^{-(x-\mu)^2/2\sigma^2} + \mu\sigma\sqrt{\frac{\pi}{2}}(\mu^2+3\sigma^2)\,\mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right). \quad \text{(A9)}$$

An approximate formula for the expectation value of x for a given ν is given by equation (6.14) of BBKS:

$$\langle x\rangle = \gamma\nu + \frac{3(1-\gamma^2)+(1.216-0.9\gamma^4)\mathrm{e}^{-\gamma^3\nu^2/8}}{[3(1-\gamma^2)+0.45+(\gamma\nu/2)^2]^{1/2}+\gamma\nu/2}. \quad \text{(A10)}$$

The expectation values for a_{12} and a_{13} are functions of $y\equiv(6+5x^2)^{-1/2}$:

$$\langle a_{12}\rangle = \sqrt{\frac{1-60y^4}{1-3y+30y^4}}$$
$$\langle a_{13}\rangle = \sqrt{\frac{1+3y-30y^4}{1-3y+30y^4}}. \quad \text{(A11)}$$

APPENDIX B: DERIVATION OF THE PREDICTED MASS RELATION

In this appendix we show how the curvature of the initial peak, x, comes into play in the non-linear evolution of total cluster mass. We start from the initial density profile, given by (19):

$$\delta_G(r) = \frac{\nu-\gamma x}{\sigma_0(1-\gamma^2)}\xi_G(r) + \frac{\nu-x/\gamma}{3\sigma_0(1-\gamma^2)}\nabla^2\xi_G(r). \quad \text{(B1)}$$

We neglect non-linear evolution of σ_0, so that γ remains constant (σ_1 and σ_2 remain well below unity). We assume that the autocorrelation function is a power law with index n, which evolves as $n(t)=n(t_i)\epsilon(t)$, where $\epsilon(t)$ is unity at $t=t_i$ and increases slowly. This gives us approximately

$$\nabla^2\xi_G(0,t) \approx \frac{\epsilon(t)[n(t_i)\epsilon(t)+1]}{n(t_i)+1}a^2(t)\nabla^2\xi_G(0,t_i) = -3\frac{\epsilon(t)[n(t_i)\epsilon(t)+1]}{n(t_i)+1}\gamma^2\sigma_0^2(t), \quad \text{(B2)}$$

where we have used equation (20) for the second part, and thus

$$\delta_G(0,t) = \left[\frac{\nu(t)-\gamma x(t)}{(1-\gamma^2)} + \eta(t)\gamma\frac{x(t)-\gamma\nu(t)}{(1-\gamma^2)}\right]\sigma_0(t), \quad \text{(B3)}$$

with

$$\eta(t) \equiv \frac{\epsilon(t)[n(t_i)\epsilon(t)+1]}{n(t_i)+1}, \quad \text{(B4)}$$

which monotonically increases (slowly) from its initial value $\eta(t_i)=1$. The second term in equation (B3) evolves somewhat faster than the first because of the increasing slope $n(t)$ of the autocorrelation function. This might seem unimportant, but the point is that the x terms, which grow relatively fast (as clearly seen in Fig. 6), no longer cancel because of the different growth rates of the two terms in equation (B3).

Rearranging the terms, we find

$$\delta_G(0,t) = \frac{\nu(t)[1-\gamma^2\eta(t)]-\gamma x(t)[1-\eta(t)]}{(1-\gamma^2)}\sigma_0(t). \quad \text{(B5)}$$

We can rewrite this in terms of growth factors for ν and x:

$$\delta_G(0,t) = [g_\nu(t)\nu_i + g_x(t)x_i]\sigma_0(t_0)a(t) \quad \text{(B6)}$$

with $\nu_i\equiv\nu(t_i)$, $x_i\equiv x(t_i)$ and

$$g_\nu(t) \equiv \frac{\nu(t)}{\nu_i}\left[\frac{1-\gamma^2\eta(t)}{1-\gamma^2}\right] \quad \text{(B7)}$$

and

$$g_x(t) \equiv \frac{x(t)}{x_i}\left[\frac{\gamma\eta(t)-\gamma}{1-\gamma^2}\right]. \quad \text{(B8)}$$

