How to plant a merger tree

Rachel S. Somerville[1,2]⋆ and Tsafrir S. Kolatt[2]

[1] *Racah Institute of Physics, The Hebrew University, Jerusalem*
[2] *Physics Department, University of California, Santa Cruz, USA*

ABSTRACT

We investigate several approaches for constructing Monte Carlo realizations of the merging history of virialized dark matter haloes ('merger trees') using the extended Press–Schechter formalism. We describe several unsuccessful methods in order to illustrate some of the difficult aspects of this problem. We develop a practical method that leads to the reconstruction of the mean quantities that can be derived from the Press–Schechter model. This method is convenient, computationally efficient, and works for any power spectrum or background cosmology. In addition, we investigate statistics that describe the distribution of the number of progenitors and their masses as a function of redshift.

Key words: galaxies: clusters: general – galaxies: formation – cosmology: theory – dark matter.

1 INTRODUCTION

In the standard picture of modern structure formation, small-amplitude Gaussian density fluctuations, which perhaps arose from quantum fluctuations and were amplified by a period of rapid inflation, become more overdense with respect to their surroundings as the Universe expands. Eventually the self-gravity acting on these regions becomes larger than the pressure of the expansion, and they collapse to form bound, virialized structures. In hierarchical models, such as the cold dark matter (CDM) family of models, the amplitude of the fluctuations decreases with increasing scale. Thus small-mass objects form first, and are then incorporated into larger structures as time progresses. These dense, gravitationally bound structures provide the environments where galaxies can form. Hierarchical structure formation thus gives a natural explanation for the very complex observed large-scale structure of the Universe, i.e., clusters, superclusters, filaments, etc.

One way to study this process is with *N*-body simulations. However, numerical simulations have familiar drawbacks. They are computationally expensive, so it is difficult or impossible to explore a wide range of models or different realizations of the same model. In addition, memory and time limitations make it impossible to attain the mass and force resolution required to study simultaneously objects from dwarf galaxies ($\sim 10^9 M_\odot$) to clusters ($\sim 10^{15} M_\odot$). Semi-analytic methods are therefore an important alternative.

The model developed by Press & Schechter (1974) provides a simple but relatively effective framework for the description of the mass history of particles in a hierarchical universe with Gaussian random-phase initial perturbations. The focus of the original Press–Schechter model was the derivation of the multiplicity function of

non-linear objects ('haloes') as a function of redshift, i.e., the 'mass function' or number density of haloes of a given mass at a redshift *z*. This prediction has been tested quite extensively and found to be in relatively good agreement with *N*-body simulations (Efstathiou et al. 1988; Gelb & Bertschinger 1994; Lacey & Cole 1994; Ma 1996; Gross et al. 1998). The Press–Schechter theory was extended to give the *conditional* probability that a particle in a halo of mass M_0 at z_0 was in a halo of mass M_1 at an earlier redshift z_1, leading to an expression for the conditional mass function (Bond et al. 1991; Bower 1991). The extended Press–Schechter formalism can also be manipulated to obtain expressions for halo survival times, formation times, and merger rates (Lacey & Cole 1993, hereafter LC93), which have also been shown to agree reasonably well with the results from *N*-body simulations (Lacey & Cole 1994).

The computation of these mean quantities within the Press–Schechter model is straightforward. However, for certain purposes one would like to go beyond this. In particular, the semi-analytic approach to modelling galaxy formation (cf. Kauffmann, White & Guiderdoni 1993, Cole et al. 1994) attempts to describe the formation history of galaxies and gas within dark matter haloes, including simplified hydrodynamics, star formation, supernova feedback, galaxy–galaxy merging, and stellar population synthesis. These models rely on the construction of a 'merger tree', which involves predicting the masses of progenitor haloes and the redshifts at which they merge to form larger haloes. Galaxies initially form in their own halo and are traced as they are incorporated into larger haloes, and eventually perhaps merge with other galaxies. A halo of a given mass may have a variety of merging histories, and the properties of galaxies that form within this halo presumably depend to some extent on the details of this history.

Most of the previous work using semi-analytic models has focused on reproducing or predicting mean quantities and qualitative trends. However, as observational data continue to improve,

⋆ E-mail: rachels@astro.huji.ac.il

one would like to be able to investigate whether the broader properties of the predicted *distribution* of model galaxies are consistent with the observations. For example, there has already been some investigation of whether the scatter in the observed Tully–Fisher relation (Eisenstein & Loeb 1996), and in the colour–magnitude and line strength versus velocity dispersion (Mg–σ) relations (Kauffman 1996), in merger models is consistent with observations. Before we can trust the models for evaluating these kinds of questions, we must ensure that the merger trees not only satisfy the mean properties readily predicted by the extended Press–Schechter model, but ideally also the full distribution function. As we shall see, accomplishing this goal is far from straightforward, and it has not been thoroughly investigated in previous work on this subject.

