Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 12 March 1999 (MN LaTeX style file v1.4)

The forest of merger history trees associated with the formation of dark matter halos

Ravi K. Sheth[1] and Gerard Lemson[1,2]
MPI-Astrophysik, K. Schwarzschildstr. 1, 85748 Garching, Germany
[1] *Max-Planck Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85740 Garching, Germany*
[2] *271 Dartmouth Street, Apt. 4F, Boston, MA 02116, USA*

Accepted 1999 January 25. Submitted 1998 December 17; in original form 1998 June 1

ABSTRACT
We describe a simple efficient algorithm that allows one to construct Monte-Carlo realizations of merger histories of dark matter halos. The algorithm is motivated by the excursion set model for the conditional and unconditional halo mass functions. The forest of trees constructed using this algorithm depends on the underlying power spectrum. For Poisson or white-noise initial power-spectra, the forest has exactly the same properties as the ensemble of trees described by Sheth. In this case, many ensemble averaged higher order statistics of the tree distribution can be computed analytically. For Gaussian initial conditions with more general power-spectra, mean properties of our forests closely resemble the mean properties expected from the excursion set approach. For these more general initial conditions, our algorithm shows how to write down simple, analytic approximations to some higher order statistical quantities associated with the forest. These higher order statistics generated using our algorithm, and the associated analytic approximations, are in good agreement with what is measured in numerical simulations of hierarchical gravitational clustering.

Key words: galaxies: clustering – cosmology: theory – dark matter.

1 INTRODUCTION

It is widely believed that the massive dark matter halos which exist today have grown from small initially Gaussian density fluctuations. Many questions of astrophysical interest can be addressed using Monte–Carlo realizations of the merger histories of such massive dark matter halos. An efficient method for generating Monte-Carlo realizations of the clustering process is essential for addressing such questions (e.g. Kauffmann & White 1993). Ideally, such an algorithm for generating merger history trees should produce results that are consistent with other known properties of the clustering process. If the algorithm also provides additional insight into the evolution of the clustering process, so much the better.

The excursion set approach of Bond et al. (1991) and Lacey & Cole (1993) provides analytic expressions for the mean number of halos of mass m that are later incorporated into a larger halo of mass $M > m$. These expressions are in good agreement with the corresponding quantities measured in numerical simulations of hierarchical clustering (Lacey & Cole 1994). Unfortunately, although the second and higher order moments of the subclump distribution are relatively easy to measure in the numerical simulations, the excursion set approach does not provide expressions for these higher order moments, for arbitrary Gaussian initial conditions. This means that the excursion set approach does not allow one to compute the higher order moments of the merger history tree (see, e.g., Fig. 1 in Sheth & Pitman 1997), which can then be compared with the simulation results.

For the special case of clustering from Poisson, or white-noise initial conditions, a model for the higher order moments of this subclump distribution exists (Sheth 1996; Sheth & Pitman 1997). Section 2 describes an algorithm which generates forests of trees that are guaranteed to have these higher order moments. The detailed proofs showing why the algorithm works are presented in Appendix A. It shows that these higher order moments are consistent with the fact that mutually disconnected regions in Poisson or white-noise Gaussian distributions are mutually independent.

To generate the forest of trees associated with arbitrary Gaussian initial conditions, we discuss two simple modifications of our white-noise algorithm. Whereas one provides a reasonable approximation to the known excursion set mean values, the other reproduces them exactly. This means that our algorithms produce forests which, in the mean, are similar to the forests which grow in numerical simulations of hierarchical gravitational clustering. We then compare the predicted higher order moments of our two algorithms to those measured in the simulations. Section 3 shows that, for a range of initially scale-free power spectra of current interest, the higher order distributions generated by one of our algorithms are, indeed, similar to those measured in the numerical simulations. This, therefore, is our preferred algorithm, and we show that the agreement with simulations is good even when the initial conditions are rather different from white noise.

A final section summarizes our results, and also discusses how our algorithm differs from those proposed by other authors. For example, although our algorithm is not of binary split type, trees constructed using binary splits may provide good approximations to ours. However, in contrast to other algorithms, our algorithm shows clearly how to write down analytic approximations to the higher order moments of the subclump distribution. This means that, in addition to allowing one to generate forests of merger history trees that are similar to those measured in numerical simulations, our results are also useful for providing an approximate analytic description of the evolution of the spatial distribution of dark matter halos that should also be reasonably accurate. This is, perhaps, the main virtue of our algorithm.

2 THE PARTITION ALGORITHM

2.1 Poisson initial conditions

Sheth (1996) describes an analytic model for the forest of merger history trees of halos of any given mass, identified at any given time, if the initial distribution was Poisson. Appendix A shows that this model is consistent with the assumption that mutually disconnected volumes are mutually independent. This means that it is straightforward to derive an algorithm for partitioning halos into subclumps.

Imagine partitioning an M-clump up into subclumps by choosing first one subclump, then choosing a second from the mass that remains, and so on until all the mass M has been assigned to subclumps. Start by choosing a random one of the M particles. On average, the chosen particle will be a member of an m_1-subclump with some probability $f(m_1, b_1|M, b_0)$, where $m \le M$ and $b_1 \le b_0$, where $f(m|M)$ is given by equation (A5). So choose m_1 by drawing a random number from this distribution. As discussed in Appendix A, the halo M can be thought of as being a region of size $V_M = Mb_0/\bar{n}$, where $\bar{n}$ denotes the average density, and $b_0 \equiv 1/(1+\delta_0)$, where δ_0 denotes the average overdensity within V_M.

If the first subclump has mass m_1, then it is associated with a volume $V_{m_1} = m_1 b_1/\bar{n}$ within V_M, the average density within V_{m_1} being greater than that within V_M. The remaining $R_1 \equiv (M - m_1)$ particles must be distributed randomly (because the initial distribution is Poisson) within the remaining volume $V_M - V_{m_1}$. Let $\delta^{(1)}$ denote the average density in this remaining volume. The overdensity in this volume can be represented by some $b^{(1)} \equiv 1/(1+\delta^{(1)})$, where

$$\bar{n}/b^{(1)} \equiv (M - m_1)/(V_M - V_{m_1}) \qquad (1)$$

Now the second subclump of the (M, b_0)-halo must be chosen from among the remaining $R_1 = M - m_1$ particles. By the same argument that led to the first choice, the probability that this second subclump has exactly m_2 particles should be $f(m_2, b_1|R_1, b^{(1)})$. Therefore, choose m_2 by drawing a random number from this distribution.

If the second subclump has mass m_2, then there are $R_2 \equiv (M - m_1 - m_2)$ remaining particles, they are randomly distributed within $(V_M - V_{m_1} - V_{m_2})$, where $V_{m_2} = m_2 b_1/\bar{n}$, and the average density within this remaining region is parametrized by some $b^{(2)}$, defined analogously to equation (1). Therefore, choose m_3 by drawing from the distribution $f(m_3, b_1|R_2, b^{(2)})$. This can be continued until all the particles have been assigned to subclumps.

In summary, to partition an (M, b_0)-halo into b_1-subhalos, proceed as follows.

```
Initialize:  specify b1, b0, M, and set R = M.
  Set b = M(b0-b1)/R + b1.
  Choose m with probability f(m,b1|R,b).
  Set R = R-m.
Iterate until R = 0.
```

The list of m's so obtained gives one possible partition of M into b_1-subclumps. Repeat to generate an ensemble of such partitions. Appendix A shows that this algorithm produces partitions with the same ensemble properties as Sheth's (1996) branching process.

Notice that, other than requiring that $0 \le b_1/b_0 \le 1$, there is no other constraint on b_1/b_0. This means that the algorithm constructs partitions with the correct ensemble properties in one time step. However, one is often interested in following the history of a halo as it becomes partitioned into smaller pieces through many small time steps. It is simple to use this partition algorithm to construct such a merger history tree. Given M and b_0, choose $b_1 < b_0$. Use the algorithm above to partition M into b_1-subclumps. Call this partition $\boldsymbol{m}$. Choose $b_2 < b_1$. Partition each subclump m of $\boldsymbol{m}$ into its constituent subclumps using the algorithm above, with the initial parameters $b_1 \to b_2$, $b_0 \to b_1$, $M \to m$ and set $R = m$. Then choose $b_3 < b_2$, and so on. The results of Sheth & Pitman (1997) and Sheth (1998) show that this many-step algorithm is guaranteed to give the same ensemble of partitions as the one-step algorithm described in detail above.

2.2 White noise Gaussian initial conditions

The partition algorithm of the previous section can be generalized to partition halos associated with Gaussian initial fluctuations.