The CDM spectrum smoothed on $4\,h^{-1}$ Mpc has an autocorrelation function with an initial slope of $n\approx-1.5$ for the range of wavenumbers in our simulations (Efstathiou 1990). As an example of what the final relation will look like we assume that $\eta(t_0)\approx1.2$. Because $\gamma(4\,h^{-1}\mathrm{Mpc})=0.729$ for CDM, we obtain

$$\delta_G(t_0) \approx 0.8\frac{\nu(t)}{\nu_i} + 0.2\frac{x(t)}{x_i}. \quad \text{(B9)}$$

The peak amplitude $\nu(t)$ just grows, but the evolution of the curvature $x(t)$ is more complicated. Van Haarlem & van de Weygaert (1993) performed cluster simulations with constraints set on the same smoothing scale, and found that a small initial x_i results in a larger final $x(t_0)$ than a larger initial x_i. For three simulations with $\nu_i=3$ and x_is corresponding to the 5, 50 and 95 percentiles of

$P(x|3\sigma_0, 4\,h^{-1}\mathrm{Mpc})$, a linear relation provides a good fit to their results:

$$x(t_0) \approx 27 - 4.6x_\mathrm{i}. \tag{B10}$$

We should note that both constants are present epoch values of unknown functions of time, but here we just look for a relation between initial and final values. This relation is obtained for $\nu_\mathrm{i} = 3$ only, but since that is roughly the average value for our model cluster sample, a linear relation for $x(t_0)$ seems to be a good approximation for all ν_i.

We now assume a linear relation for $\nu(t_0)$ as well, i.e. $\nu(t_0) \approx a\nu_\mathrm{i} + b$, where again the fitting parameters are functions of t_0. This linear relation is a good approximation because we only consider a relatively small range of ν_i, so that the non-linear relation between $\nu(t_0)$ and ν_i can locally be well approximated by a linear function.

Combining both linear relations we can write the expected final total mass as

$$M_\mathrm{G}(\nu_\mathrm{i}, x_\mathrm{i}, t_0) = (2\pi R_\mathrm{G}^2)^{3/2}\bar{\rho}\left[c_0 + c_1\nu_\mathrm{i} + c_2 x_\mathrm{i}\right]. \tag{B11}$$

The constant c_1 will be larger than one, as $\nu(t)$ is a monotonically growing function. On the other hand, we expect c_2 to be negative on the basis of equation (B10).

APPENDIX C: RELIABILITY OF THE CLUSTER MODELS

Our models could be susceptible to various systematic effects and errors. The first source of error that comes to mind is the N-body code itself, since the integration of the particle trajectories is performed with a certain accuracy. However, the energy is conserved sufficiently well for all models: about 1 per cent change in $E/U(0)$ (the total energy divided by the initial total potential energy). The shape of the initially spherical distribution of particles changes little and its borders do not move inwards. This means that we have indeed included the complete region of influence of the cluster, i.e. the turn-around radius is well within the simulation volume. The autocorrelation function lacks some of the large-scale power at late times, but that is only the case for σ_8 larger than the value of 0.46 that we adopt. However, we will need to take larger simulation volumes for scenarios with more power on large scales, like the CDM $\Omega_0 = 0.2$ scenario.

Systematic errors are most likely due to the approximate way in which we model the formation and evolution of galaxies. For example, mass loss of galaxies is not modelled, and dynamical friction may therefore be unrealistically large. However, the galaxy formation parameters were chosen such that only the cores of particle groups that were found to be virialized were transformed into a single galaxy particle (see also van Kampen 1995). This allows the remaining outer particles to be stripped, even though this need not necessarily happen. The crude way in which merging is modelled is probably the biggest defect. Merging may therefore be underestimated in some regions and overestimated in others. In future work we will try to add a separate merger criterion, even though this will introduce more parameters to the modelling.

This paper has been typeset from a TEX/LATEX file prepared by the author.