

Merger history trees of dark matter haloes in moving barrier models

Jorge Moreno,¹★ Carlo Giocoli²★ and Ravi K. Sheth¹★

¹*Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104-6396, USA*

²*Dipartimento di Astronomia, Università degli Studi di Padova, Vicolo dell'osservatorio 2 I-35122 Padova, Italy*

Accepted 2008 July 28. Received 2008 July 23; in original form 2008 May 19

ABSTRACT

We present an algorithm for generating merger histories of dark matter haloes. The algorithm is based on the excursion-set approach with moving barriers whose shape is motivated by the ellipsoidal collapse model of halo formation. In contrast to most other merger-tree algorithms, ours takes discrete steps in mass rather than time. This allows us to quantify effects which arise from the fact that outputs from numerical simulations are usually in discrete time bins. In addition, it suggests a natural set of scaling variables for describing the abundance of halo progenitors; this scaling is not as general as that associated with a spherical collapse. We test our algorithm by comparing its predictions with measurements in numerical simulations. The progenitor mass fractions and mass functions are in good agreement, as is the predicted scaling law. We also test the formation-redshift distribution, the mass distribution at formation, and the redshift distribution of the most recent major merger; all are in reasonable agreement with N -body simulation data, over a broad range of masses and redshifts. Finally, we study the effects of sampling in discrete time snapshots. In all cases, the improvement over algorithms based on the spherical collapse assumption is significant.

Key words: methods: numerical – galaxies: haloes – cosmology: theory – dark matter.

1 INTRODUCTION

Most models of galaxy formation in a hierarchical universe assume that the merger history of the surrounding dark matter halo plays an important role in determining the properties of a galaxy (e.g. White & Rees 1978; White & Frenk 1991; Baugh 2006 and references therein). Although halo-merger histories can be measured using N -body simulations, these can be time consuming and computationally intensive (Springel et al. 2005). This has fuelled considerable study of the formation and merger histories of dark matter haloes from a Monte Carlo perspective. Monte Carlo merger trees have the advantage of being fast and one may easily probe mass regimes inaccessible to current N -body simulations. Moreover, unlike N -body experiments, the cosmology and initial conditions may be easily modified.

The excursion-set framework (Bond et al. 1991; Lacey & Cole 1993), which is motivated by the pioneering work of Press & Schechter (1974), provides the basis for current models of halo assembly. Initially, this framework was based on the assumption that haloes form from a spherical collapse, of the type first described by Gunn & Gott (1972). Fast algorithms for generating halo merger trees, in which haloes were assumed to form from a spherical collapse, were developed in the 1990s (Kauffmann & White

1993; Sheth & Lemson 1999; Somerville & Kolatt 1999; Cole et al. 2000) (see Zhang, Fakhouri & Ma 2008b, for a review). However, spherical collapse overpredicts (underpredicts) the abundance of haloes in the low (high) mass regime. To address these issues, Sheth & Tormen (1999) extended the excursion-set framework to include ellipsoidal collapse (Sheth, Mo & Tormen 2001; Sheth & Tormen 2002). This clearly showed that merger-trees which assume spherical collapse are inadequate.

Hiotelis & Del Popolo (2006) describe a merger-tree algorithm which extends some of the older algorithms to incorporate aspects of the ellipsoidal collapse results. In addition, a number of new algorithms have recently been published (Neistein & Dekel 2008; Parkinson, Cole & Helly 2008); although efficient and accurate, such methods side-step the idea of ellipsoidal collapse altogether. Moreover, these methods are calibrated to match N -body simulations, and are therefore limited by the accuracy and scope of these simulations.

The aim of the present work is to provide a merger history tree algorithm, which is based explicitly on the excursion-set formalism with ellipsoidal collapse. The most significant difference between the algorithms we derive and all the others described above is that it takes discrete steps in mass rather than time. This feature allows us to study a few problems which are more difficult to address with the other methods.

After we completed this project, Zhang et al. (2008b) presented an alternative algorithm with ellipsoidal collapse (and discrete-time snapshots). This algorithm generates progenitors across many mass

★E-mail: jmoreno@physics.upenn.edu (JM); carlogiocoli@unipd.it (CG); shethrk@physics.upenn.edu (RKS)

bins and then assigns them to final haloes. In this sense, it is very similar to the Kauffmann & White (1993) merger tree, but solves many of its problems. Although this tree is quite different from ours, it also attempts to describe the merger history of a halo from the excursion-set formalism. These two approaches show that generating merger trees without tuning to N -body simulations is a quite non-trivial, yet interesting, challenge.

A review of background material and a description of our algorithm are given in Section 2. Section 3 compares our algorithm with excursion-set theory predictions, and with measurements in N -body simulations. The tests include the progenitor mass fractions and mass functions, the formation-redshift distribution, the mass distribution at formation and the redshift distribution of the most recent major merger. Section 4 summarizes our findings and discusses possible applications of our algorithm, an outline of which is in Appendix A.

2 MERGER TREES IN THE EXCURSION-SET APPROACH

In the excursion-set approach, the problem of estimating halo abundances is mapped to one of estimating the distribution of the number of steps a Brownian-motion random walk must take before it first crosses a barrier of specified height (Bond et al. 1991). In this approach, the height of the barrier plays a crucial role.

2.1 Constant and moving barriers

The Press–Schechter mass function is associated with barriers of constant height – such barriers arise naturally in models in which haloes form from a spherical collapse model. In contrast, in the ellipsoidal collapse model, barriers of the form

$$B(S, \delta_c) = \sqrt{q} \delta_c \left[1 + \beta \left(\frac{S}{q \delta_c^2} \right)^\gamma \right] \quad (1)$$

are more natural. Here, δ_c is the overdensity required for spherical collapse – it is a monotonically increasing function of redshift, and it is given by $\delta_{c0}/D(z)$, where $\delta_{c0} \equiv \delta_c(z=0) \sim 1.686$ and $D(z)$ is the linear-growth factor. S is a monotonically decreasing function of halo mass, given by

$$S(m) \equiv \sigma^2(m) = \frac{1}{2\pi^2} \int dk k^2 P(k) \tilde{W}^2(kR), \quad (2)$$

where $P(k)$ is the initial power spectrum of density fluctuations, linearly evolved to the present time, \tilde{W} is the Fourier transform of $W(x) = (3/x^3)(\sin x - x \cos x)$, $R = (3m/4\pi\bar{\rho})^{1/3}$, and $\bar{\rho}$ is the comoving background density. At large R , the overdensity contained in the associated volume is practically zero. As R [and $m(R)$] decreases, $S(R)$ increases, and δ_R executes a random walk. We refer the reader to Lacey & Cole (1993) for more details.

Consider a barrier $B[S, \delta_c(z)]$, as in equation (1), and an ensemble of random walks which start from the origin: $(S, \delta) = (0, 0)$. The excursion-set approach maps the distribution of S , the number of steps a random walk must take to first cross such a barrier, to the fraction of mass in m haloes at redshift z . This quantity is associated with the so-called unconditional mass function. The conditional mass function of high-redshift progenitors of a more massive final M halo at some lower redshift Z is modelled using walks which start from $[S_M, \delta_c(Z)]$ instead.

The shape of the mass functions (unconditional and conditional) is determined by the shape of the barrier, encoded in the (q, β, γ) parameters. The spherical collapse model is associated with $(q, \beta,$

$\gamma) = (1, 0, 0)$, whereas ellipsoidal collapse has $(0.707, 0.47, 0.615)$. When $\gamma > 1/2$, not all walks are guaranteed to cross the ellipsoidal collapse barrier (Sheth & Tormen 2002). Moreover, the barriers associated with two different times may intersect; of course, this never happens for the spherical collapse barriers. Sheth & Tormen suggest that this intersection of barriers may represent the possibility that haloes can fragment. This complicates the algorithm we describe below, so we instead study the limiting case of ‘square-root’ barriers for which $\gamma = 1/2$:

$$B(S, \delta_c) = \sqrt{q} \delta_c + \beta \sqrt{S}. \quad (3)$$

The predicted halo abundances associated with $(q, \beta, \gamma) = (0.55, 0.5, 0.5)$ are very similar to those in simulations (Mahmood & Rajesh 2005; Moreno, Giocoli & Sheth 2008). Moreover, the first crossing distribution of this barrier is known (Breiman 1967).