Let $\delta(t)$ denote the critical density required for a spherical perturbation to collapse by time t (cf. equation A4). Thus, δ is both an overdensity variable and a pseudo-time variable, as was b in the previous section. Also, let $\sigma^2(M)$ denote the variance in the initial density fluctuation field when smoothed over regions that contain mass M on average. In what follows, we will sometimes write $S(M)$ instead of $\sigma^2(M)$. The dependence of the variance on M is specified by the power spectrum; we will only be interested in initial power spectra for which $S(M)$ decreases monotonically as M increases (e.g. Press & Schechter 1974).

In the excursion set approach (Bond et al. 1991; Lacey

& Cole 1993), the fraction of mass that is in halos which have mass in the range dm about m at cosmological time t is

$$f(m,\delta)\,\mathrm{d}m = \frac{\sqrt{\nu}\,\mathrm{e}^{-\nu}}{\sqrt{\pi}}\,\frac{\mathrm{d}\nu}{\nu}, \tag{2}$$

where $\nu = \delta^2/2S$, $S = \sigma^2(m)$ and δ decreases as cosmological time increases, as described above. The average number density of such halos is

$$n(m,\delta)\,\mathrm{d}m = (\bar{\rho}/m)\,f(m,\delta)\,\mathrm{d}m, \tag{3}$$

where $\bar{\rho}$ denotes the average density.

Similarly, consider a halo that has mass M at time δ_0. Let $f(m,\delta_1|M,\delta_0)\,\mathrm{d}m$ denote the fraction of the mass of this halo that was in subhalos of mass m at the earlier time $\delta_1 > \delta_0$. Then $f(m|M)\,\mathrm{d}m$ is given by equation (2), but with $\nu \to (\delta_1-\delta_0)^2/2(s-S)$, where $s = \sigma^2(m)$, and $S = \sigma^2(M)$. Thus, the mean number of m halos within an M halo is

$$N(m,\delta_1|M,\delta_0)\,\mathrm{d}m = (M/m)\,f(m,\delta_1|M,\delta_0)\,\mathrm{d}m \tag{4}$$

(e.g. Lacey & Cole 1993). This means that $f(m,\delta_1|M,\delta_0)$ can be written as $f(m|M,D)$, where $D = (\delta_1-\delta_0)$. Just as in the Poisson case, M occupies a region V, such that $M/V = \bar{\rho}(1+\delta_0)$, where $\bar{\rho}$ is the background density. Similarly, the subhalo m occupies v within V, with $m/v = \bar{\rho}(1+\delta_1)$. Thus, the density in the remaining volume is

$$\frac{M-m}{V-v} = \bar{\rho}(1+\delta'), \quad \text{so} \quad (\delta_1-\delta') = \frac{(\delta_1-\delta_0)}{1-(m/M)} \tag{5}$$

to lowest order in the δ terms. Finally, since disconnected volumes are mutually independent in the white-noise case, just as in the Poisson case, the same partition algorithm that worked there will also work here:

```
Initialize:  specify d1, d0, M, and set R = M.
  Set D = (d1-d0)/(R/M).
  Choose m with probability f(m|R,D).
  Set R = R-m.
Iterate until R is as small as desired.
```

A simple transformation of equation (2) shows that to choose m according to $f(m|R,D)$ one needs to generate a Gaussian random variable. There are efficient ways to do this, so this algorithm is fast.

This algorithm allows one to generate partitions of M whose distributional properties are described by the white-noise analogue of equation (A7). As before, there is no restriction on the size of the 'time step': $(\delta_1-\delta_0) > 0$ can be set as large as desired (although the algorithm is slower for larger time steps, since, for large $\delta_1-\delta_0$ most partitions are into many tiny pieces). And again, constructing a merger history tree by embedding this partition algorithm in a loop over time-steps is straightforward. In either case, the algorithm generates subclump distributions whose mean values are exactly the same as those required by the excursion set approach (equation 4). Moreover, this partition algorithm shows explicitly that the higher order moments of the subclump distribution are easy to compute. For example, the cross-correlation between m_1 and m_2 halos that are within the same M halo is simply the product

$$c(12|0) = N(m_1,\delta_1|M,\delta_0)\,N(m_2,\delta_1|M-m_1,\delta'), \tag{6}$$

where $N(m|M)$ was defined by equation (4) and δ' was defined by equation (5). For white-noise $c(12|0) = c(21|0)$. These higher order moments are the same as those one obtains by applying Stirling's approximation to the corresponding Poisson expressions, and taking the $\delta \ll 1$ limit.

2.3 Arbitrary Gaussian initial conditions

The assumption that disconnected volumes are mutually independent is wrong when the initial conditions differ from white-noise. Therefore, it is not obvious that the algorithm above should be used to generate partitions for arbitrary initial conditions. On the other hand, for arbitrary initial conditions, nothing is known about the higher order moments of the subclump distribution; only the mean is known from the excursion set approach. Below, we take two different approaches to constructing a partition algorithm for initial conditions that are not white-noise. Which algorithm to use is determined by comparison with numerical simulations in Section 3.

The first approach is motivated by the observation that, when expressed as functions of the variance rather than the mass, all excursion set quantities are independent of power spectrum. This suggests that, with a suitable transformation, we should be able to use the algorithm above, since it is known to be exact for the white-noise spectrum. We do this as follows. Run the algorithm assuming a white-noise power-spectrum. However, treat each chosen μ not as a subclump having mass $m = \mu$, but as a region of volume $v/V = \mu/M$ containing mass μ, that is populated by some number ν of objects that all have the same mass m. Thus, $\nu = \mu/m$. The value of m is got by requiring that $\sigma^2(m)$ equals the white-noise value, namely $\sigma^2(m) = 1/\mu$. To illustrate, suppose that $\sigma^2(m) = m^{-\alpha}$. Then ν is $\mu\,\sigma^{2/\alpha}(m) = \mu^{(\alpha-1)/\alpha}$. For white noise, $\alpha = 1$, so $\nu = 1$; thus, the region v contains exactly one halo as in the previous subsection. For general α, ν is neither unity nor even integer. Nevertheless, this approach generates partitions which are guaranteed to have the excursion set value in the mean, but which, as a result of the fact that m halos always come in groups of ν, may have unnaturally high values for the higher order moments. Since this approach simply uses the white-noise algorithm, the higher order moments can be calculated analytically, essentially by transforming the white noise expressions. For example, if $\sigma^2(m) = m^{-\alpha}$, then the cross-correlation between halos is

$$C(12|0)\,\mathrm{d}m_1\mathrm{d}m_2 = \left(\frac{m_1}{M}\frac{m_2}{M}\right)^{(\alpha-1)} c(12|0)\ \mathrm{d}\mu_1\mathrm{d}\mu_2 \tag{7}$$

where $c(12|0)$ is given by equation (6) but with the white-noise power spectrum.

The second approach is to assume, with no real justification, that the algorithm of the previous subsection can be used directly. In this case, the ensemble of partitions will differ for different power-spectra only because $\sigma^2(m)$ and $\sigma^2(M)$ in $f(m|M)$ depend on the power-spectrum. Since, in general, disjoint volumes are not independent, this means, of course, that there is no longer any guarantee that the algorithm will generate the excursion set mean values, either when used as a one step partition algorithm, or when the partition algorithm has been embedded in a loop over timesteps. Another way to see this is to note that $f(m|M)$ factors conveniently (analogously to equation A22 of the Poisson case) only for a white-noise power-spectrum. Since there is no guarantee that the algorithm will produce the excursion set statements in the mean, we also have no real analytic estimate for the higher order moments associated with this algorithm. Later, when we show the results using either the ensemble of one-step partitions or multi-step merger trees associated with this approach, we will compare the Monte-Carlo results with the formulae given in the previous subsection (4 and 6), but with the relation $\sigma^2(M)$ depending on power spectrum.

2.4 Monte-Carlo realizations of the forest

In this section, we compare various properties of the Monte-Carlo forests generated using our partition algorithms with the analytic formulae for the mean and higher order moments of the subclump distribution. Comparison with N-body simulations is the subject of the next section.

We show results for partitions associated with initially scale free power-spectra: $P(k) \propto k^n$. In this case, $\sigma^2(m) \propto m^{-\alpha}$, where $\alpha = (n+3)/3$. For such power-spectra, expressions like $N(m|M)$ and $c(mm|M)$ are functions of two parameters only: m/M and $(\delta_1 - \delta_0)/\sigma(M)$. It is convenient to define $\delta = \delta_c(1+z) = \sigma(M_*)(1+z)$, so that $(\delta_1 - \delta_0)/\sigma(M) = (z_1 - z_0)\,(M/M_*)^{\alpha/2}$. The curves below are plotted as functions of $(m/M)^\alpha$. They were produced for a range of choices for $(z_1 - z_0)$, with M set equal to M_*. The scaling above shows that by rescaling $(z_1 - z_0)$ appropriately, the same curves represent other values of M/M_*. For example, if $M = \gamma M_*$, then these same curves represent the subclump distribution at $(z_1 - z_0)/\gamma^{\alpha/2}$.