In this paper we discuss some of the practical and theoretical difficulties of using the extended Press–Schechter model to create Monte Carlo merger histories of dark matter haloes. We mention some limitations of the previously proposed methods for the construction of merger trees, and the motivation for developing a new approach. We discuss several unsuccessful approaches, in an attempt to clarify some aspects of this problem as well as to prevent others from following the same dead ends. In addition, the formalism we present may be useful in future work on this subject, as the final approach that we present seems to be effective and convenient, but it is still not rigorous. A primary motivation for embarking on this project was to develop a method that reproduces the full joint probability distribution, rather than just the mean. Our method is not guaranteed to do so, and further investigation of this question requires a comparison with *N*-body simulations. This will be presented in a companion paper (Somerville et al. in preparation).

In Section 2 we give a brief introduction to the Press–Schechter formalism. In Section 3 we summarize some of the previous methods for creating merger trees. In Section 4 we describe the source of some difficulties one encounters in attempting to use the extended Press–Schechter formalism to construct merger trees. We develop a simple model of the joint probability distribution in Section 5 and attempt to use it to construct merger trees. In Section 6 we present a completely different approach which eventually leads to our successful method, described in Section 6.3. In Section 7 we investigate the distribution of progenitor number and mass given by our successful method. We summarize and conclude in Section 8.

Readers who are interested only in the successful method may skip to Section 6.

2 THE PRESS–SCHECHTER FORMALISM

The Press–Schechter model (Press & Schechter 1974) is based on a combination of linear growth theory, spherical collapse theory, and the properties of Gaussian random fields. Suppose that we have smoothed the initial density distribution on a scale R using some spherically symmetric window function $W_M(r)$, where $M(R)$ is the average mass contained within the window function. There are various possible choices for the form of the window function (cf. LC93), and the relation between M and R will clearly depend on this choice. We use a real-space top–hat window function, $W_M(r) = \Theta(R - r)(4\pi R^3/3)^{-1}$, where Θ is the Heaviside step function. In this case $M = 4\pi\rho_0 R^3/3$, where ρ_0 is the mean mass density of the universe. The mass variance $S(M) \equiv \sigma^2(M)$ may be calculated from

$$\sigma^2(M) = \frac{1}{2\pi^2} \int P(k) W^2(kR) k^2 \, dk \,, \tag{1}$$

where $P(k)$ is the mass power spectrum, and $W(kR)$ is the Fourier transform of the real-space top-hat:

$$W(kR) = \frac{3[\sin(kR) - kR\cos(kR)]}{(kR)^3} \,. \tag{2}$$

The 'excursion set' derivation due to Bond et al. (1991) leads naturally to the extended Press–Schechter formalism that we will use extensively in this paper. The smoothed field $\delta(M)$ is a Gaussian random variable with zero mean and variance S. The value of δ executes a random walk as the smoothing scale is changed. Adopting an ansatz similar to that of the original Press–Schechter model, we associate the fraction of matter in collapsed objects in the mass interval $M, M + dM$ at time t with the fraction of trajectories that make their *first upcrossing* through the threshold $\omega \equiv \delta_c(t)$ in the interval $S, S + dS$. This may be translated to a mass interval through equation (1). The halo multiplicity function (here in the notation of LC93) is then

$$f(S, \omega)dS = \frac{1}{\sqrt{2\pi}} \frac{\omega}{S^{3/2}} \exp\left(-\frac{\omega^2}{2S}\right) dS \,. \tag{3}$$

The *conditional* mass function, the fraction of the trajectories in haloes with mass M_1 at z_1 that are in haloes with mass M_0 at z_0 ($M_1 < M_0$, $z_0 < z_1$) is

$$f(S_1, \omega_1 \mid S_0, \omega_0)dS_1 =$$
$$\frac{1}{\sqrt{2\pi}} \frac{(\omega_1 - \omega_0)}{(S_1 - S_0)^{3/2}} \exp\left[-\frac{(\omega_1 - \omega_0)^2}{2(S_1 - S_0)}\right] dS_1 \,. \tag{4}$$

The probability that a halo of mass M_0 at redshift z_0 had a progenitor in the mass range $(M_1, M_1 + dM_1)$ is given by (LC93)

$$\frac{dP}{dM_1}(M_1, z_1 \mid M_0, z_0)dM_1 =$$
$$\frac{M_0}{M_1} f(S_1, \omega_1 \mid S_0, \omega_0) \left|\frac{dS}{dM}\right| dM_1 \,, \tag{5}$$

where the factor M_0/M_1 converts the counting from mass weighting to number weighting.

All of the results presented in this paper have been calculated for an $\Omega = 1$ universe with $H_0 = 50$ km s^{-1} Mpc^{-1}. The power spectrum is obtained from the fitting formula of Bardeen et al. (1986) with $\Gamma = 0.21$ and $\sigma_8 = 0.6$. This is the τCDM model of Efstathiou, Bond & White (1992), and has been chosen because the slope of the power spectrum on galaxy scales is consistent with observations. However, our results are equally valid for any assumed power spectrum or cosmology. The one exception that we know of is the case of a universe with a 'hot' dark matter component, such as a massive neutrino (CHDM- or MDM-type models). The standard extended Press–Schechter formalism does not properly treat the evolution due to the changing free-streaming length of the neutrino in such models. All of the expressions given are valid for a general cosmology unless otherwise noted.