2.2 A conditional scaling symmetry

Recall that the unconditional mass function is associated to the first crossing distribution associated with walks which start from the origin: $(S, \delta) = (0, 0)$. When constant barriers are used, it can be expressed self-similarly as

$$f(m|z) dm = f(S|\delta_c) dS = f(\nu) d\nu, \quad (4)$$

where $\nu \equiv \delta_c^2/S$. The conditional mass function of m haloes at z that end up in bound objects of mass $M > m$ at $Z < z$ is given by $f(m, z|M, Z) dm = f(S_m, \delta_1|S_M, \delta_0) dS_m$, where $\delta_1 = \delta_c(z)$, $\delta_0 = \delta_c(Z)$. In other words, the conditional mass function is associated with the first crossing distribution of a barrier of height δ_1 by random walks with origin at (S_M, δ_0) . Because a straight line is straight whatever the origin of the coordinate system, the conditional mass function, in the spherical collapse model has the same functional form as that of the unconditional mass function, provided one sets $\nu = (\delta_1 - \delta_0)^2/(S_m - S_M)$.

However, for the square-root barrier, a walk which starts from $(\sqrt{q} \delta_0 + \beta \sqrt{S_M}, S_M)$ must cross a barrier of shape $\sqrt{q}(\delta_1 - \delta_0) + \beta \sqrt{S_m - S_M + S_M}$. This is not quite of the same form as equation (3). As a result, the conditional mass function is not simply a rescaled version of the unconditional one. Rather, in this model,

$$f(m, z|M, z_0) dm = f(S_m/S_M|\eta) d(S_m/S_M), \quad (5)$$

where

$$\eta \equiv \frac{\delta_1 - \delta_0}{\sqrt{S_M}} \quad (6)$$

(see Breiman 1967 or Giocoli et al. 2007 for the exact expressions). Thus, final haloes of different masses will have similar progenitor-mass functions when expressed in terms of S_m/S_M , provided they have similar values of η . While this scaling is like that for the constant barrier model, in the square-root barrier model, the progenitor mass function is *not* a function of the combination $\nu^2 = \eta^2/(S_m/S_M - 1)$. This is interesting, because Sheth & Tormen (2002) have shown that the conditional mass function in simulations is not well-fit by a function of ν . In what follows, we will present evidence that it is, however, a function of η and S_m/S_M separately, so the qualitatively different scaling associated with the square-root barrier is indeed seen in simulations.

2.3 Mass histories and merger trees

Fig. 1 illustrates how the mass-growth history of an object is encoded in the excursion-set approach if objects form from a spherical

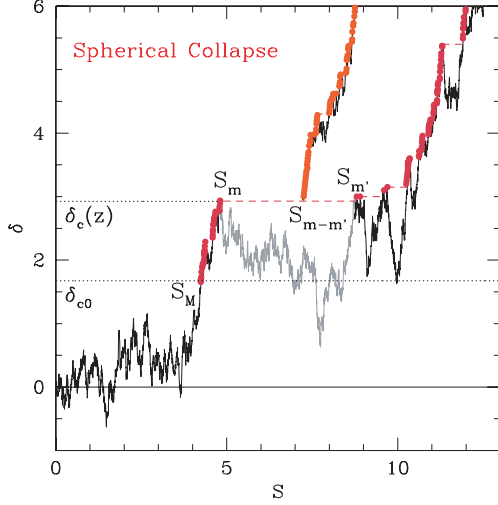


Figure 1. A random walk and its associated mass history. The dark filled circles represent the history of a halo of mass M at redshift $z = 0$. A merger $(m', m - m') \rightarrow m$ at redshift z is depicted by the $S_m \rightarrow S_{m'}$ jump at height $\delta_c(z)$. A new branch associated with $(m - m')$ is connected at $[S_{m-m'}, \delta_c(z)]$. The light-filled circles denote the mass history of this object.

collapse (also see fig. 1 of Lacey & Cole (1993)). The jagged line shows a random walk which starts from the origin: $(S, \delta) = (0, 0)$. Imagine drawing a horizontal line with height $\delta_{c0} = 1.686$ and marking the *smallest* value of S at which the walk intersects this ‘barrier’ of constant height (δ_{c0} corresponds to the present time and $\delta_c > \delta_{c0}$ corresponds to higher redshifts). The dotted-horizontal line denotes such barrier. Then increase the height of this barrier, and record how this value of S changes as δ increases. Such mass history points are depicted as dark filled circles in Fig. 1. The dashed lines show that S will occasionally jump from a small value to a larger one. Since S is a proxy for mass, and δ_c for time, such a jump is a proxy for an instantaneous change in mass: a merger. Note that the random walk steps under such jumps are not part of the mass history (e.g. the gray portion with $S_m < S < S_{m'}$).

The key to our merger tree algorithm, which is described in detail in Appendix A, is to recognize that these jumps mean that there are a set of other walks which one might associate with this one – one for each jump. One such walk is illustrated by the second jagged curve, which starts at about the middle of the panel. If the jump from S_m to $S_{m'}$ occurred when the barrier height was $\delta_c(z)$, then this other walk starts from $[S_{m-m'}, \delta_c(z)]$. The ‘merger history’ associated with this new *branch* is represented by the light-shade filled circles in Fig. 1. For every such jump, a new random walk must be drawn. For each jump within each of those new walks, the same process applies – more walks must be drawn. In summary, the bundle of such walks encodes the entire merger history of a present-day object. Notice that jumps can occur at any z – there is no constraint that they happen at discrete times. However, if one is interested in the mass function of progenitors at some fixed z , one simply reads-off the list of values of S at which this bundle of walks first cross $\delta_c(z)$.

So far, we have discussed how to generate trees in the spherical collapse model. Fig. 2 shows the same walk as before, but now the mass growth history associated with the walk is given by its intersection with square-root barriers of gradually increasing height. This shows clearly that the jumps in mass, and the times at which they occur, are modified. But the overall logic remains the same. Each jump

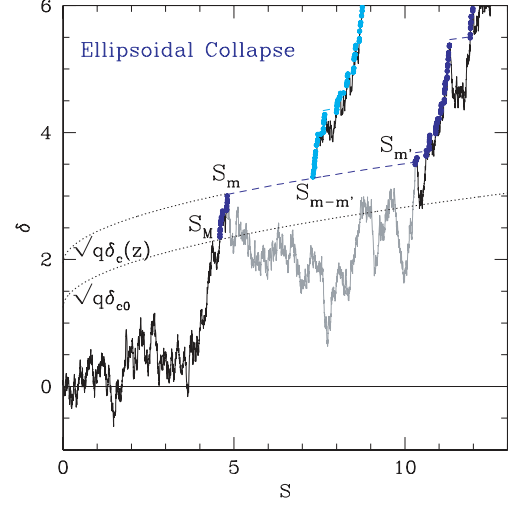


Figure 2. The same random walk as in Fig. 1, but now with square root rather than constant barriers, illustrating that the mass-accretion history depends on the barrier shape. In our algorithm, the new object with mass $(m - m')$ is now connected at $(S_{m-m'}, \sqrt{q}\delta_c(z) + \beta\sqrt{S_{m-m'}})$.

gives rise to a new walk that starts from $\{S_{m-m'}, B[S_{m-m'}, \delta_c(z)]\}$, where B is given by equation (3).

The natural generalization to spherical collapse is to incorporate the original $\gamma > 1/2$ barrier of Sheth et al. (2001). Such a choice would complicate this algorithm significantly. First of all, as Fig. 2 illustrates, the shape of the square-root barrier remains the same with different redshifts, except for an overall vertical shift. This is not the case with $\gamma > 1/2$. As δ_c increases (increasing redshift), the term $\delta_c^{1/2-2\gamma}$ makes the barrier in equation (1) increase less rapidly with S . A consequence of this is that the barriers associated with different redshifts cross. In the absence of crossing-barriers (e.g. constant and square-root barriers), one may uniquely map any point in the (S, δ) plane to (m, z) . The crossing of barriers invalidates this property, making the identification of jumps with mergers at a given-time ill defined.

3 COMPARISON WITH SIMULATIONS

In this section, we compare the statistical properties of our merger history trees with expectations from the excursion-set theory which they are supposed to reproduce, and with measurements in the GIF2 (German Israel Fund 2) N -body simulation (Gao et al. 2004b). The simulation followed the evolution of 400^3 particles in a periodic cubic box $110 h^{-1}$ Mpc on a side in a flat Λ cold dark matter (Λ CDM) background cosmology with parameters $(\Omega_m, \sigma_8, h, \Omega_b h^2, n) = (0.3, 0.9, 0.7, 0.0196, 1)$. Fifty simulation snapshots were output, equally spaced in $\log(1 + z)$. At each snapshot, haloes were identified using the spherical overdensity criterion, adopting for virial mass the definition of Eke, Cole & Frenk (1996) (i.e. with virial density at $\sim 324\bar{\rho}$ at redshift zero). The particle mass is $m_p = 1.73 \times 10^9 h^{-1} M_\odot$ and only objects with at least ten particles are considered. $M_*(z)$, defined by $\delta_c^2(z) = S[M_*(z)]$, is the typical mass scale at redshift z . It is common practice to express halo masses in terms of $M_* = M_*(z = 0)$ (the z -dependence is suppressed for the present time). For this cosmology and initial power spectrum, $M_* = 8.7 \times 10^{12} h^{-1} M_\odot \simeq 5030 m_p$. The simulation data and halo catalogues are available at <http://www.mpa-garching.mpg.de/Virgo>. See

Giocoli, Tormen & van den Bosch (2008a) for more details regarding the post-processing of the simulation.