The previous section discussed two partition algorithms. In the white noise case, they are the same. Since the first algorithm discussed above works by simply transforming the white noise partition algorithm, if the analytic formulae (equations 4 and 7) describe the Monte Carlo $n = 0$ results well, then they will also work when $n \neq 0$. Therefore, in what follows, we have chosen to show the results we obtained using the second of our partition algorithms, where we also show the $n = 0$ results.

Fig. 1 shows results for initial power spectra with slopes $n = 0$, -1, and -2. In all panels, the thick solid curves show the analytic formulae with $M = M_*$ and $(z_1 - z_0) = 0.01$, 0.03, 0.1, 0.3, 1 and 3. Symbols show the corresponding quantities, averaged over 4,000 random partitions of M, generated using the second algorithm described above. The different panels in each figure show the cumulative quantities $F(>m|M)$, $N(>m|M)$, and $\mathrm{Var}(>m|M)/N(>m|M)$. Both $F(>m|M)$ and $N(>m|M)$ are got by integrating the excursion set equations from m to M. $\mathrm{Var}(>m|M)/N(>m|M)$ is the variance in the number of subclumps more massive than m, divided by the mean, $N(>m|M)$; if the number of subclumps were Poisson distributed among the halos, this quantity would be unity. The solid curve for $\mathrm{Var}(>m|M)/N(>m|M)$ in Fig. 1 is got by integrating equation (6) over the relevant range in m_1 and m_2, doing the necessary integral over (4) to get the mean $N(>m|M)$, so computing the variance, and then dividing by the mean. For $n = 0$, the symbols are well fit by our analytic formulae. The fits to the mean excursion set quantities (top two panels) are excellent, as is the fit to the second order moment (bottom panel). This shows that the algorithm works. As discussed above, this means that we are able to describe the results of the first algorithm analytically for all power spectra.

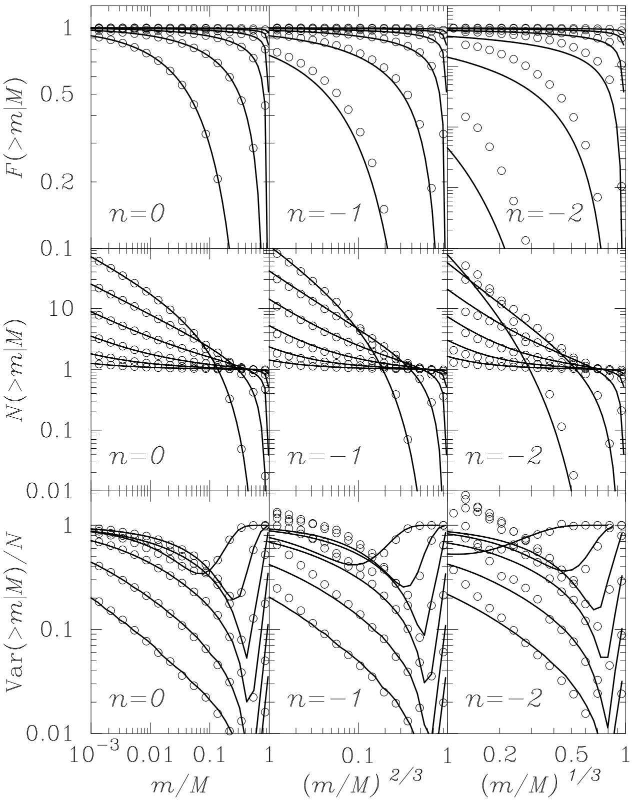

Figure 1. Results from using the second algorithm described in the text to partition an $M = M_*$ halo into subclumps at an earlier time, specified by $(z_1 - z_0) = 0.01$. 0.03, 0.1, 0.3, 1 and 3. In the top and middle panels, curves that are higher on the right correspond to smaller values of $(z_1 - z_0)$. The trend is the opposite in the bottom panel. Symbols show the Monte-Carlo results, curves show analytic formulae for comparison.

The ratio of the variance to the mean almost always has a minimum—this is a consequence of mass conservation. Consider the range $m/M > 0.5$ for a given redshift interval $z \equiv z_1 - z_0$. Let $p(n,z)$ denote the probability that M has n subclumps in this range. Mass conservation requires that there be at most one subclump in this range. Therefore $p(n,z) = 0$ for all $n > 1$, so the mean is $\mu_1 = p(1,z)$, the variance is $\mu_2 = p(1,z) - p(1,z)^2$, and the ratio μ_2/μ_1 is $1 - p(1,z)$, which is always less than unity. Since $p(1,z)$ refers to all subhalos that are more massive than m, it will certainly not decrease as m/M decreases from 1 to 0.5 (since decreasing m/M corresponds to increasing the allowed range of subhalo masses). Therefore, over this range, μ_2/μ_1 will decrease as m/M decreases. The rate of decrease will depend on the redshift interval, since, for a given $m/M > 0.5$, $p(1,z)$ will decrease as z increases. This just reflects the fact that it is much less likely that $m/M \approx 1$ at, say, $z = 5$ than at, say, $z = 0.5$. When $m/M \ll 1$, then the restriction implied by mass conservation becomes less severe, so the scatter in the subhalo counts can increase, and may even exceed the Poisson value. Fig. 1 shows that, for white noise, this Poisson value is never exceeded. The ratio sometimes exceeds unity for other power-spectra, at small masses. For the first of our algorithms, at least, this is not surprising. In that case, when $n < 0$, then subclumps of mass m always come in groups of $\nu > 1$, so it is as though less massive clumps are clustered,

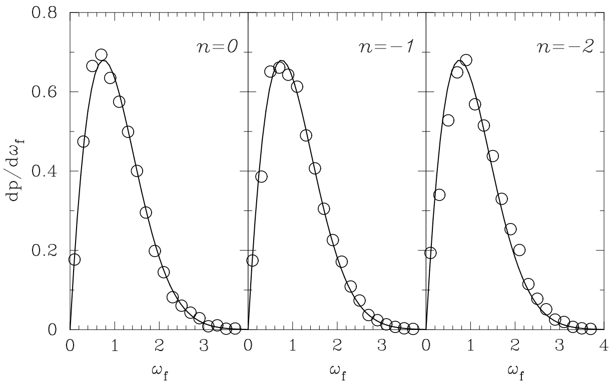

Figure 2. The scaled distribution of formation times. Open circles show the distribution generated using our merger tree algorithm, solid lines show equation (8), originally derived by Lacey & Cole (1993). The solid lines provide good fits to the distribution measured in numerical simulations (Lacey & Cole 1994).

and so the variance in the counts of these clumps exceeds the mean.

However, the figure shows that when $n \neq 0$, then our second algorithm does not reproduce the excursion set mean values exactly, though for $m/M \geq 0.01$ or so the agreement bewteen the symbols and the curves is fair. The algorithm generates partitions with slightly too many massive subclumps relative to less massive subclumps. This means that the discrepancy between the algorithm and the theoretical curves (when $n \neq 0$) is more obvious in the plots of $F(>m|M)$ than in those of $N(>m|M)$. Nevertheless, over this mass range, the variance given by equation (6) provides a fairly good description of the variance in the Monte-Carlo ensemble.

Embedding this algorithm into a loop over small time steps to generate ensembles of merger history trees gives very similar results. Since the difference between the symbols and the curves in Fig. 1 depends on $(z_1 - z_0)$, we might reasonably expect that the Monte-Carlo trees generated by looping over timesteps will depend on the choice of Δz. We have explored this possibility: for the range $0.001 \leq \Delta z \leq 0.05$, this dependence is negligible. This is reassuring, particularly because this means that, at least over the mass range $m/M \geq 0.01$ or so, equation (6) provides a good analytic approximation to the variance of the subclump distribution.

3 COMPARISON WITH NUMERICAL SIMULATIONS

This section compares the ensemble of trees generated by embedding the second of our two partition algorithms in a loop over time steps with that measured in numerical simulations of clustering from scale-free initial conditions. First we argue that the distribution of halo formation times generated by our merger tree algorithm is in reasonable agreement with what was measured by Lacey & Cole (1994) in their simulations. Then we consider a variety of other statistical tests, and compare the results generated using our merger tree algorithm with those which occured in scale-free simulations kindly made available by S. White.