3 PREVIOUS METHODS

The development of techniques for constructing Monte Carlo realizations of the merging history of dark matter haloes (Cole & Kaiser 1988; Cole 1991; Kauffmann & White 1993; LC93) using the extended Press–Schechter formalism has allowed a great deal of progress to be made in the use of semi-analytic methods for studying galaxy formation and evolution. Kauffmann & White developed a method for constructing merger trees which addresses the problem of simultaneously reproducing the average number of haloes given by equation (5) and imposing the constraint that the

mass of a halo be equal to the sum of the masses of its progenitors at every stage. To do this, they impose a grid in mass and redshift. For the first step in redshift, they then create a list of haloes, where the number of haloes with mass M_i is given by $N(M_i) = N_{ens}\mathrm{d}N/\mathrm{d}M(M_i)\Delta M_i$, rounded to the nearest integer. N_{ens} is some large number of ensembles, typically $N_{ens} \sim 100$. The progenitors are randomly assigned to ensembles, starting with the largest and working in order of decreasing mass. The probability of assignment is proportional to the amount of remaining free mass, with the constraint that the total mass of the progenitors cannot exceed the mass of the parent[1]. This process is repeated for all the steps in the redshift grid.

This algorithm is guaranteed to reproduce exactly the mean number of haloes of each mass at each redshift for the set of ensembles. However, it is possible to encounter a situation in which the next halo does not fit into any of the ensembles. In order to get around this problem, mass conservation is enforced only in an approximate way (G. Kauffmann, private communication). There are also some practical drawbacks to this method. It is necessary to generate a large number of ensembles and store them, which is somewhat inconvenient. An arbitrary grid in halo mass and redshift must be imposed. Also, because the function $\mathrm{d}N/\mathrm{d}M$ is very sharply peaked around M_0 for small M_0 or small redshift intervals Δz, the algorithm as described in Kauffmann & White (1993) is sensitive to the binning used and is prone to numerical problems. The algorithm breaks down for certain choices of power spectrum. Finally, although the mean of the distribution is reproduced by construction, the partitioning of the haloes into individual ensembles is ad hoc and may or may not reproduce the higher moments of the distribution.

A different approach, referred to as the 'Block model', has been proposed by Cole & Kaiseer (1988) and Cole (1991). A major drawback with this approach is that the halo masses always grow in discrete steps of factors of 2. This is problematic for the purpose of semi-analytic galaxy formation modelling, in which one would like to follow individual galaxies with fairly fine time resolution. LC93 propose a generalization of the block model which removes this condition, but we show in Section 6.1 that this method produces halo mass distributions that are severely discrepant with the Press–Schechter model.

4 DIFFICULTIES OF TREES

The first choice we must make if we want to construct a merger tree is whether to start from high redshift and merge together small clumps until the desired redshift is reached (as in an N-body simulation or, presumably, in the real Universe), or to start from the present day and work backwards in time, 'disintegrating' the haloes into their progenitors like a film run backwards. The extended Press–Schechter formalism provides expressions applicable to both situations. An important consideration is that we would like to eventually anchor our approach using the observed $z = 0$ properties of galaxies, which are presumably the most secure. In addition, because we lack any information about spatial correlations, if we go forward in time, we do not know which small clumps to combine with which. It may be possible to get around this problem somehow, but for now we pursue the 'disintegration' approach, in which we postulate the existence of a 'parent' halo of a given mass M_0 at redshift z_0 and break it into its 'progenitors' working backwards in time.

It might appear a simple matter to construct a merger tree by simply picking the masses of the progenitors of our parent halo at some earlier redshift z_1 from the expression for $\mathrm{d}P/\mathrm{d}M$ $(M_1, z_1 | M_0, z_0)$ given by equation (5) above, then repeating this process starting from each progenitor in turn for the next step back in time. Two difficulties immediately arise in implementing this approach. First, from inspection of equation (5), the *number* of haloes clearly diverges as the mass goes to zero. However, note that the *mass* contained in small haloes (equation 4) does not diverge as $M \rightarrow 0$. In order to pick masses from the number-weighted probability function numerically, it is necessary to introduce a cut-off mass, or effective mass resolution, M_l.

The second problem is that the progenitor masses must simultaneously be drawn from the distribution $\mathrm{d}P/\mathrm{d}M(M)$ *and* add up to the mass of the parent, M_0. The problem is that $\mathrm{d}P/\mathrm{d}M$ is just the *average* number of haloes that one can make out of the mass $M_0 f(M) \mathrm{d}M$. What we really want is the *joint* probability function for the set of progenitors $\{M_1, \dots, M_n\}$, $\mathrm{d}P_n/\mathrm{d}M$ $(\{M_1, \dots, M_n\}, z_1 | M_0, z_0)$ with any value of $n < M_0/M_l$ [2]. An obvious problem with the use of the single halo probability rather than the joint probability is that there is no guarantee that we will not at some stage pick a progenitor that does not 'fit' in the halo: i.e., $M > M_0 - \sum_i M_i$, where M_i are the masses of all the previously picked progenitors. In addition, since the expression $P(M)$ gives only the probability that there was a progenitor of mass M at an earlier time, we do not know a priori how many progenitors were present at the redshift z_1.

Having noted these points, we can write down some basic requirements for our merger tree construction algorithm.

(i) The procedure must account for the mass contained in haloes below the mass resolution M_l, which we will refer to as the 'accreted mass'. However, the results must be independent of the value of M_l.