To compare our merger histories with those in the GIF2 simulation, we generated 2000 realizations of our tree for each final halo mass bin M of interest. In all cases, the minimum mass considered was $m_{\text{dust}} = M/1000$, and the merger histories of haloes with mass below this were not followed (we call this minimum mass the ‘branching-mass resolution’). We used random walks with 10^5 steps in between S_M and S_{dust} to ensure that the mass change between the steps was less than m_{dust} . Having a small step size is essential to faithfully reproduce random walks in a computer. Moreover, if the step size is too large, we run the risk of missing branches. Numerical tests showed that the outputs converged to our results for small step sizes. See the Appendix, for more details on the implementation of our Monte Carlo tree. Recall that our tree does not take discrete steps in time. Nevertheless, for fair comparison with the measurements from the GIF2 simulation, the tree data were stored in the same discrete redshift bins as were output from the simulation. We use ‘Cont’ to denote the original tree data and ‘Snap’ for the data stored in redshift snapshots. Sections 3.3 and 3.4 study some merger-related quantities which are sensitive to the differences between these two ways of storing trees (Figs 8, 9 and 10).

3.1 The progenitor mass function

Figs (3)–(5) show the progenitor mass fractions and mass functions at five different redshifts ($z = 0.5, 1, 2, 3, 5$), for haloes identified at $z = 0$ with final masses given by $M/M_* = 0.06, 0.6$ and 6 . The

corresponding values of η (equation 6) are shown in each panel. In all three figures, filled circles show measurements in the GIF2 simulation, and open circles show results from our square-root trees. We probe the $m < m_p$ regime with our trees to verify consistency with analytic excursion-set predictions (the smooth curves in all the panels). The expressions we use are given explicitly in Appendix A of Giocoli et al. (2007). The short-dashed curve shows the constant barrier (β, q) = (0, 1) prediction associated with spherical collapse (Lacey & Cole 1993). The solid curves show the exact square-root barrier solution with $(\beta, q, \gamma) = (0.5, 0.55, 0.5)$; this is a complicated affair, involving sums of Parabolic Cylinder functions (Breiman 1967). Dashed curves show the considerably simpler approximation to the solution which is due to Sheth & Tormen (2002); this approximation is excellent over the entire range of interest. The long-dashed curves show this same approximation for the ellipsoidal collapse barrier: $(\beta, q, \gamma) = (0.707, 0.47, 0.615)$. The square-root barrier prediction agrees well with the $\gamma = 0.615$ curve, except in the high-mass regime. This discrepancy becomes evident when $\eta > 1$ and it is amplified with increasing η (see below).

Before we ask how our merger tree algorithm compares with simulations, we note that it produces progenitor mass functions that are well described by the theory curves over a wide range of masses and redshifts. At high redshifts, our tree data lie slightly below the theory curves at both high and low m/M , and slightly above in between, although where the crossover points occur depends on z and M . In all other regimes, our Monte Carlo trees match the square-root barrier predictions. Any additional disagreement with the GIF2

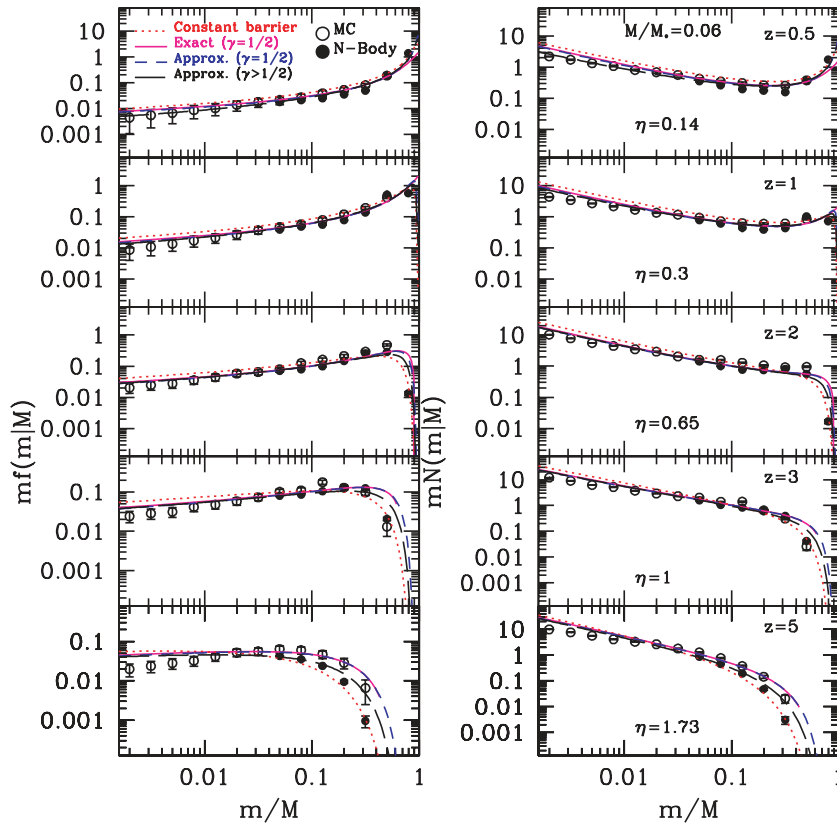


Figure 3. The progenitor mass fraction (left-hand side) and mass function (right-hand side) at redshifts $z = (0.5, 1, 2, 3, 5)$, for haloes of mass $M/M_* = 0.06$ at $z = 0$. Filled circles show measurements in the GIF2 simulation, and open circles are from our square-root trees. The smooth solid and dashed curves show the exact square-root barrier solution, and the series approximation, respectively. The long-dashed curve shows the ellipsoidal collapse model with $\gamma > 1/2$, and the short-dashed curve is the constant barrier prediction. Values of the scaling parameter η (equation 6) are also shown (see Fig. 6).

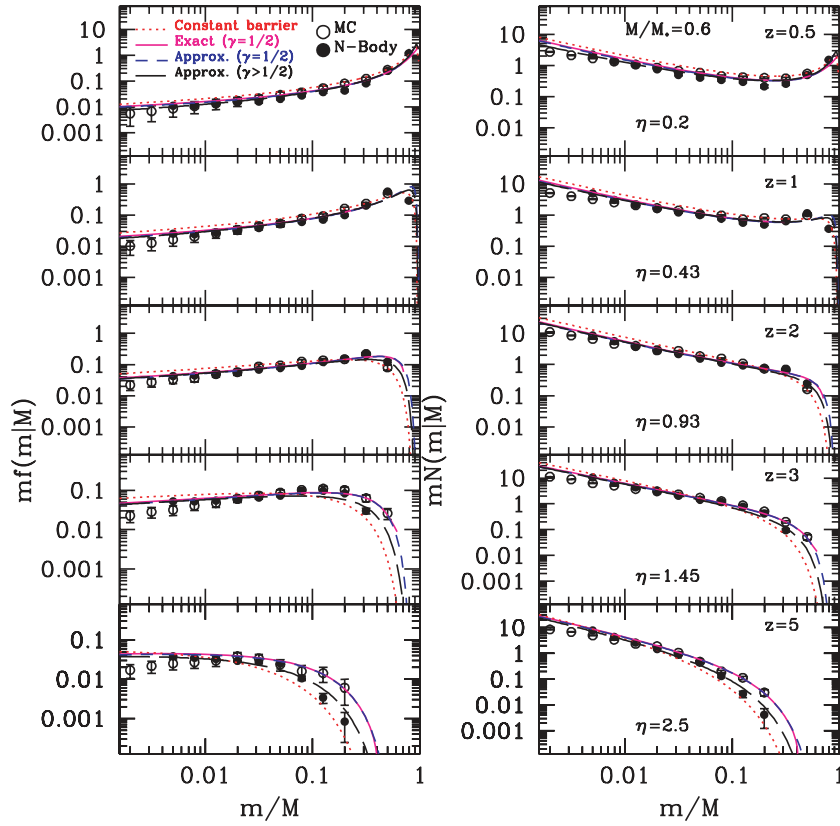


Figure 4. Same as Fig. 3, but with $M/M_* = 0.6$.