Suppose we consider the mass of the largest subclump of the largest subclump of the largest subclump and so on. Define the formation time $z_{\rm f}$ of an M_0 halo as the first time that this mass drops below half of M_0. Lacey & Cole (1993) showed that the distribution of formation times defined in this way, suitably rescaled, depends only weakly on the power spectrum. They showed that, for white-noise initial conditions, this scaled distribution is

$$\mathrm{d}p/\mathrm{d}\omega_{\rm f} = 2\,\omega_{\rm f}\,\mathrm{erfc}(\omega_{\rm f}/\sqrt{2}), \tag{8}$$

where $\omega_{\rm f} \equiv z_{\rm f}\,(M_0/M*)^{\alpha/2}/\sqrt{2^\alpha - 1}$. Lacey & Cole (1994) showed that this distribution provides a good fit to the scaled formation time distribution measured in numerical simulations of hierarchical clustering. The solid curves in each panel of Fig. 2 show this analytic formula. The open circles in Fig. 2 show the corresponding scaled distribution of formation times (averaged over 10^4 realizations of the merger tree) generated by our merger tree algorithm. The symbols are reasonably well described by the formula. Since the formula describes the simulation results reasonably well also, we conclude that our merger tree algorithm generates formation time distributions that are similar to those in simulations.

Next, we will consider a variety of other statistical tests; some are variants of those used by Kauffmann & White (1993), and others are variants of those used by Sheth (1996) and Sheth & Pitman (1997). The numerical simulations used here are the same as those studied by Mo & White (1996), where they are described in more detail. They are normalized so that the number of particles in an M_* halo is $M_*(a) = (a/\delta_{\rm c0})^{6/(n+3)}$, where a is the expansion factor since the initial time, and n is the slope of the initial power spectrum. We show results for $n = 0$ and $n = -1.5$ below. We will sometimes quote results in terms of M_*: for all figures below, we use $\delta_{\rm c0} = 1.7$, and we study the merger trees of M_0 halos identified at a time when an M_* halo contains 471 and 166 particles for $n = 0$ and -1.5, respectively.

In what follows we choose a minimum mass threshold m. At a given output time z_0 we identify all halos with mass $M \geq m$. In the remainder of this section, $z_0 \equiv 0$. At an earlier output time $z_1 > z_0$ we identify all subclumps of each halo identified at z_0 that, at z_1 are more massive than m. For each halo, we store the number of such subclumps as well as its square. Averaging over all z_0 halos that have the same M/m allows us to compute the mean and the variance of the cumulative subclump distribution. In the previous section we showed these quantities as functions of m/M as m varied at fixed M, for representative values of M and $z_1 - z_0$. In Figs. 3 and 4, we plot these quantities versus M/m for fixed m as M varies. In all the plots below, when $z_0 = 0$, then $m/M_* = 10/471$ for the $n = 0$ simulations, and $m/M_* = 5/166$ for $n = -1.5$.

Figs. 3 and 4 show the first two moments of the subclump distribution as a function of M/m. The histogram shows the number of halos in each mass bin that were found in the simulations. The thin solid curves in show the expected number computed using the unconditional mass function $n(M)\,\mathrm{d}M \propto f(M)\,\mathrm{d}M/M$ of equation (2). There are two reasons for showing this histogram. This first is to demonstrate that the unconditional mass function is in reasonable agreement with the simulations. The second is that the histogram is intended to give an idea of the mass range

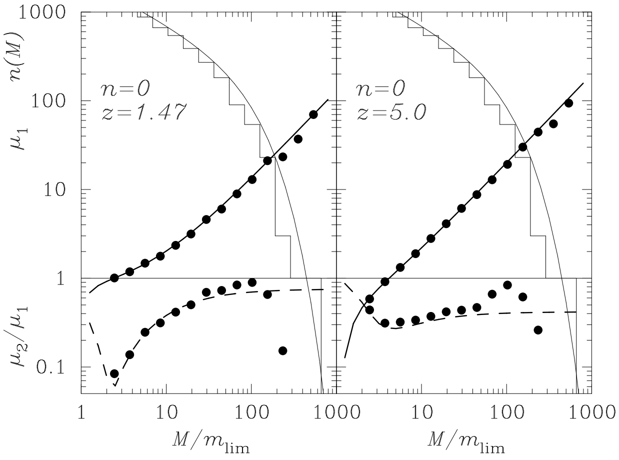

Figure 3. The mean μ_1, and the variance divided by the mean μ_2/μ_1, of the number of subclumps that, at z, have mass greater than $m = 10$ that are within a halo with mass M/m. Curves show the analytic results; for $n = 0$ these curves describe the results of our merger tree algorithm exactly. The filled circles show the corresponding quantities measured in the $n = 0$ simulations. The histogram shows the unconditional mass function measured in the simulations, and the thin solid curve shows the associated Press–Schechter mass function.

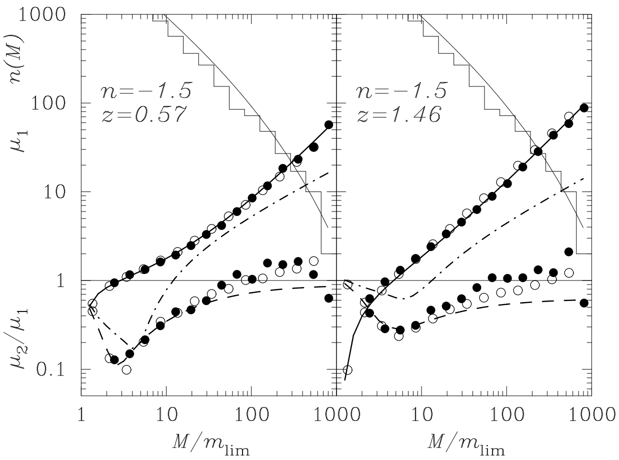

Figure 4. The same as the previous figure, but for $m = 5$ and $n = -1.5$. Filled circles show the simulation results, open circles show the corresponding quantities generated using our tree algorithm, curves show the analytic formulae. We have shown two curves for the variance; the dashed curve shows equation (6) and the dot-dashed shows equation (7). Whereas equation (7) describes the results of the first of our algorithms exactly, equation (6) fits our second algorithm exactly only when $n = 0$.

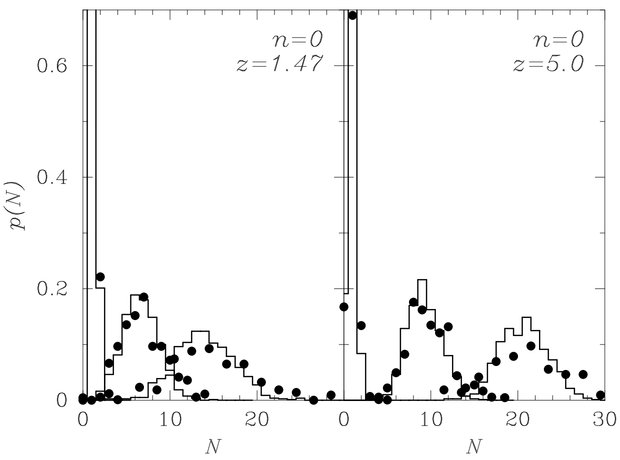

Figure 5. The distribution of the number of subclumps that, at z, have mass greater than $m/M_* = 10/471$ that are within a halo with mass $M_0/M_* = 38/471 = 0.08$ (left-most), $447/471 = 0.95$ (middle) and $1024/471 = 2.17$ (right-most). Histograms show the distribution measured using our algorithm, and filled circles show the corresponding distribution measured in the $n = 0$ simulations.

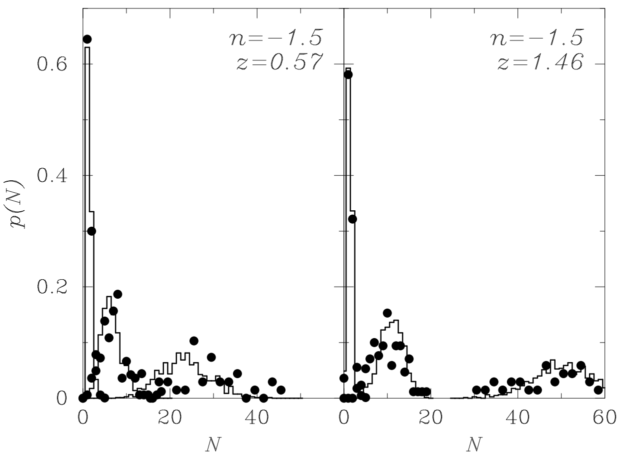

Figure 6. The distribution of the number of subclumps that, at z, have mass greater than $m/M_* = 5/166$ that are within a halo with mass $M_0/M_* = 28/166 = 0.17$ (left-most), $339/166 = 2.04$ (middle) and $1780/166 = 10.7$ (right-most). Histograms show the distribution measured using our algorithm, and filled circles show the corresponding distribution measured in the $n = -1.5$ simulations.

over which the halo sample in the simulations is approximately statistically complete (e.g., some of the statistical quantities measured in the simulations may not be accurate measures of the true quantities in the mass range where the number of halos in a given mass bin is small.)