(ii) The procedure should treat all progenitors equally, independently of the sequence in which they are chosen.

(iii) The procedure must simultaneously reproduce the distribution of the number of progenitors and their masses, while conserving mass.

(iv) The algorithm should be numerically robust and must be possible to implement in a computationally efficient and convenient way.

We now demonstrate some problems that arise in several seemingly straightforward approaches to building the trees. The bold solid line in Fig. 1 shows the prediction of the extended Press–Schechter model for the quantity $\mathrm{d}N/\mathrm{d}M$, the number of progenitors with mass M for a parent halo with $M_0 = 5 M_l$ after a single step in redshift from $z_0 = 0$ to $z_1 = 0.2$ (also called the conditional mass function). This is the quantity that a successful merging tree method must reproduce. In this figure and hereafter unless otherwise noted, all masses are given in units of the mass resolution M_l. Here we have used $M_l = 1.0 \times 10^{10} \mathrm{M}_\odot$, but the results are independent of this value.

In our first attempted method, the progenitor masses are chosen from equation (5) until the mass reservoir M_0 is exhausted. The probability is set to zero for $M < M_l$. We have tried two ways of addressing the problem of mass 'overflow' described above. One approach is to choose progenitors until the total mass exceeds M_0,

---

[1] Note that in our terminology, the 'parent' is younger than its 'progenitors', because we always work backwards in time.

[2] From now on we will drop the differential notation $\mathrm{d}P/\mathrm{d}M$ and refer to the probability given by equation (5) as simply $P(M)$. We will also frequently drop the explicit dependence on the redshift and the parent mass M_0 where this is unambiguous.
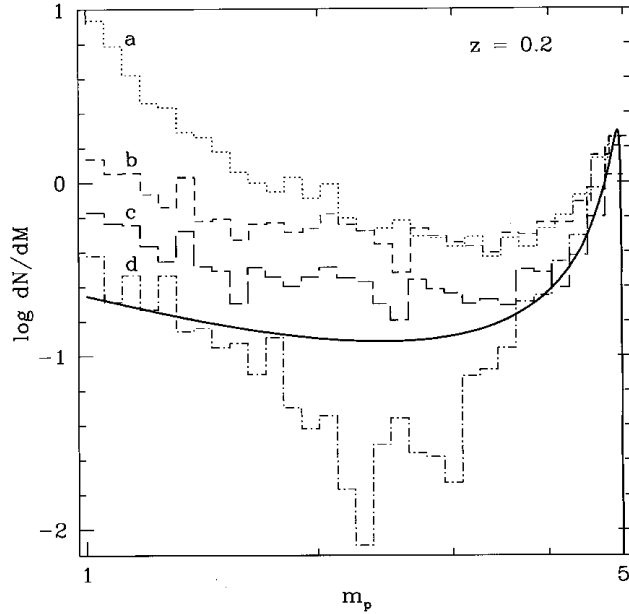
Figure 1. The number of progenitors with mass m_p (conditional mass function) for a single step in redshift ($z_0 = 0$ to $z_1 = 0.2$). The bold solid line is the prediction of the extended Press–Schechter theory (equation 5). The histograms were obtained by picking masses from the distribution $P(M, z_1 | M_0, z_0)$ until the parent mass M_0 was exhausted. The dotted histogram (a) shows the results using the Naive Method with Cut-off, and the short-dashed histogram (b) shows the results of using the Naive Method with Truncation (see text). The long-dashed histogram (c) uses Accretion Model Method 1, and the dot-dashed histogram (d) uses Accretion Model Method 2 (see text). All masses are in units of the minimum progenitor mass M_l.

and then to truncate the mass of the last progenitor. We will refer to this as the Naive Method with Truncation'. Another approach is to impose an upper mass cut-off, so that the probability is effectively set to zero for any values of M that exceed the available mass. In effect, we then choose the ith progenitor from the modified distribution

$$P_i(M_i) = P(M_i)\Theta(M_i - M_l)\Theta\left(M_0 - \sum_{j=1}^{i} M_j\right), \qquad (6)$$

where $P(M)$ is the original probability function from equation (5) renormalized for its new range. No attempt is made to compensate for the contribution of masses below M_l. We refer to this as the 'Naive Method with Cut-off'. The results for the quantity dN/dM (averaged over many ensembles) is shown by the histograms marked a (Cut-off) and b (Truncation) in Fig. 1.

Both procedures clearly fail to reproduce correctly the conditional mass function predicted by the extended Press–Schechter model (equation 5). Although the introduction of the upper mass cut-off to prevent choosing masses that are too large to fit in the current ensemble is somewhat more elegant than the brute force truncation, we see that it produces a shift towards smaller masses, which leads to large excess of small-mass haloes. It should be noted that this figure shows only one step in redshift. The excess multiplies with each subsequent step in redshift, so that even a relatively small discrepancy quickly becomes very serious. It should also be noted that these problems are most pronounced when $M_0 \approx M_l$ (as in the case shown in the figure). If $M_0 \gg M_l$, the disagreement is not as bad.