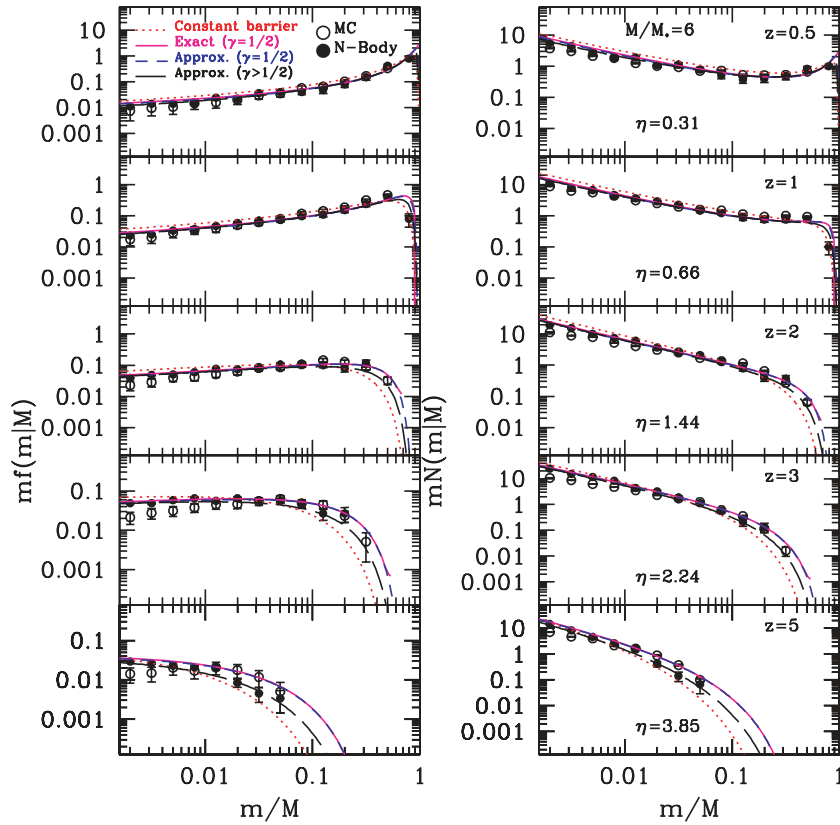


Figure 5. Same as Fig. 3, but with $M/M_* = 6$.

Table 1. The η -symmetry (equation 6) used for the comparisons shown in Fig. 6.

M/M_*	z	η	–	M/M_*	z	η
0.06	1	0.3	–	6	0.5	0.31
0.06	2	0.65	–	6	1	0.66
0.6	3	1.45	–	6	2	1.44

simulation measurements (compare open and filled circles) is due to limitations of the $\gamma = 1/2$ model.

Finally, recall that the square-root and constant barrier models make specific predictions for how the conditional mass functions should scale with final halo mass and time. Table 1 lists pairs with similar η , yet quite distinct values of M and z . Fig. 6 compares the associated conditional mass functions. The black squares and long-dashed lines refer to the left-hand side of Table 1 (low final masses), whereas the gray triangles and short-dashed lines refer to the right-hand side (high final masses). Notice that the curves are remarkably similar to one another, as are the symbols. This is true despite the fact that the values of η are not perfectly identical, and that $f(m, z|M, Z) dm = f(S_m/S_M|\eta) d(S_m/S_M) \simeq f(m/M|\eta) d(m/M)$. The results for low-mass haloes (black squares) are truncated at higher m/M than they are for larger M (gray triangles), simply because only haloes with at least ten particles are considered. Evidently, the conditional mass functions are indeed functions of η and S_m/S_M separately, rather than of the combination ν .

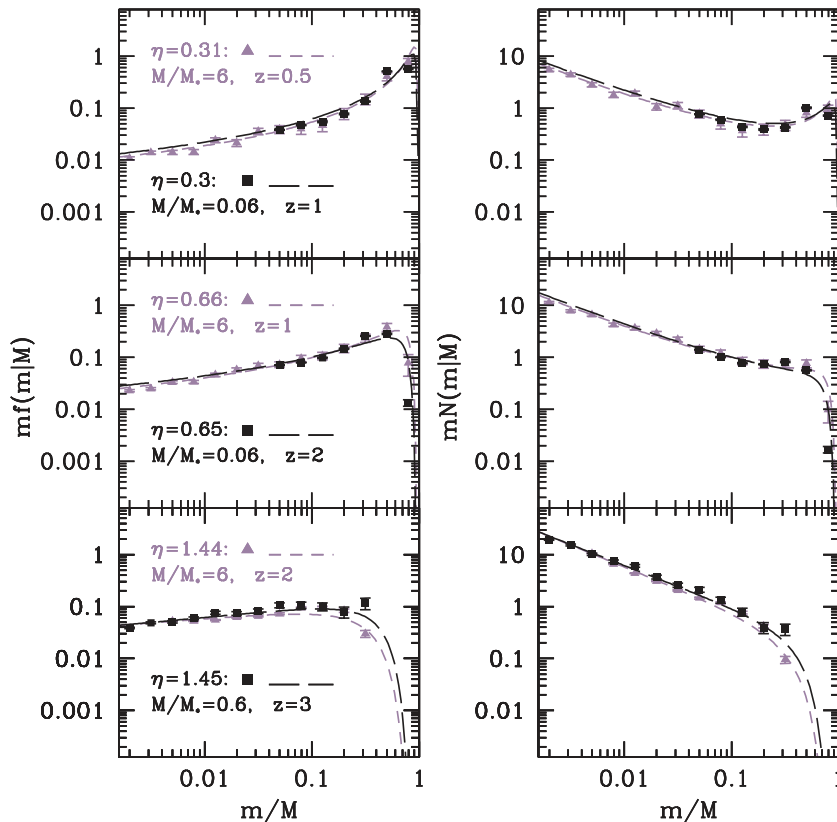


Figure 6. The η -symmetry. Different combinations of M and z with similar η (equation 6 and Table 1). N -body simulation measurements and the Sheth & Tormen (2002) result with $\gamma > 1/2$ are shown.

3.2 The distribution of formation redshifts

Following Lacey & Cole (1993), a halo is said to have ‘formed’ when it first acquires half of its final mass. For a given halo mass, there is a distribution of formation redshifts. This distribution is expected to peak at earlier times for lower mass haloes. The excursion-set model with constant barriers provides a simple expression for this distribution of formation times:

$$p(\omega_F) d\omega_F = 2\omega_F \operatorname{erfc}(\omega_F/\sqrt{2}) d\omega_F, \quad (7)$$

where

$$\omega_F \equiv \frac{\delta_c(z_F) - \delta_c(z_0)}{\sqrt{S(M/2) - S(M)}} = \frac{\eta}{\sqrt{S(M/2)/S(M) - 1}}, \quad (8)$$

with η given by equation (6).

The filled circles in Fig. 7 show the formation redshift distributions for haloes with masses in the range $0.9M \leq M \leq 1.1M$ with $\log_{10}(M/M_*) = \{1, \dots, -1.5\}$ in steps of -0.5 in the GIF2 simulation. We use the first snapshot when at least half the mass is in a single progenitor as the formation time, and make no attempt to interpolate our simulation formation redshifts between these discretely spaced output times (Harker et al. 2006; Giocoli et al. 2007). The open circles and triangles show the corresponding formation time distributions from our square-root and constant barrier trees. Recall that we do not discretise redshift in our tree, so the question of interpolation does not arise.

For the ellipsoidal collapse model, Giocoli et al. (2007) showed that the formation redshift is well described by equation (7) if one replaces $\omega_F \rightarrow \sqrt{q}\omega_F$. The smooth curves show this with $q = 1, 0.707$ and 0.55 (short dashed, long dashed and dashed), which represent