The thick solid curves show the conditional mass function $\mu_1 \equiv N(> m|M)$ computed by doing the necessary integral over equation (4). The thick dashed curves show $\mu_2/\mu_1 \equiv \mathrm{Var}(> m|M)/N(> m|M)$ computed using equation (6), and the thick dot–dashed curves show this same quantity computed using equation (7). (For $n = 0$ these two curves overlap exactly.) Recall that the curves in Figs. 1 depend on the combination $(z_1 - z_0)^2 S_*/S$. Although the curves there were made for $M = M_*$ halos, if M/M_* is different, then the same curves represent a different value of $z_1 - z_0$ for scale-free initial conditions. So, with a little thought, one can convince oneself that the curves there are consistent with those shown here.

The filled circles show the corresponding quantities measured in the numerical simulations. Open circles show these quantities generated by our merger history algorithm. When $n = 0$, the tree results are fit by the analytic formulae exactly, so we only show the theoretical curves. When $n = 0$, these curves fit the simulation results reasonably well. When $n = -1.5$, the merger tree algorithm is not fit exactly by the analytic curves, and the two analytic curves for the variance differ from each other. Figure 4 shows that the numerical simulation results are fit better by equation (6) than by (7), and still better by the results generated by our algorithm.

Figs. 5 and 6 show the distribution of the number of progenitors of an M_0 halo that, at z, are more massive than m. The histogram shows this quantity generated by averaging over 800 realizations of the merger history tree using our algorithm, and solid points show the corresponding quantity measured in the simulations. For both values of n studied, three representative values of M_0 are shown: these approximately bracket the range $0.1 \le M/M_* \le 10$. For both values of n, the histograms generated by our algorithm are in reasonable agreement with the numerical simulation results. Whereas the previous figures showed that our algorithm produces forests with approximately the correct mean and second moment, these figures suggest that our higher order moments are also accurate.

Another measure of the higher order moments is shown in Figs. 7 and 8. They show the relative sizes of the largest subclumps produced by our algorithm and those measured in the simulations. The histograms show the distribution of the size of the largest subclump M_1/M_0 for the same values of M_0 as in the previous figures. (The topmost panels show the lowest values of M_0.) The symbols show the corresponding simulation results. The open circles in the top right show the mean of M_2/M_1 given M_1/M_0, where M_2 denotes the size of the second largest subclump, averaged over 800 realizations of the merger tree. The solid circles and error bars show the associated quantities measured in the simulations. For both values of n our algorithm is in good agreement with the simulations, though the agreement when $n = 0$ is somewhat better than when $n = -1.5$.

We conclude this section with the observation that, for most of the tests we considered, our algorithm produces forests of merger history trees that are reasonably similar to those which occur in numerical simulations of hierarchical clustering.

4 DISCUSSION

We described an algorithm which allowed us to partition halos at a given time into subhalos at an arbitrary earlier time. Our algorithm is exact for haloes that form from Poisson or white-noise Gaussian initial conditions. In these cases, analytic expressions for the higher order moments of the subclump distribution are available (Sheth 1996). We discussed two possible ways of modifying our algorithm to partition halos which form from more general Gaussian initial conditions. This showed how to write down simple analytic approximations to the associated higher order moments of the subclump distribution that were reasonably accurate even for these more general initial conditions.

We showed that it was simple and self consistent to embed our partition algorithm into a loop over time steps, and so to generate a forest of merger trees. We compared this forest of trees with the ensemble of trees which grew in numerical simulations of hierarchical gravitational clustering from a range of scale-free Gaussian initial conditions. For the initial conditions we studied, the agreement between our trees and those of the simulations was fair. This comparison also showed that our analytic formulae for the higher order moments of the subclump distribution provide reasonable fits to the numerical simulation results, even when the initial conditions are different from white noise. For example, the second order moment is well described by our equation (6).

Our approach falls within the framework of the approach pioneered by Press & Schechter (1974) and extended

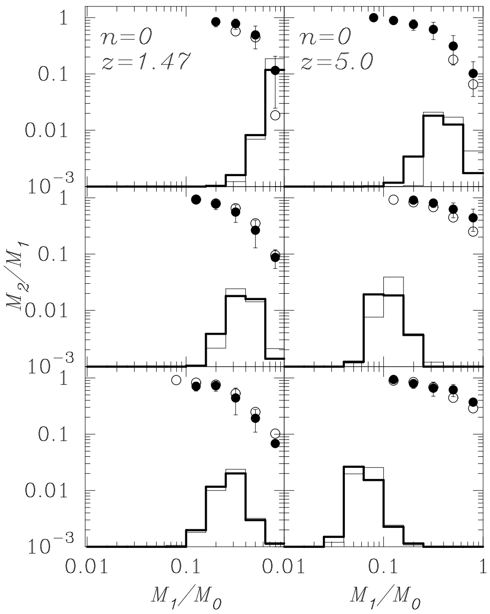

Figure 7. The relative sizes of the first and second largest subclumps, M_1/M_0 and M_2/M_1, for the same three values of M_0 as in Fig. 5. Thin histograms show the distribution of M_1/M_0 measured using our algorithm, and bold histograms show the corresponding distribution measured in the $n = 0$ simulations. Open circles in the top right show the mean of M_2/M_1 given M_1/M_0 generated by our algorithm; filled circles show the corresponding mean and rms scatter measured in the simulations.

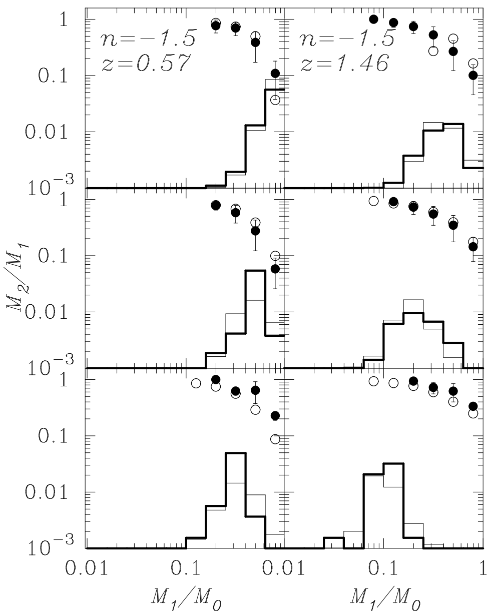

Figure 8. Same as the previous figure, but for $n = -1.5$, and for the same three choices of M_0 shown in Fig. 6.

by Bond et al. (1991) and Lacey & Cole (1993). It is predicated on the view that, when combined with the spherical collapse model, the statistical properties of the initial conditions reduce the problem of partitioning a halo's mass among its constituent subclumps to one of combinatorial statistics. The assumption is that, at this level of detail, the more complicated physics of mass transfer and disruption may safely be ignored. The agreement with the simulation results suggests that the more complicated physics which our approach ignores is, for the most part, irrelevant.

Although it falls within the general excursion set framework of Bond et al. (1991) and Lacey & Cole (1993), our algorithm differs from other merger history tree algorithms in several ways. Firstly, ours is really a partition algorithm: in principle, there is no limit on the size of the time-step between when a halo is identified and when its subclump distribution is studied. In this respect, it is similar to the algorithm proposed by Kauffmann & White (1993). In the Poisson case, the higher order moments generated by our algorithm differ from theirs (Sheth & Pitman 1997), so it is likely that these higher order moments are also different for more general initial conditions.

To illustrate, Fig. 9 shows this difference for the white noise case. It is similar in style to Fig. 1; the symbols show the result of using their algorithm to construct the merger trees, and the curves show the analytic formulae described in the text–recall that they describe our algorithm exactly. At high redshift, the Kauffmann–White algorithm produces values of the variance that are well in excess of the Poisson value. This is not seen in the simulations (Fig. 3). We have not tried to quantify the difference between their model and ours for other power spectra, since we have already shown that ours produces forests that are similar to those in the numerical simulations. Moreover, our algorithm is simpler, faster, and has more physical as well as statistical motivation.