5   THE ACCRETION MODEL

We must introduce a self-consistent way of treating progenitors above and below the cut-off mass, M_l. It should be noted that although the contribution of masses $M < M_l$ may be negligible for haloes with $M_0 \gg M_l$, as the mass of the parent approaches the resolution limit it necessarily becomes a significant fraction of the total progenitor mass. Because every halo, regardless of its size, must be broken down into smaller and smaller pieces until all of the pieces fall below the mass resolution (this is what makes the process finite), the correct treatment of small haloes is crucial for reconstructing the formation history of haloes of all masses out to arbitrarily high redshift.

The basic idea behind this approach is never to treat any mass below M_l in terms of progenitor number, but rather to find a complimentary description for it as accreted diffuse matter. We now introduce an arbitrary distinction in terminology to reflect this division. Let *progenitors* by definition have mass greater than the fixed mass resolution M_l. The aggregate contribution of all haloes with $M < M_l$ will be referred to as *accreted mass*. A fully rigorous procedure should use the *joint* probability for progenitor number (not mass) above M_l, and accreted mass (not number of haloes) below this mass scale.

We can now define a few more useful quantities. Given the mass of the parent halo M_0 and the redshift step $z_0 \rightarrow z_1$, the average number of progenitors (recalling our definition above), is

$$\bar{N} \equiv \langle N_p(M | M_0) \rangle = \int_{M_l}^{M_0} dM \frac{M_0}{M} P(M, z_1 | M_0, z_0). \qquad (7)$$

We can also calculate the average fraction of M_0 that dwelt in the form of progenitor haloes of mass $M > M_l$:

$$\bar{f}_p = \int_{M_l}^{\infty} dM \, P(M, z_1 | M_0, z_0), \qquad (8)$$

and the complimentary quantity for the average fraction of M_0 that came from 'accreted' mass, $\bar{f}_{acc} = 1 - \bar{f}_p$.

Before we proceed, we would like to warn the reader that the contents of the remainder of this section are rather detailed and probably of interest only to the specialist. The formalism developed in the rest of this section is not used directly in the successful method that we will eventually derive. We urge the impatient reader to skip directly to Section 6.

From the above predictions we can try to evaluate what went wrong with our previous procedures (the 'Naive' methods). The predicted average number of progenitors for the case considered in Fig. 1 ($M_0 = 5M_l$, $z_0 = 0$, $z_1 = 0.2$) is $\bar{N} = 1.14$. The actual average number for 100 Monte Carlo realizations using the Naive Method with Truncation is $\bar{N} = 2.1$, and for the Naive Method with Cut-off it is $\bar{N} = 2.3$. Thus we see that the mean of the distribution is shifted towards larger numbers of small-mass haloes. This motivates our goal in this section, which is to find the probability function $\mathcal{P}_N$ of having N progenitors at redshift z_1 given the mass M_0 at a later redshift z_0, with the imposed cut-off of M_l.

Given M_0, z_0, and z_1, let $M_{1,p}$ be the mass of a progenitor, where by definition $M_{1,p} > M_l$. However, during the time interval $\Delta z \equiv z_1 - z_0$, this progenitor accretes a mass $M_{1,acc}$ due to merging with small haloes of mass $M < M_l$ which are not counted as progenitors. Therefore its effective contribution to M_0 is $M_1 = M_{1,p} + M_{1,acc}$. We now define a modified probability function

$$\tilde{P}(M_1 | M_0) = \int_{M_l}^{M_1} dM_{1,p} \, P(M_{1,p} | M_0) P_{acc}(M_{1,acc} | M_{1,p}). \qquad (9)$$

The weighting function P_{acc} is proportional to the probability for $M_{1,p}$ to accrete a mass $M_{1,acc}$ during the specified redshift interval. This probability *is not* simply $P(M_{1,acc}, z_1|M_0, z_0)$, because M_{acc} will in general be comprised of many small haloes. It should be noted that for the same reason, M_{acc} is not necessarily smaller than M_l. We return to the determination of the function P_{acc} in a moment.

The probability for having one and only one progenitor (with additional accreted mass) given a halo of mass M_0 is

$$\mathcal{P}_1 = \tilde{P}(M_0|M_0) \,. \tag{10}$$

The probability for exactly two progenitors is

$$\mathcal{P}_2 = \int_{M_l}^{M_0-M_l} dM_1 \, \tilde{P}_2(M_1|M_0) \,, \tag{11}$$

where

$$\tilde{P}_2(M_1|M_0) \equiv \tilde{P}_1(M_1|M_0)\tilde{P}_1(M_0 - M_1|M_0) \,. \tag{12}$$

The generalization to N progenitors $\mathcal{P}_N$ is obtained recursively via

$$\mathcal{P}_N = \int_{M_l}^{M_0-(N-1)M_l} dM_1 \tilde{P}_N(M_1, ..., M_N|M_0) \tag{13}$$

$$= \int_{M_l}^{M_0-(N-1)M_l} dM_1 \, \tilde{P}_1(M_1|M_0)$$

$$\dots \int_{M_l}^{M_0-\sum_{i=1}^{j-1} M_i - (N-j-1)M_l} dM_j \tilde{P}_1(M_j|M_0)$$

$$\dots \int_{M_l}^{M_0-\sum_{i=1}^{N-2}(M_i)-(N-2)M_l} dM_{N-1}$$

$$\times \tilde{P}_1(M_{N-1}|M_0) \, \tilde{P}_1\Big(M_0 - \sum_{i=1}^{N-1}(M_i)\Big| M_0\Big) \,. \tag{14}$$

The probability of having no progenitors of mass bigger than M_l, $\mathcal{P}_0$, is evaluated at the end by the requirement of

$$\sum_{i=0}^{N} \mathcal{P}_i = 1 \,, \tag{15}$$

where N is sufficiently large that $\mathcal{P}_N \to 0$. The average number of progenitors must satisfy

$$\langle N_p \rangle = \sum_{n=0}^{N} n \mathcal{P}_n \,, \tag{16}$$

and may be compared with the independent prediction of equation (7).