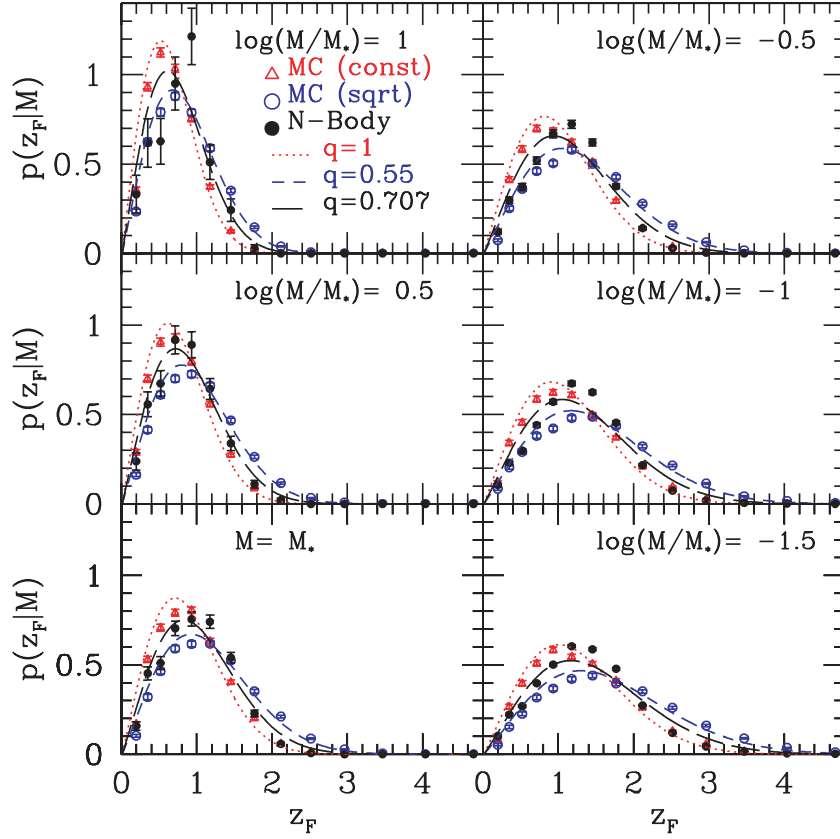


Figure 7. Distribution of formation redshifts. Filled circles show simulation data, open circles and triangles show results from the square root and constant barrier trees. Smooth curves show equation (7) with $q = 1, 0.707$ and 0.55 (short dashed, long dashed, and dashed), corresponding to the predicted distribution for constant barriers (spherical collapse), moving (ellipsoidal collapse) and square-root barriers.

the (constant, $\gamma = 0.615$, and square-root) barrier predictions. For higher values of q , the peaks are located at lower redshifts and the widths of the curves decrease. Strictly speaking, equation (7) only holds for white-noise initial conditions. Nevertheless, as pointed out Lacey & Cole (1993), it remains a reasonable approximation to CDM case. Furthermore, note that it provides an excellent description of the formation times generated by our trees. However, no choice of q provides particularly good agreement with the GIF2 simulation, a discrepancy noted by previous authors (Lin, Jing & Lin 2003; Hiortelis & Del Popolo 2006; Giocoli et al. 2007). This is likely a consequence of the excursion-set assumption that different steps in the walk are uncorrelated (Sheth & Tormen 2002). See Pan et al. (2008) and references therein for how one might improve on this.

3.3 The mass distribution at formation

The previous subsection studied halo formation, where formation was defined as the first time that the mass of one of the progenitors exceeds half the total. Therefore, this mass can have any value between $1/2$ and 1 times the final mass, and one can study the distribution of masses at, and just prior to, formation. The excursion-set constant barrier model makes a prediction for this distribution (Nusser & Sheth 1999). The mass distribution at formation is expected to be

$$p(\mu) d\mu = \frac{2}{\pi} \sqrt{\frac{1-\mu}{2\mu-1}} \frac{d\mu}{\mu^2}, \quad \text{where } 1/2 \leq \mu \leq 1, \quad (9)$$

and $\mu \equiv m/M$, and the distribution just before formation is

$$q(\mu) d\mu = \frac{1}{\pi(1-\mu)} \left(\sqrt{\mu/(1-2\mu)} - \sqrt{1-2\mu} \right) \frac{d\mu}{\mu^2}, \quad (10)$$

where $1/4 \leq \mu \leq 1/2$. We have found that, to a very good approximation, $\mu p(\mu) \rightarrow \mu q(\mu)$ if one replaces $\mu \rightarrow 1/4\mu$ (solid and dashed curves in Fig. 8, respectively). Let m_B and m_A denote the masses before and after formation, respectively. Roughly speaking, the symmetry about $M/2$ indicates that having a specific ratio of m_B to $M/2$ before formation is equally likely to having the same ratio of $M/2$ to m_A after formation.

Although these expressions were derived assuming a white-noise power spectrum, they are expected to be relatively independent of $P(k)$. Sheth & Tormen (2004) showed that they did indeed match numerical simulations well for different cosmologies and initial power spectra. Fig. 8 shows that they also work well for square-root barriers.

The agreement between the theory curves (smooth curves) and our Monte Carlo trees (jagged lines, labelled ‘Cont’) is excellent. All panels show that agreement with simulation data is also quite good. However, there is a systematic discrepancy: the cusp at $\mu = 1/2$ appears to be less pronounced in the simulation, with correspondingly lower tails. A similar discrepancy was seen by Sheth & Tormen (2004), who suggested that the fact that the simulations only provide discrete snapshots in time may be smoothing out the peak. By sampling our trees at the simulation snapshots (open symbols, labelled ‘Snap’), we have attempted to model the magnitude of this effect. Fig. 9 illustrates that the cusp has indeed been smoothed,

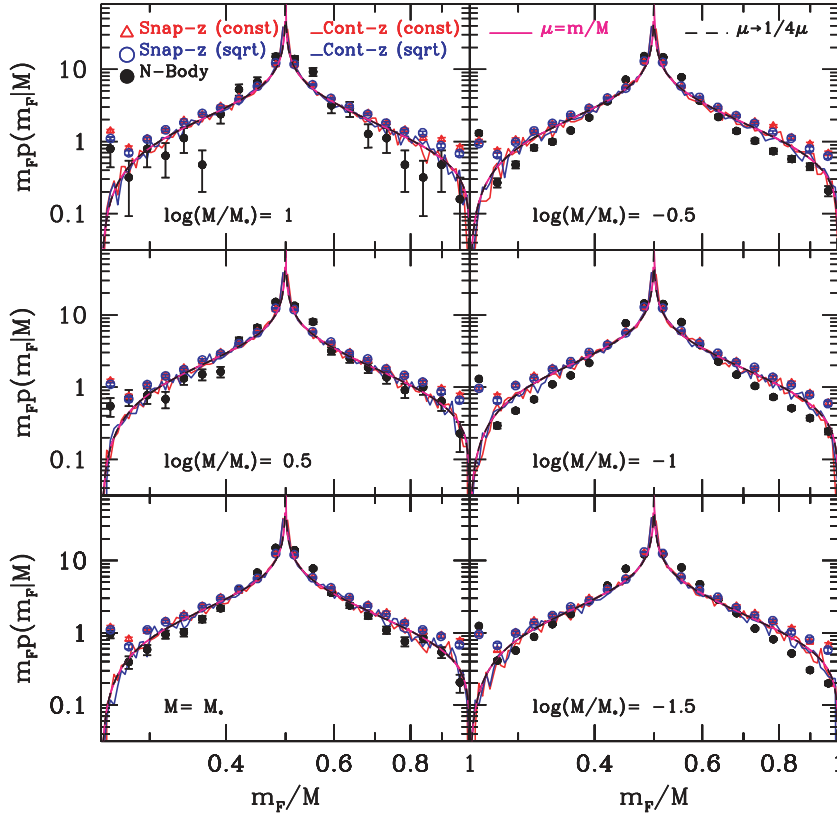


Figure 8. Distribution of the mass at formation for several final masses. The left half of each panel shows the mass just prior to formation, whereas the right half shows the mass just after formation. Filled circles show simulation data, open circles and triangles are from the square-root and constant barrier trees. The solid curve shows $\mu q(\mu)$ (right half) and $\mu p(\mu)$ (left half) (equations 10 and 9, respectively). The dashed curve shows these same expressions with $\mu \rightarrow 1/4\mu$.

but this is a dramatic effect only around $0.49 \leq \mu \leq 0.51$. The discrepancies further from the peak remain (Fig. 8).

3.4 The redshift distribution of the last major merger

Mergers of galaxies with similar masses are expected to produce strong short-lived periods of star formation: *starbursts* (Hernquist & Mihos 1995). Recent numerical studies suggest that the mass ratio of the galaxies involved plays an important role: merger-induced bursts occur when the galaxies have similar masses (Gao et al. 2004a; Springel & Hernquist 2005; Cox et al. 2008). Moreover, as suggested by Maller et al. (2006), a galaxy’s Hubble-type is strongly correlated with its last major merger.

Understanding such mergers requires understanding the mergers their host haloes undergo. Consider a merger $(m', m' - m) \rightarrow m$, with $m' > m - m'$. For ease of comparison with Parkinson et al. (2008), we will define a ‘major’ merger as one in which $(m - m')/m' \geq 1/3$. The filled circles in Fig. 10 show the redshift distribution of the last major merger on to the main branch. (The last major merger does not necessarily happen on the main branch. However, fig. 3 of Parkinson et al. (2008) suggests that, in most cases, it does. Presumably, this is because the assembly of haloes in recent times is dominated by mergers.) Curves show measurements in our full trees (‘Cont’), and open symbols show the result of only sampling the trees at the GIF2 simulation outputs (‘Snap’). For the discretely sampled data, only mergers involving haloes with at least ten particles are considered. Note that the anomalously low data point that is second from the left-hand side in all panels

appears to be an effect of seeing the tree at discrete snapshots – the smooth curves show no such dip. This feature is also present in the simulation analysed by Parkinson et al. (2008); we expect it to disappear if more finely spaced snapshots are analysed.