Binary split algorithms based on the excursion set approach have been proposed by Lacey & Cole (1993) and by Sheth & Pitman (1997). In the limit $\delta_1 - \delta_0 \equiv \delta_{10} \ll 1$, our one-step partition algorithm produces partitions in which most of the mass is in the largest two pieces. To see this, recall that each choice in the partition process is of some $s - S(R) = (\delta_{10}/x)^2$, where x is a Gaussian random variable with zero mean and unit variance. Since x is Gaussian, most of the time $s - S(R) \approx \delta_{10}^2$. If $\delta_{10} \ll 1$, then $s \approx S(R)$ after most draws of the random number x. This means that, most of the time, the first choice is of a subhalo with most of the mass of M. Iterating this argument, the second choice is of a subhalo which contains most of the remaining mass. Since a given range in $s - S$ correponds to a smaller mass range for white-noise than for power-spectra with more negative slopes, the above is more true for white-noise than for these other power-spectra. Therefore, at least for white-noise, a binary split algorithm, with $f(m|M)$ for the split rule, and with succesive splits spaced at $\Delta\delta = \delta_{10} \ll 1$, should provide a good approximation to the merger trees produced here.

A different approach to just such a binary split algorithm is described by Sheth & Pitman (1997). They showed that, in the white noise case, the binary split algorithm described in the previous paragraph will generate forests that are exactly the same as those generated using our algorithm. They also discuss why binary split algorithms that are consistent with the excursion set approach are easier to construct in the white-noise case than in the general case. [The binary split algorithm proposed by Lacey & Cole (1993) is different from the one described above. However, as they showed, their algorithm fails to produce the correct distribution of halo formation times.]

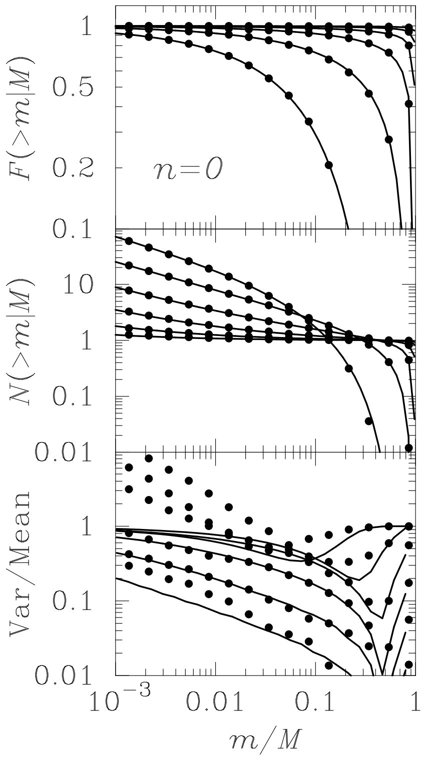

Figure 9. Same as Fig 1, but now symbols show results obtained using the Kauffmann–White algorithm for $n = 0$ initial conditions. The solid curves describe the results of our algorithm exactly. The curves do not fit the higher redshift symbols–the scatter among partitions generated by our two algorithms is different. However, our algorithm fits the scatter measured in numerical simulations well.

One great virtue of our approach is that it shows clearly how to write analytic estimates of the higher order moments of the subclump distribution—it is not so easy to do this with the binary split approach. Analytic approximations to these higher order moments, that are also reasonably accurate, are extremely useful. This is because the higher order moments associated with the forest of merger history trees of dark matter halos can be related to the higher order moments of the spatial distribution of these halos (Mo & White 1996). Since the higher order moments associated with our merger tree algorithm, and also of the merger trees in the numerical simulations, were reasonably well fit by our equation (6), it is possible to write down analytic approximations to spatial quantities like the halo–halo correlation function that are also reasonably accurate. Such expressions greatly simplify the study of the spatial distribution of dark matter halos. This is the subject of ongoing work.

ACKNOWLEDGMENTS

Thanks to Jim Pitman for discussing his results on size–biased (see, e.g., Perman, Pitman & Yor 1992) distributions which, in the white-noise case, lead to this same partition algorithm (see Corollary 5 and Lemma 12 of Aldous & Pitman 1998). Many thanks also to S. Zaroubi and M. Haj for their hospitality when this project was nearing completion.

REFERENCES

Aldous D., Pitman J., 1998, Berk. Stat. Tech. Rep., 489, to appear in Ann. Prob.

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Borel E., 1942, Compt. rend., 214, 452

Epstein R. I., 1983, MNRAS, 205, 207

Kauffmann G., White S. D. M., 1993, MNRAS, 261, 921

Lacey C., Cole S., 1993, MNRAS, 262, 627

Lacey C., Cole S., 1994, MNRAS, 271, 676

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347

Perman M., Pitman J., Yor M., 1992, Prob. Th. Rel. Fields, 92, 21

Press W., Schechter P., 1974, ApJ, 187, 425

Sheth R. K., 1995, MNRAS, 276, 796

Sheth R. K., 1996, MNRAS, 281, 1277

Sheth R. K., Pitman J., 1997, MNRAS, 289, 66

Sheth R. K., 1998, MNRAS, 295, 869

APPENDIX A: THE POISSON MODEL

This Appendix demonstrates that the algorithm described in the main text produces forests of merger history trees that have the same ensemble properties as those of the branching process described by Sheth (1996). Section A1 sets notation by providing a summary of old results. These are used to derive new results in Sections A2 and A3.

A1 Excursion set and branching process results

The evolution of clustering from an initially Poisson distribution has been studied using the excursion set (Epstein 1983) and cloud-in-cloud formalisms (Sheth 1995). In these analyses, all matter is bound into clumps, sometimes called halos. The probability that a clump has m associated particles is

$$\eta(m,b) = \frac{(mb)^{m-1}\mathrm{e}^{-mb}}{m!}, \tag{A1}$$

where $m \geq 1$ and $0 \leq b \leq 1$. Equation (A1) is known as the Borel distribution (Borel 1942). The probability $f(m,b)$ that a randomly chosen particle belongs to an m-clump is given by multiplying $\eta(m,b)$ by m, and dividing by the mean clump size $\sum_m m\,\eta(m,b) = 1/(1-b)$:

$$f(m,b) = m\,(1-b)\,\eta(m,b). \tag{A2}$$

Since all matter is in clumps, and $f(m,b)$ is the fraction of mass in m clumps,

$$\sum_{m=1}^{\infty} f(m,b) = 1. \tag{A3}$$

The evolution of clustering is described by the evolution of the b parameter, which is the ratio of the average density in the universe to the critical overdensity that is required for an object to collapse and virialize at cosmological time t. That is,

$$b(t) = \frac{1}{1+\delta_{\rm c}(t)} \qquad \text{where} \quad \delta_{\rm c}(t) \propto \frac{1.686}{a(t)}, \tag{A4}$$

t denotes cosmological time, and $a(t)$ is the factor by which the universe has expanded since some fiducial initial epoch, say, t_i. Notice that the critical density decreases as the universe expands, so that b increases with expansion. Since b increases monotonically as cosmological time increases, it can also be treated as a pseudo-time variable. Thus, b plays two roles: it can be thought of as describing a spatial density, or it can be thought of as describing time. Therefore, $f(m,b)$ can be thought of as the fraction of the total mass that is in m-clumps at time b, or it can be thought of as the fraction of the total mass that is associated with regions within which the average overdensity is given by b.

The excursion set description also allows one to compute the probability that a randomly chosen particle of a clump having exactly M particles at the time corresponding to b_0 was a member of an m-clump at the earlier epoch $b_1 < b_0$. This probability is

$$\begin{aligned} f(m,b_1|M,b_0) &= M\left(1-\frac{b_1}{b_0}\right)\binom{M}{m}\frac{m^m}{M^M} \\ &\quad\times\left(\frac{b_1}{b_0}\right)^{m-1}\left[M-m\frac{b_1}{b_0}\right]^{M-m-1}, \end{aligned} \tag{A5}$$

where $1 \leq m \leq M$ and $0 \leq b_1/b_0 \leq 1$ (Sheth 1995). This implies that the average number of progenitor m-subclumps identified at b_1 that are later (at $b_0 > b_1$) in an M-clump is

$$N(m,b_1|M,b_0) = (M/m)\,f(m,b_1|M,b_0). \tag{A6}$$

Although the excursion set formalism provides information about the mean number of progenitor subclumps of any given parent clump, it does not provide a complete description of the entire merger history tree (cf. Fig. 1 in Sheth & Pitman 1997).