5.1 The accretion probability

The accretion weighting function $P_{acc}(M_{acc} | M_p; z_1, z_0)$ is an important missing ingredient in these expressions. It is proportional to the probability for a progenitor with mass M_p to accrete a mass M_{acc} during the redshift interval $\Delta z = z_1 - z_0$. It should reflect our expectation that it is very unlikely for a small-mass halo to accrete a very large amount of mass, as this would require the simultaneous merging of a very large number of haloes with $M < M_l$. Similarly, a large halo will be unlikely to accrete a very small amount of mass, because its cross-section for merging is large. In this section we incorporate these qualitative expectations into a reasonable guess for the accretion probability function P_{acc}.

Consider a halo with a mass M_1 at a redshift z_1. From the spherical collapse model (e.g. White & Frenk 1991), we expect the virial mass to increase due to the infall of previously

uncollapsed material. The mass at a later time corresponding to a redshift $z_2 < z_1$ is

$$M_2 = M_1 + \int_{z_1}^{z_2} dz \frac{dM}{dt}(M_1) \frac{dt}{dz} \,, \tag{17}$$

where the accretion rate from the spherical infall model is

$$\frac{dM}{dt}(M) = \frac{V_c^3}{2\pi G} \,, \tag{18}$$

where V_c is the circular velocity of the halo (the last formula is strictly true only in a universe with no cosmological constant, but is a good approximation even if $\Lambda \neq 0$). This change in mass includes mergers with haloes of all masses. We still need to estimate how much of the mass change $\Delta M = M_2 - M_1$ is due to mergers with haloes with mass less than the resolution limit, i.e., 'accretion'. To do this we use the expression for the mean fraction of accreted mass ($\bar{f}_{acc} \equiv 1 - \bar{f}_p$), starting from the mass M_2 and going back in time from z_2 to z_1. In this way we estimate the average mass accreted by M_1 to be $\bar{M}_{acc} = \bar{f}_{acc}(M_2) M_2$.

For higher moments of the distribution function $P_{acc}(M_{acc}|M_p; z_1, z_0)$ we shall assume that the accretion is mainly due to the infall of blobs (i.e., haloes of mass $< M_l$) with typical mass $M_b \ll M_l$. For the average accretion of equation (17) we expect

$$N_b \simeq \frac{\bar{M}_{acc}}{M_b} \,. \tag{19}$$

For any total accreted mass, regardless of the value of M_b, the number of blobs is proportional to the average accreted mass. If the number of the blobs is Poisson-distributed, then the second moment of the accreted mass distribution function is proportional to the total mass accreted. In the limit $N_b \gg 1$ the distribution should approach a Gaussian distribution. We have already argued that the mean of this distribution should be $\bar{M}_{acc}$. We can make a rough guess for the width of the distribution, $\sigma_{acc}^2 = \beta(\Delta M - \bar{M}_{acc})$, where ΔM represents the mass change predicted by the spherical infall model. This should be an upper limit on the accreted mass. Due to the uncertainties involved in the derivation of this expression, the parameter β is left free and can be tuned as needed (we used $\beta = 2$ for the results presented here). We now have a reasonable guess for the functional form of the accretion probability function for a progenitor of mass M_p over the redshift interval z_1 to z_0:

$$P_{acc}(M_{acc}|M_p; z_1, z_0) \propto \frac{1}{\sigma_{acc}^{1/2}} \exp\left[-\frac{(M_{acc} - \bar{M}_{acc})^2}{2\sigma_{acc}^2}\right]. \tag{20}$$

Although this expression is admittedly ad hoc, one can see that it contains the correct qualitative behaviour. The mean accreted mass increases with the progenitor mass and with Δz as expected.

5.2 Merger trees with the accretion model

We can now imagine a new approach for constructing the merger trees which addresses the two main sources of the problems in the previous approach: the failure to account for accreted mass, and the incorrect distribution of the number of progenitors. Given the probability function $\mathcal{P}_N$ for each M_0 and time-step, we pick the number of progenitors from this distribution. We assign a mass to each of these progenitors from the distribution $P_i(M)$ (equation 6) as before. The Θ function prevents us from choosing a mass larger than the available mass at any stage. The accreted mass is automatically obtained from the residue of this procedure. The results of this algorithm (which we will call Accretion Model Method 1) are shown by the histogram c in Fig. 1. We see that the results have