For high masses, the data from the square-root trees peak at about the same redshifts as the simulations; the constant barrier, spherical collapse trees peak at lower redshifts. This improvement relative to the spherical collapse case is similar to that in the modified GALFORM trees of Parkinson et al. (2008). However, our square-root trees tend to lie above the simulation at low redshifts, and below at higher redshifts. The discrepancy with simulation becomes increasingly worse at small masses, although it is possible that the GIF2 results for our two smallest mass bins are not reliable – the high redshift mergers involve haloes with few particles. Because we require haloes to have more than 10 particles, we are likely to miss major mergers once the typical mass becomes of this order.

4 DISCUSSION

We presented an algorithm for generating merger histories of dark matter haloes. This algorithm is based on the excursion-set approach (Figs 1 and 2, and related discussion), and can handle moving barriers of the sort that are associated with ellipsoidal collapse (equation 1). We illustrated its use by generating the forest of trees associated with a square-root barrier (equation 3). The halo mass function associated with this barrier is known to provide a reasonable description of halo abundances in the GIF2 simulation against

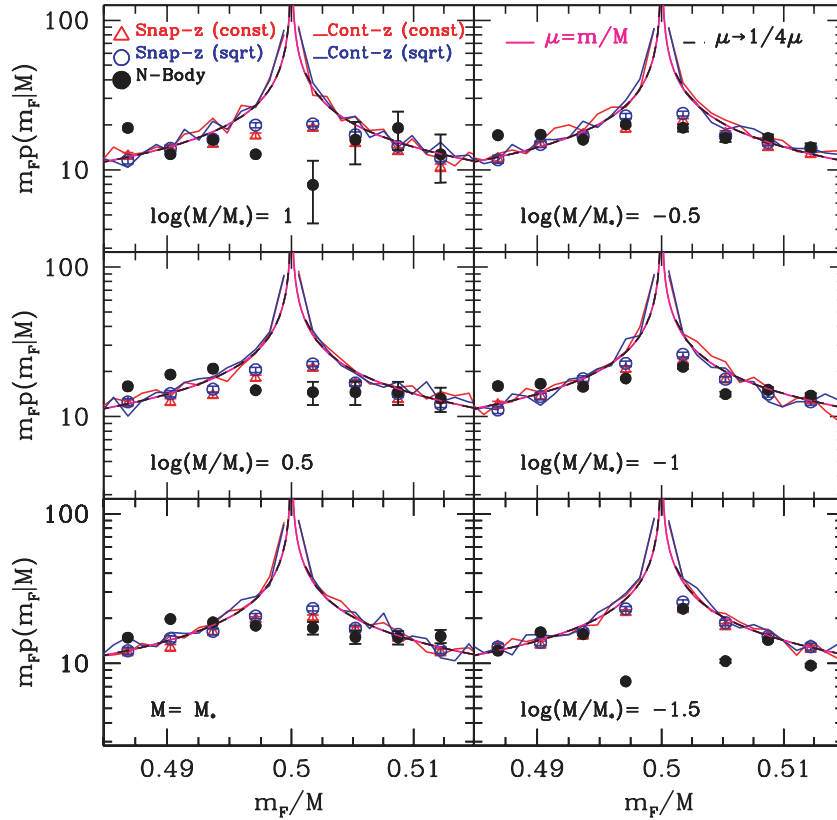


Figure 9. Same as Fig. 8, but showing only the region around $m/M = 1/2$. The peak in the simulations (filled symbols) is less pronounced than in the merger trees (jagged lines). Open circles show the result of sampling the merger trees at the same redshifts as the simulation snapshots: this makes a dramatic difference around $0.49 \leq \mu \leq 0.51$, suggesting that the sharp cusp predicted by the theory will also be present in simulations with sufficiently closely spaced outputs. The smaller discrepancies further from the peak remain.

which we test our model (Mahmood & Rajesh 2005; Moreno et al. 2008).

Our algorithm produces merger histories which yield the progenitor mass functions which are commonly computed using excursion-set theory – demonstrating that our approach is nearly self-consistent – over a broad range of masses and redshifts (Figs 3–5). These progenitor mass functions are also in reasonable agreement with measurements in simulations; they are a significant improvement on trees based on the constant barrier, spherical collapse model.

Our algorithm is different from others in the literature because it is based on taking discrete steps in mass, whereas all others (full N -body simulations included) take discrete steps in time. We also used our algorithm to show that while the distribution of times at which haloes assemble half their mass depends quite strongly on the barrier shape (Fig. 7), the distribution of masses at formation does not (Fig. 8). Excursion-set related formulae for this ‘universal’ distribution (equations 9 and 10) provide an excellent description of our merger trees; the agreement with simulations is good, but not perfect.

The algorithm is described in detail in Appendix A. In essence, it requires that one be able to generate one-dimensional random walks quickly. Since this reduces to being able to generate long strings of Gaussian variables, and fast routines for this exist, it is reasonably fast. Significant speed-ups can be obtained if one exploits known properties of random walks. For example, in the present context, all steps in the walk which lie below the current threshold value

are not interesting (e.g. gray-jagged portions of the random walks during the $S_m \rightarrow S_{m'}$ jump in Figs 1 and 2). If the distribution of the number of steps it takes to first exceed the current value is known, then one need not generate all these steps, one can instead draw a number from this distribution, and simply jump this number of steps. Incorporating such changes into our algorithm is the subject of ongoing work.

The excursion-set approach successfully describes many properties of the hierarchical growth of dark matter haloes. However, with the exception of white-noise initial conditions, the progenitor distributions it predicts cannot be made consistent with the intuitively attractive notion that, in sufficiently small time-steps, mergers are binary (Sheth & Pitman 1997; Benson, Kamionkowski & Hassani 2005; Benson 2008; Neistein & Dekel 2008; Zhang, Ma & Fakhouri 2008a). This accounts for some of the discrepancy between our trees and the excursion-set based theory curves. A number of authors have attempted to alleviate this by relaxing the assumption of binary mergers in their merger tree algorithms. For example, one of the two algorithms in Sheth & Lemson (1999) reproduces the spherical collapse based results exactly, for arbitrary power spectra. In this algorithm, mergers occur between groups of objects rather than just two but, as they pointed out, this was a rather contrived solution. Somerville & Kolatt (1999), Neistein & Dekel (2008) and Zhang et al. (2008b) discuss other possibilities. We make no attempt to address this issue with our algorithm.

The appendix also shows that building the main branch is straightforward (Fig. A2). We expect our approach to facilitate studies

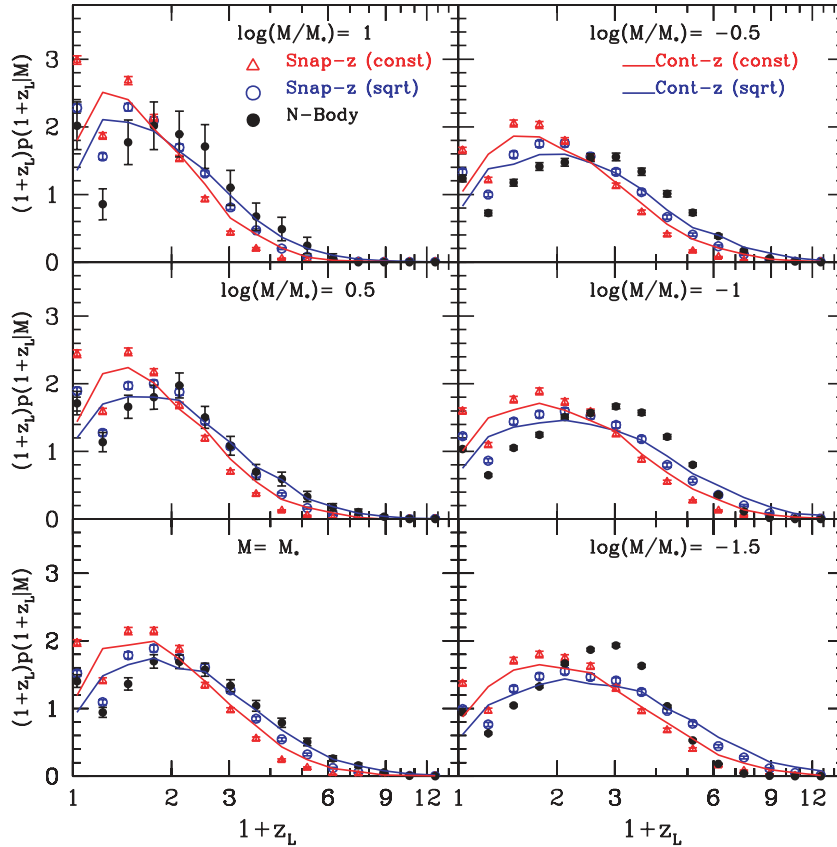


Figure 10. The redshift distribution of the most recent major merger, for several final masses. A major merger is defined as one in which the minor component has at least 1/3 of the mass of the major component. Filled circles show the GIF2 simulation data, open symbols show the corresponding measurements in the same ‘snapshot’ versions of our trees, and smooth curves show the ‘continuous’ distributions which would be seen with arbitrarily closely spaced output times.

involving the mass-accretion history of a halo (Avila-Reese, Firmani & Hernández 1998; Nusser & Sheth 1999; van den Bosch 2002; Wechsler et al. 2002; Gao et al. 2005; Maulbetsch et al. 2007; Stewart et al. 2008; Giocoli et al. 2008a; Giocoli, Pieri & Tormen 2008b). This is the subject of work in progress.