Let $p(\boldsymbol{n},b_1|M,b_0)$, with $\boldsymbol{n} = (n_1,\cdots,n_M)$, $\sum_{j=1}^M n_j = n$ and $\sum_{j=1}^M j\,n_j = M$, denote the probability that an M-clump at the epoch b_0 was previously in n subclumps at the epoch b_1, of which there were n_1 single particle subclumps, n_2 subclumps with exactly two particles, n_j subclumps with exactly j particles, and so on. Then the branching process model of Sheth (1996) yields the formula

$$p(\boldsymbol{n},b_1|M,b_0) = \frac{(Mb_{01})^{n-1}\mathrm{e}^{-Mb_{01}}}{\eta(M,b_0)}\prod_{j=1}^{M}\frac{\eta(j,b_1)^{n_j}}{n_j!}, \tag{A7}$$

where $b_{01} = (b_0 - b_1)$. Sheth & Pitman (1997) showed that this same formula follows from another construction, which they termed the additive coalescent. Appendix A2 below shows that this formula is consistent with the assumption that mutually disconnected regions within a Poisson distribution are mutually independent.

Notice that equation (A7) depends only on the ratio b_1/b_0, and not on the actual values of b_1 or b_0. This dependence on the ratio b_1/b_0 implies that the merger histories of M-clumps formed at the epochs b_0 and b_0' are related to each other simply by rescaling the time variable. By taking relevant sums over the partition probability function, one can show explicitly that the branching process expressions for the clump size distribution and the merger probabilities $f(m|M)$ are identical to the excursion set results (equations A1 and A5). However, the partition probability function also allows one to compute quantities that cannot be estimated using the excursion set formalism. For example, the excursion set approach says that the mean number of m-subclumps per M-clump is $(M/m)\,f(m|M)$. The partition probability function allows one to compute the higher order moments of the n_j distribution also. For example,

$$\left\langle \frac{n_i!}{(n_i-\alpha)!}\frac{n_j!}{(n_j-\beta)!}, b_1 \,\middle|\, M, b_0 \right\rangle$$
$$= \Big[Mb_{01}\Big]^{\alpha+\beta}\ \frac{\eta^\alpha(i,b_1)\,\eta^\beta(j,b_1)\ \eta(R,B)}{\eta(M,b_0)},\quad \text{if } R>0$$
$$= \Big[Mb_{01}\Big]^{\alpha+\beta-1}\ \frac{\eta^\alpha(i,b_1)\,\eta^\beta(j,b_1)}{\eta(M,b_0)},\quad \text{if } R=0, \qquad \text{(A8)}$$

where $b_{01}\equiv(b_0-b_1)$ and

$$RB \equiv (M-\alpha i-\beta j)B = Mb_0 - \alpha i b_1 - \beta j b_1. \qquad \text{(A9)}$$

When $\alpha=1$ and $\beta=0$ this is equivalent to the relation obtained from equation (A5). When $\alpha=\beta=1$, equation (A8) gives an estimate of the cross-correlation between i- and j-subclumps that are within the same M-clump.

A2 The partition probability function

This section describes a simple interpretation of the partition probability function (equation A7). It shows that this partition formula is consistent with the assumption that mutually disconnected volumes in a Poisson distribution are mutually independent.

Since b_0 can be related to a density (equation A4), an (M,b_0)-halo is associated with a region in the initial Poisson distribution that has size

$${}_0V_M = M/[\bar n\,(1+\delta_0)] = Mb_0/\bar n, \qquad \text{(A10)}$$

where $\bar n$ denotes the average density. This means that $p(n_1\cdots n_M, b_1|M,b_0)$ of equation (A7) denotes the probability that a region containing exactly M particles at average density $\bar n(1+\delta_0)$ has n subregions, each with average density $\bar n(1+\delta_1)$, where $\delta_1\ge\delta_0$, of which n_1 such subregions contain only one particle (so they each have size ${}_1V_1$), n_2 contain two particles (so they each have size ${}_1V_2$), etc. Notice that the sum of the sizes of these n sub-volumes is ${}_1V_M < {}_0V_M$, since $b_1\le b_0$. The remaining volume ${}_0V_M - {}_1V_M = M(b_0-b_1)/\bar n$ is empty.

Now suppose that the region ${}_0V_M$ containing M particles is known to contain n b_1-subregions, of which at least one such subregion has size ${}_1V_m$ and contains exactly m particles. The probability that ${}_0V_M$ contains n_1 b_1-regions of size ${}_1V_1$, n_2 b_1-regions of size ${}_1V_2$, etc., given that, within ${}_0V_M$, there is at least one b_1-subregion of size ${}_1V_m$ containing m particles, and $n-1$ other b_1-subregions is

$$\frac{(Mb_{01})^{n-1}\,\mathrm{e}^{-Mb_{01}}}{\eta(M,b_0)}\ \frac{\eta(m,b_1)}{\langle n_m,b_1|M,b_0\rangle}\prod_{i=1}^{M-m}\frac{\eta(i,b_1)^{n_i}}{n_i!}, \qquad \text{(A11)}$$

where $b_{01}=b_0-b_1$, $\sum_i n_i = n-1$, and $\sum_i i\,n_i = M-m$. The term

$$\langle n_m,b_1|M,b_0\rangle = \sum_{\text{all parts.}} n_m\,p(n_1\cdots n_M,b_1|M,b_0), \qquad \text{(A12)}$$

which is the sum over all partitions of M that have at least one m-part, is included in the denominator since this is a conditional probability. However, equation (A8) shows that

$$\langle n_m,b_1|M,b_0\rangle = Mb_{01}\,\frac{\eta(m,b_1)\,\eta(M-m,b')}{\eta(M,b_0)}, \qquad \text{(A13)}$$

where b' is defined by the relation

$$\frac{M-m}{{}_0V_M - {}_1V_m} \equiv \bar n(1+\delta') = \frac{\bar n}{b'}. \qquad \text{(A14)}$$

Thus, b' parameterizes the density in the remaining volume ${}_0V_M - {}_1V_m$. Substituting equation (A13) for $\langle n_m,b_1|M,b_0\rangle$ into (A11) yields

$$\frac{(Mb_{01})^{n-2}\,\mathrm{e}^{-Mb_{01}}}{\eta(M-m,b')}\prod_{i=1}^{M-m}\frac{\eta(i,b_1)^{n_i}}{n_i!}. \qquad \text{(A15)}$$

However,

$$(M-m)\,(b'-b_1) = M(b_0-b_1) = Mb_{01}, \qquad \text{(A16)}$$

so (A15) can be written as

$$\frac{[(M-m)(b'-b_1)]^{n-2}\mathrm{e}^{-(M-m)(b'-b_1)}}{\eta(M-m,b')}\prod_{i=1}^{M-m}\frac{\eta(i,b_1)^{n_i}}{n_i!}, \qquad \text{(A17)}$$

where $\sum_i n_i = n-1$, and $\sum_i i\,n_i = M-m$. Clearly, this is the same as

$$p(n_1\cdots n_{M-m}, b_1|M-m,b'), \qquad \text{(A18)}$$

where the n_is sum to $(n-1)$, and the sum of in_i is $(M-m)$.

This shows explicitly that if the region ${}_0V_M$ containing exactly M particles, so the average density within it is $M/{}_0V_M = \bar n(1+\delta_0)$, is known to contain exactly n subregions, each with average density $\bar n(1+\delta_1)$, and it is also known that one of these subregions has size ${}_1V_m$ and contains exactly m particles, so the remaining $M-m$ particles are arranged into $n-1$ subregions, each with the same overdensity δ_1, in the remaining volume ${}_0V_M - {}_1V_m$, then the distribution of sizes of the $n-1$ subregions (each with overdensity parameter b_1) within ${}_0V_M - {}_1V_m$ is given by the same partition probability function as for the total volume ${}_0V_M$, with trivial adjustments for the fact that the average density within the remaining volume ${}_0V_M - {}_1V_m$ will be different from that within ${}_0V_M$ itself, and for the fact that we need now only consider partitions of $M-m$ into $n-1$ parts, rather than of M into n parts. In other words, if $p(\boldsymbol{n},b_1|M,b_0)$ is viewed as describing the distribution of b_1-subregions within the region ${}_0V_M$, then it is fully consistent with the fact that non-overlapping volumes in a Poisson distribution are mutually independent.

This property of the partition formula could have been inferred directly from the additive coalescent interpretation of equation (A7) that is provided by Sheth & Pitman (1997).

A3 The partition algorithm

The partition algorithm described below follows from the fact that equation (A7) is consistent with the assumption that mutually disconnected regions within a Poisson distribution are mutually independent.