improved dramatically from the Naive Method with Truncation, where the upper mass cut-off was also used but the number of progenitors was not specified. Also the average number of progenitors is now $\bar{N} = 1.08$, in much better agreement with the expected value of $\bar{N} = 1.14$. However, there is still an inconsistency in this procedure, which leads to the remaining discrepancy. We have still assigned the mass to the progenitors based on the *single* halo probability function $P(M)$, when, as we have argued before, we really should have used the joint probability function $P_n(\{M_1, \ldots, M_n\})$. This can, in principle, be calculated using the same approach that we used to obtain the $\mathcal{P}_N$ function. For example, the joint probability for $N = 2$, taking into account the accretion weighting, is just the integrand of the expression for $\mathcal{P}_2$:

$$
\begin{aligned}
P_2(\{M_1, M_2\}) = {} & P(M_1)P(M_2)\,\Theta(M_0 - M_1 - M_2) \\
& \times \int_0^{M_0 - M_1 - M_2} \mathrm{d}M_{\mathrm{acc}}\, P_{\mathrm{acc}}(M_{\mathrm{acc}}\,|\,M_1) \\
& \times P_{\mathrm{acc}}(M_0 - M_1 - M_2 - M_{\mathrm{acc}}\,|\,M_2)\,.
\end{aligned}
\tag{21}
$$

This expression may be generalized to the N-halo joint probability as before.

We now pick the number of progenitors from $\mathcal{P}_N$ and assign the masses from the joint probability function equation (21) (Accretion Model Method 2). This approach goes a long way towards curing the problems we have noted, as we see from histogram d in Fig. 1. However, the shape of the mass function is not quite right. We attribute this to the inconsistency introduced by our ad hoc accretion probability weighting. We find that changing the form of this function significantly affects the results obtained for the mass function. We obtain better results for a lognormal distribution than for the Gaussian distribution used here. If we could somehow obtain some external information on the form of the accretion weighting, for example from N-body simulations, it might be possible to produce a successfully working method. However, this scheme is also rather cumbersome and computationally expensive. The calculation of $\mathcal{P}_N$ involves the computation of $(N-1)!$ integrals, and must be repeated for every parent halo mass, M_0 and redshift interval Δz. The joint probability function for the ith progenitor will depend on the masses of the previously chosen progenitors, and thus must be recalculated at each stage. This quickly becomes prohibitively computationally expensive when large numbers of progenitors are allowed.

We can address the second problem by reducing the time-step or redshift interval Δz. As Δz is decreased, the form of $\mathcal{P}_N$ steepens and becomes peaked at smaller N. For a small enough choice of Δz, $\mathcal{P}_{N \geq 3} \to 0$. We will refer to this condition as the two-progenitor limit. The idea is not to allow any processes that involve more than a single bifurcation. This demand allows one to calculate only three functions for each time-step ($\mathcal{P}_0$, $\mathcal{P}_1$, & $\mathcal{P}_2$), by using the analytic model of equation (14). At each stage, the time-step will now depend on the parent mass M_0. The larger the halo, the smaller the time-step necessary to satisfy this condition.

Although going to the two-progenitor limit might make the procedure computationally feasible, for the moment our lack of knowledge about the accretion probability weighting, and the sensitivity of the results to this function, lead us to abandon this approach. Perhaps the formalism we have developed here, and the simple approach we have presented for modelling the joint probability function, can be refined in the future. However, we do not pursue it in this paper.

6 A PRACTICAL SOLUTION

6.1 Binary merger trees without accretion

In the absence of external information about the behaviour of the accreted mass component, we are forced to treat it within our Monte Carlo procedure. As we have discussed, we do not want to use the number-weighted probability for masses below M_l because of the divergence of this expression at small masses. However, the mass-weighted probability, equation (4), does not diverge. If we pick a mass M_1 from the mass-weighted expression $f(M_1, z_1\,|\,M_0, z_0)$, this is equivalent to discovering that a single trajectory, or particle, from the parent halo M_0 was in a halo with mass M_1 at z_1. Once again, this is the single trajectory probability, and if we continue to select masses they will not in general fit together in any sensible combination that can lead to M_0. We attempt to evade this problem by choosing a very small time-step and so going to the two progenitor limit, as before. It is convenient to use $\omega \equiv \delta_{\mathrm{c}}(z) = \delta_{\mathrm{c},0}/D(z)$ as our time variable, and $S(M) \equiv \sigma^2(M)$ as our mass variable, as in LC93. These can be translated back to redshift and mass by inversion of the appropriate expressions. The probability for a step ΔS in a time-step $\Delta\omega$ is (LC93, equation 2.29)

$$
P(\Delta S, \Delta\omega)\mathrm{d}\Delta S = \frac{1}{\sqrt{2\pi}} \frac{\Delta\omega}{(\Delta S)^{3/2}} \exp\left[-\frac{(\Delta\omega)^2}{2\Delta S}\right] \mathrm{d}\Delta S.
\tag{22}
$$

If we make a change in variables, $x \equiv \Delta\omega/(2\sqrt{\Delta S})$, this becomes a Gaussian distribution in x with zero mean and unit variance. We can see from this expression that if we choose the time-step such that

$$
\Delta\omega \lesssim \sqrt{\frac{\mathrm{d}S}{\mathrm{d}M}(M_0)\Delta M_{\mathrm{c}}},
\tag{23}
$$

where $\Delta M_{\mathrm{c}} \ll M_0$, then a step ΔS corresponding to a change in mass ΔM larger than the mass resolution becomes a 2σ event. We must choose this time-step carefully – if it is too big, then the two-progenitor approximation will break down badly. If it is too small, then the results become dominated by numerical noise. The above expression is approximate, but provides a rule of thumb. Note that it scales with M_0 through the differential $\mathrm{d}S/\mathrm{d}M(M_0)$, so larger parent haloes will require smaller time-steps.