We mentioned in the introduction that halo merger histories play an important role in galaxy formation models. Models of the mergers of supermassive black holes (Menou, Haiman & Narayanan 2001; Volonteri, Haardt & Madau 2003; Yoo et al. 2007; Volonteri, Haardt & Gültekin 2008), merger-induced starbursts (Hernquist & Mihos 1995) and quasars (Kauffmann & Haehnelt 2000) also have the assembly of haloes as their backbone. Halo assembly histories also play a key role in studies of the brightest cluster galaxies (De Lucia & Blaizot 2007), luminous red galaxies (Almeida et al. 2008; Conroy, Ho & White 2007; Masjedi, Hogg & Blanton 2008; Wake et al. 2008), satellites and the intracluster light (Conroy, Wechsler & Kravtsov 2007; Skibba, Sheth & Martino 2007), and the nature of substructure in galaxy clusters. This last is important for interpreting the Sunyaev-Zel’dovich (Holder, McCarthy & Babul 2007), and strong lensing (Natarajan, De Lucia & Springel 2007) signals.

One of the advantages of using Monte Carlo merger trees is that one may easily change the underlying cosmology and initial power spectrum. Recently, there have been attempts to describe spherical collapse and non-linear growth in modified-gravity theories (Laszlo & Bean 2008; Schäfer & Koyama 2008). In principle, such modifications can easily be incorporated into our algorithm. We hope that our algorithm will prove useful in some of these studies.

ACKNOWLEDGMENTS

We wish to thank E. Neistein, K. Stewart, V. Avila-Reese, J. Pan, M. Martino, J. Newman and J. Hyde for useful discussions and positive feedback. We also thank our referee, O. Fakhouri, for his insightful criticisms and genuine interest in improving this work. J. Moreno would like to thank V. Moreno and the kids for their patience and understanding during the completion of this project. This work was supported in part by CONACyT–Mexico, and by Grant 2002352 from the US–Israel BSF.

REFERENCES

- Almeida C., Baugh C. M., Wake D. A., Lacey C. G., Benson A. J., Bower R. G., Pimbblet K., 2008, *MNRAS*, 386, 2145
- Avila-Reese V., Firmani C., Hernández X., 1998, *ApJ*, 505, 37
- Baugh C. M., 2006, *Rep. Prog. Phys.*, 69, 3101
- Benson A. J., Kamionkowski M., Hassani S. H., 2005, *MNRAS*, 357, 847
- Benson A. J., 2008, *MNRAS*, submitted (arXiv:0802.3713)
- Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, *ApJ*, 379, 440
- Breiman L., 1967, in Cam. L. L., ed. *Proc. 5th Berk. Symp. Math. Statist. Prob. Vol. 2*, Univ. California Press, Berkeley, p. 9
- Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168
- Conroy C., Ho S., White M., 2007, *MNRAS*, 379, 1491
- Conroy C., Wechsler R. H., Kravtsov A. V., 2007, *ApJ*, 668, 826
- Cox T. J., Jonsson P., Somerville R. S., Primack J. R., Dekel A., 2008, *MNRAS*, 384, 386
- De Lucia G., Blaizot J., 2007, *MNRAS*, 375, 2
- Eke V. R., Cole S., Frenk C. S., 1996, *MNRAS*, 282, 263

- Gao L., Loeb A., Peebles P. J. E., White S. D. M., Jenkins A., 2004a, *ApJ*, 614, 17
- Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004b, *MNRAS*, 355, 819
- Gao L., White S. D. M., Jenkins A., Frenk C. S., Springel V., 2005, *MNRAS*, 363, 379
- Giocoli C., Moreno J., Sheth R. K., Tormen G., 2007, *MNRAS*, 376, 977
- Giocoli C., Tormen G., van den Bosch F. C., 2008a, *MNRAS*, 386, 2135
- Giocoli C., Pieri L., Tormen G., 2008b, *MNRAS*, 387, 689
- Gunn J. E., Gott J. R. I., 1972, *ApJ*, 176, 1
- Harker G., Cole S., Helly J., Frenk C., Jenkins A., 2006, *MNRAS*, 367, 1039
- Hernquist L., Mihos J. C., 1995, *ApJ*, 448, 41
- Hiotelis N., Del Popolo A., 2006, *AP&SS*, 301, 167
- Holder G. P., McCarthy I. G., Babul A., 2007, *MNRAS*, 382, 1697
- Kauffmann G., White S. D. M., 1993, *MNRAS*, 261, 921
- Kauffmann G., Haehnelt M., 2000, *MNRAS*, 311, 576
- Lacey C., Cole S., 1993, *MNRAS*, 262, 627
- Laszlo I., Bean R., 2008, *Phys. Rev. D*, 77, 024048
- Lin W. P., Jing Y. P., Lin L., 2003, *MNRAS*, 344, 1327
- Mahmood A., Rajesh R., 2005, preprint (astro-ph/0502513)
- Maller A. H., Katz N., Kereš D., Davé R., Weinberg D. H., 2006, *ApJ*, 647, 763
- Masjedi M., Hogg D. W., Blanton M. R., 2008, *ApJ*, 679, 260
- Maulbetsch C., Avila-Reese V., Colín P., Gottlöber S., Khalatyan A., Steinmetz M., 2007, *ApJ*, 654, 53
- Menou K., Haiman Z., Narayanan V. K., 2001, *ApJ*, 558, 535
- Moreno J., Giocoli C., Sheth R. K., 2007, *MNRAS*, submitted (arXiv:0712.4100)
- Natarajan P., De Lucia G., Springel V., 2007, *MNRAS*, 376, 180
- Neistein E., Dekel A., 2008, *MNRAS*, 383, 615
- Neistein E., Dekel A., 2008, *MNRAS*, 388, 1792
- Nusser A., Sheth R. K., 1999, *MNRAS*, 303, 685
- Pan J., Wang Y., Chen X., Teodoro L., 2008, *MNRAS*, 389, 461
- Parkinson H., Cole S., Helly J., 2008, *MNRAS*, 383, 557
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Schäfer B. M., Koyama K., 2008, *MNRAS*, 385, 411
- Sheth R. K., Lemson G., 1999, *MNRAS*, 305, 946
- Sheth R. K., Pitman J., 1997, *MNRAS*, 289, 66
- Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119
- Sheth R. K., Tormen G., 2002, *MNRAS*, 329, 61
- Sheth R. K., Tormen G., 2004, *MNRAS*, 349, 1464
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Skibba R. A., Sheth R. K., Martino M. C., 2007, *MNRAS*, 382, 1940
- Somerville R. S., Kolatt T. S., 1999, *MNRAS*, 305, 1
- Springel V., Hernquist L., 2005, *ApJ*, 622, L9
- Springel V. et al., 2005, *Nat*, 435, 629
- Stewart K. R., Bullock J. S., Wechsler R. H., Maller A. H., Zentner A. R., 2008, *ApJ*, 683, 597
- van den Bosch F. C., 2002, *MNRAS*, 331, 98
- Volonteri M., Haardt F., Madau P., 2003, *ApJ*, 582, 559
- Volonteri M., Haardt F., Gültekin K., 2008, *MNRAS*, 384, 1387
- Wake D. A., Sheth R. K., Nichol R. C., Lam T. Y. et al., 2008, *MNRAS*, 387, 1045
- Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, *ApJ*, 568, 52
- White S. D. M., Rees M. J., 1978, *MNRAS*, 183, 341
- White S. D. M., Frenk C. S., 1991, *ApJ*, 379, 52
- Yoo J., Miralda-Escudé J., Weinberg D. H., Zheng Z., Morgan C. W., 2007, *ApJ*, 667, 813
- Zhang J., Ma C.-P., Fakhouri O., 2008a, *MNRAS*, 387, L13
- Zhang J., Fakhouri O., Ma C.-P., 2008b, *MNRAS*, 389, 1521