Imagine partitioning an M-clump up into subclumps by choosing first one subclump, then choosing a second from the mass that remains, and so on until all the mass M has been assigned to subclumps. Let m_i denote the number of particles in the ith subclump. Let R_i denote the mass remaining after the ith pick:

$$R_i \equiv M - \sum_{j=1}^{i} m_j, \quad \text{and} \quad R_0 \equiv M, \tag{A19}$$

and parametrize the average density in the volume that remains after the ith pick by $b^{(i)}$. Then

$$\frac{1}{b^{(i)}} = \frac{R_i}{\bar{n}V_M - \sum_{j=1}^{i} \bar{n}V_{m_j}} = \frac{R_i}{Mb_{01} + R_i b_1}. \tag{A20}$$

Notice that $1/b^{(0)} = R_0/(Mb_{01} + R_0 b_1) = 1/b_0$ as required, and that

$$R_i\,(b^{(i)} - b_1) = R_i\left(\frac{Mb_{01} + R_i b_1}{R_i} - b_1\right) = Mb_{01} \tag{A21}$$

is independent of i. In this notation, equation (A5) becomes

$$\begin{aligned} &f(m_i, b_1|R_{i-1}, b^{(i-1)}) \\ &= m_i\,(b^{(i-1)} - b_1)\,\frac{\eta(m_i,b_1)\,\eta(R_i,b^{(i)})}{\eta(R_{i-1},b^{(i-1)})}, \\ &= \frac{m_i}{R_{i-1}}\,Mb_{01}\,\frac{\eta(m_i,b_1)\,\eta(R_i,b^{(i)})}{\eta(R_{i-1},b^{(i-1)})} \quad \text{if} \quad 1 \le m_i < R_{i-1} \end{aligned}$$

and it is

$$\begin{aligned} &= \frac{\eta(m_i,b_1)}{\eta(R_{i-1},b^{(i-1)})}\,\mathrm{e}^{-R_{i-1}(b^{(i-1)}-b_1)} \\ &= \frac{\eta(m_i,b_1)}{\eta(R_{i-1},b^{(i-1)})}\,\mathrm{e}^{-Mb_{01}} \quad \text{if} \quad m_i = R_{i-1}. \end{aligned} \tag{A22}$$

The probability that after n picks all M particles have been assigned to a subclump (i.e., $m_n = R_{n-1}$ and $R_n = 0$), is simply the product of the individual n choices above. If $\boldsymbol{m}$ denotes the set $(m_1, \cdots, m_n)$, then

$$\begin{aligned} p(\boldsymbol{m}|M) &= \prod_{i=1}^{n} f(m_i, b_1|R_{i-1}, b^{(i-1)}) \\ &= \frac{(Mb_{01})^{n-1}\,\mathrm{e}^{-Mb_{01}}}{\eta(M,b_0)} \prod_{i=1}^{n} \frac{m_i}{R_{i-1}}\,\eta(m_i,b_1), \end{aligned} \tag{A23}$$

where the final expression uses the fact that $R_0 = M$, $b^{(0)} = b_0$, and $m_n = R_{n-1}$.

Equation (A23) describes the probability that M was partitioned into these n subclumps in the order $(m_1, \cdots, m_n)$. The probability that the sequence of picks resulted in the set $(m_1, \cdots, m_n)$, whatever the order of the m_i, is given by summing equation (A23) over all $n!$ permutations of the n picks m_i. Let $\pi[\boldsymbol{m}]$ denote this set of permutations. Then

$$\begin{aligned} \sum_{\pi[\boldsymbol{m}]} p(\boldsymbol{m}|M) &= \sum_{\pi[\boldsymbol{m}]} \frac{(Mb_{01})^{n-1}\,\mathrm{e}^{-Mb_{01}}}{\eta(M,b_0)} \\ &\quad \times \left(\prod_{i=1}^{n} \frac{m_i}{R_{i-1}}\right)\left(\prod_{i=1}^{n} \eta(m_i,b_1)\right), \end{aligned} \tag{A24}$$

since, for a given $\boldsymbol{m}$, all terms in each permutation are the same, except for the product $\prod_{i=1}^{n}(m_i/R_{i-1})$. So, the sum over permutations becomes

$$\begin{aligned} \sum_{\pi[\boldsymbol{m}]} p(\boldsymbol{m}|M) &= \frac{(Mb_{01})^{n-1}\,\mathrm{e}^{-Mb_{01}}}{\eta(M,b_0)} \left(\prod_{i=1}^{n} \eta(m_i,b_1)\right) \\ &\quad \times \sum_{\pi[\boldsymbol{m}]} \left(\prod_{i=1}^{n} \frac{m_i}{R_{i-1}}\right). \end{aligned} \tag{A25}$$

A little thought shows that the contribution from the terms in the sum is unity. Appendix B shows this explicitly. Thus,

$$\sum_{\pi[\boldsymbol{m}]} p(\boldsymbol{m}|M) = \frac{(Mb_{01})^{n-1}\,\mathrm{e}^{-Mb_{01}}}{\eta(M,b_0)} \prod_{i=1}^{n} \eta(m_i,b_1). \tag{A26}$$

Now, if all particles are indistinguishable, then subclumps that have the same number of particles are indistinguishable from each other, so these permutations should not be counted separately. We can account for this by dividing the product above by $\prod_{i=1}^{M} n_i!$, where n_i denotes the number of i-subclumps in the partition, $\sum_i n_i = n$, and $\sum_i i\,n_i = M$. Thus, the probability that M was partitioned into $\boldsymbol{n}$ is the same as that given by equation (A7).

Before concluding this section, we note that Sheth (1998) describes a model for clustering from compound Poisson initial conditions. For these more general initial conditions, the analogue of equation (A5) factors similarly to how it does for the Poisson case. This means that a similar algorithm for generating the associated forest of merger trees can be used there also.

APPENDIX B: SUMMING OVER PERMUTATIONS

Consider a partition $\boldsymbol{m}$ of M into n parts: $\boldsymbol{m} \equiv (m_1, \cdots, m_n)$, where $\sum_i^n m_i = M$. Let $R_i \equiv M - \sum_{j=1}^{i} m_j$. This Appendix shows that the sum over all permutations $\pi[\boldsymbol{m}]$ of $(m_1, \cdots, m_n)$, of $\prod_{i=1}^{n}(m_i/R_{i-1})$ is unity. This result is used to simplify equation (A25).

Before proceeding, it is useful to define

$$r_{i\cdots j\cdots k} \equiv M - m_i - \ldots - m_j - \ldots - m_k. \tag{B1}$$

To illustrate the argument, suppose that $n = 3$. Since $M = m_1 + m_2 + m_3$,

$$\frac{m_i}{r_{jk}} = 1 \quad \text{and} \quad \frac{(m_i + m_j)}{r_k} = 1, \tag{B2}$$

for all combinations of $1 \le i,j,k \le 3$ in which $i \ne j \ne k$. Explicitly, the sum over the $n! = 6$ permutations of (m_1, m_2, m_3) is

$$\frac{m_1}{M}\frac{m_2}{r_1}\frac{m_3}{r_{12}} + \frac{m_1}{M}\frac{m_3}{r_1}\frac{m_2}{r_{13}} + \frac{m_2}{M}\frac{m_1}{r_2}\frac{m_3}{r_{21}} + \frac{m_2}{M}\frac{m_3}{r_2}\frac{m_1}{r_{23}} + \frac{m_3}{M}\frac{m_1}{r_3}\frac{m_2}{r_{31}} + \frac{m_3}{M}\frac{m_2}{r_3}\frac{m_1}{r_{32}}.$$

The first identity of (B2) shows that this is the same as

$$\frac{m_1}{M}\frac{m_2}{r_1} + \frac{m_1}{M}\frac{m_3}{r_1} + \frac{m_2}{M}\frac{m_1}{r_2} + \frac{m_2}{M}\frac{m_3}{r_2} + \frac{m_3}{M}\frac{m_1}{r_3} + \frac{m_3}{M}\frac{m_2}{r_3}.$$

But this is just

$$\frac{m_1}{M}\frac{(m_2+m_3)}{r_1} + \frac{m_2}{M}\frac{(m_1+m_3)}{r_2} + \frac{m_3}{M}\frac{(m_1+m_2)}{r_3},$$

and use of the second identity of (B2) reduces this to

$$\frac{m_1}{M} + \frac{m_2}{M} + \frac{m_3}{M},$$

which is unity, by definition.

It is easy to see that, for all values of n, the requirement that $M = \sum_i^n m_i$ provides relations analogous to those in (B2), and that repeated use of these relations will reduce the sum of $\prod_{i=1}^n (m_i/R_{i-1})$ over all $n!$ permutations of the m_i to unity.