APPENDIX A: THE ALGORITHM

Consider a dark matter halo of mass M_0 at redshift Z_0 (Fig. A1). Realisations of its merger history may be constructed with our tree

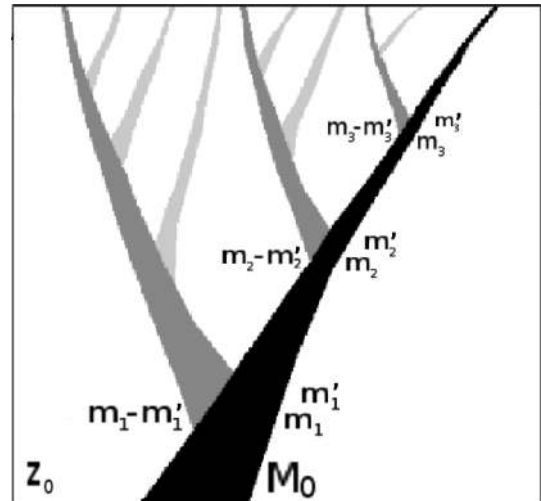


Figure A1. A sample merger history tree of a halo with mass M_0 at Z_0 . The mass history (black branch) is constructed with a random walk. Mass jumps larger than m_{dust} are identified, and branches are connected there (medium-gray). This process is repeated for each branch, and new branches are connected (light-gray) until the tree is complete.

algorithm. Below we explain how each branch is built and how these branches are connected.

A1 The branch: from random walks to mass histories

The mass growth history of a halo is contained in a random walk (Section 2.3). First we discuss the constant-barrier (spherical collapse) model and then the square-root barrier (ellipsoidal collapse) case.

A random walk is essentially a collection of steps and heights: $\{s_0, \dots, s_i, \dots\}$ and $\{h_0, \dots, h_i, \dots\}$ (e.g., the jagged line in Fig. 1). Consider a horizontal barrier of height $\delta_c(z)$. This line may intersect the walk at several values of S . The smallest of such of values, S_m , indicates what mass the halo had at redshift z . As z increases, so does δ_c – and m decreases accordingly, as expected in any hierarchical model of halo assembly.

As we increase the height of the barrier, a subset $\{(S_0, H_0), \dots, (S_j, H_j), \dots\}$ of the walk is chosen. These points contain the mass history (e.g. the dark-filled circles in Fig. 1). Every point in this subset has the following property: they are higher than all their predecessors. This is illustrated by the point at $[S_m, \delta_c(z)]$ Fig. 1: it has the maximum height in the range $S_m \leq S \leq S_m$. In other words, for any point (s_i, h_i) on the walk to be selected as part of the history, it must satisfy the following condition:

$$h_i > h_k, \quad \forall k < i. \quad (\text{A1})$$

Now we discuss what modifications are necessary when the square-root barrier model of ellipsoidal collapse is used (Fig. 2). Consider a barrier (equation 3) with δ -intercept $\sqrt{q}\delta_c(z)$. This curve may intersect the random walk at several places. We pick the smallest of these to find the mass at that redshift. As z increases, the δ -intercept increases, but the overall shape of the barrier remains unchanged. This is how the mass history $\{(S_0, H_0), \dots, (S_j, H_j), \dots\}$ is selected. From this subset we can construct a string $\{\sqrt{q}D_0, \dots, \sqrt{q}D_i, \dots\}$

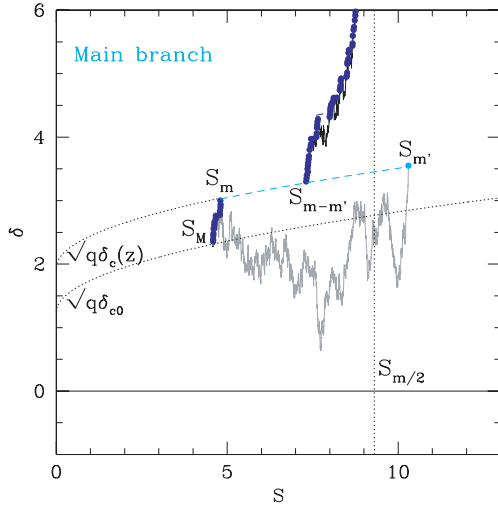


Figure A2. Main branch algorithm. The symbols and styles are the same as in Fig. 2. The dotted vertical line denotes $S = S_{m/2}$. The main branch is built by following the most massive piece at each split: $S_{m'} > S_{m/2}$ indicates that $m' < m/2$, implying that $m - m' > m'$, whose mass history we must follow.

of ‘ δ -intercepts’, where

$$D_i = \frac{1}{\sqrt{q}}(H_i - \beta\sqrt{s_i}).^1 \quad (\text{A2})$$

¹ Every point on the history has the following property: its δ -intercept is higher than that of its predecessors. In other words, for a point (s_i, h_i) on the walk to belong to the mass history, condition (A1) is replaced by

$$d_i > d_k, \quad \forall k < i, \quad (\text{A3})$$

where

$$d_i = \frac{1}{\sqrt{q}}(h_i - \beta\sqrt{s_i}). \quad (\text{A4})$$

For a given cosmology and initial power spectrum one may map excursion-set variables into physical ones: $[S(m), \delta_c(z)] \leftrightarrow (m, z)$. With this prescription, the mass accretion history is obtained: $\{(M_0, Z_0), \dots, (M_j, Z_j) \dots\}$. The masses in the history are such that $S(M_j) = S_j$ and the redshifts are given by $\delta_c(Z_j) = D_j$. When barriers are constant, $D_j = H_j$.

A2 The tree: connecting the branches

Consider a halo of mass M_0 at redshift Z_0 . We are interested in constructing its merger history tree. The first step is to draw a

¹ Note that this mapping is ill-defined when $\gamma > 1/2$, because these barriers cross.

random walk with origin at

$$s_0 = S(M_0), \quad h_0 = B[S(M_0), \delta_c(Z_0)], \quad (\text{A5})$$

where $B(S, \delta_c)$ is given by equation (3). The mass history is then collected: $\{(M_0, Z_0), \dots, (M_j, Z_j), \dots\}$ (see section A1). In Fig. A1, this is represented by the black branch.

The mass history can also be seen as a series of jumps in mass: $\{(M_0 \leftarrow M'_0, Z_0), \dots, (M_j \leftarrow M'_j, Z_j), \dots\}$, with $M'_j = M_{j+1}$. Such jumps are interpreted as binary mergers: $M'_j + (M_j - M'_j) \rightarrow M_j$. In practice, we only care about the mass jumps above the branching-mass resolution. Denote this subset of jumps as $\{\dots, (m_j \rightarrow m'_j, z_j), \dots\}$, where

$$m_j - m'_j > m_{\text{dust}}. \quad (\text{A6})$$

These jumps are shown in Fig. A1. The next step is to construct the history of each halo with mass $(m_j - m'_j)$ that falls on to the mass history of the M_0 -halo. This is done by generating random walks originating from

$$s_0 = S(m_j - m'_j), \quad h_0 = B[S(m_j - m'_j), \delta_c(z_j)], \quad (\text{A7})$$

Such mass histories correspond to the medium-gray branches in Fig. A1. The above process is repeated for each of these branches: retrieve their mass history, identify large enough mass jumps, and attach new branches there (light-gray in Fig. A1). Eventually, all the new branches will only involve mass jumps that do not satisfy condition (A6). At this point the tree is complete and the algorithm stops.

A3 The main branch: a simple modification

In a merger history tree, the ‘main’ branch of a halo is obtained by following the most massive progenitor at each mass split. In this section, we show that the mass history algorithm described in section A1 can be easily modified to construct the main branch. This process requires continuous monitoring of the walk and its associated history, and is illustrated in Fig. A2 (compare to Fig. 2).

Recall that mass decreases with increasing S . For the portions in the mass history consisting of jumps where the mass loss is less than half, the algorithm is unchanged (e.g. the portion with $S_M \leq S \leq S_m$ in Fig. A2). Occasionally, there are jumps where more than half the mass is lost. Such is the case for the $S_m \rightarrow S_{m'}$ jump illustrated in Fig. A2, with $S_{m'} > S_{m/2}$ [i.e. $m' < m/2$ and $(m - m') > m/2$]. To construct the main branch in that situation, one must simply follow $(m - m')$, not m' . In other words, instead of continuing the walk at $(S_{m'}, \sqrt{q}\delta_c(z) + \beta\sqrt{S_{m'}})$, one must continue from $(S_{m-m'}, \sqrt{q}\delta_c(z) + \beta\sqrt{S_{m-m'}})$ (i.e., the dark filled circles in Fig. A2).

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.