In the simplest version of this algorithm, we start from a parent halo with mass M_0 at z_0 and obtain the time-step $\Delta\omega$ from equation (23). We work backwards in time from this point. We choose a Gaussian random variable with unit variance and translate this to a step ΔS using the transformation mentioned above. The new halo mass at the earlier time $t(\omega + \Delta\omega)$ is then $M(S + \Delta S)$. LC93 argue that for a small enough time-step, all mergers may be treated as binary. This makes the process very simple – at each stage we break the halo into two pieces with mass M and $\Delta M \equiv M_0 - M$, where M is chosen from the probability function $f(M)$. If the progenitor obtained in this way is larger than M_l, we treat it as the next parent and repeat the procedure. If it is smaller than M_l, then we treat it as accreted mass and do not follow its history. This is essentially the same algorithm proposed in LC93 at the top of page 641, and is similar to a generalized version of the block model of Cole & Kaiser (1988) and Cole (1991).

This approach has several advantages. It is simple and may be coded recursively in a few lines. Because it mainly involves picking Gaussian random deviates, it is also very fast. Rather than being imposed on an artificial grid in redshift like previous methods, it reflects the intrinsic merging time-scales of haloes of different mass contained in the extended Press–Schechter theory. Unfortunately, the mass function of haloes obtained in this way begins to develop
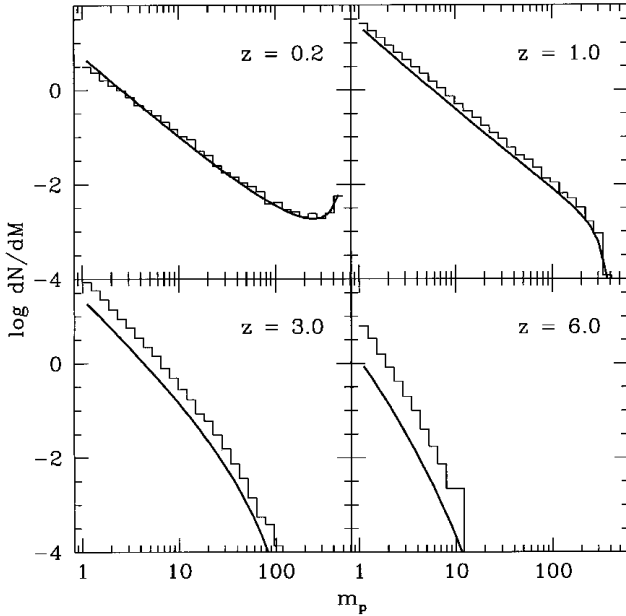
Figure 2. The number of progenitors with mass M for a halo with initial mass $M_0 = 500$, at various redshifts as shown in the figure. The solid lines are the predictions of the extended Press–Schechter theory. The histograms show the results for merger trees constructed using the Binary Tree Method without accretion, which is essentially the algorithm proposed by Lacey &Cole (1993). The trees have an excess of haloes compared to the Press–Schechter model, and the discrepancy increases with redshift.
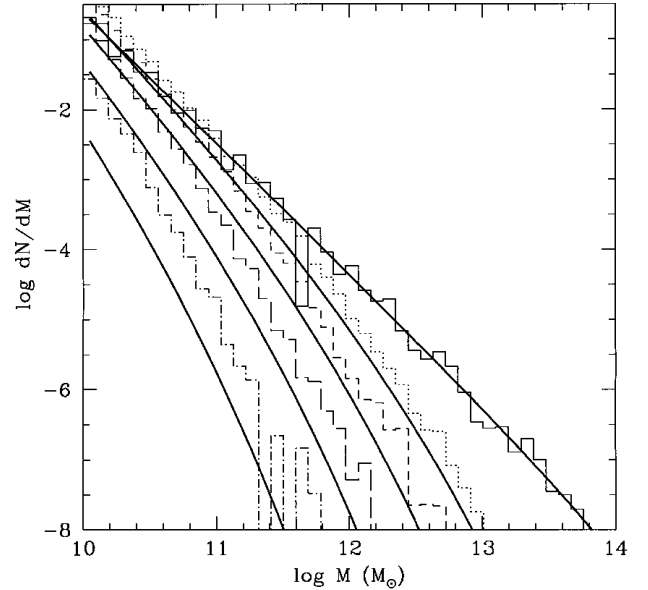

Figure 3. The total mass function obtained from the Binary Tree Method without accretion. All mergers are assumed to involve exactly two haloes. The total mass function is obtained by combining a grid of haloes from $1M_l$ to $5 \times 10^4 M_l$ ($1.0 \times 10^{10} M_\odot$ to $5.0 \times 10^{15} M_\odot$) weighted with the Press–Schechter number density at $z = 0$. The bold lines are the predictions of the Press–Schechter theory, at $z = 0.16, 2.5, 3.6, 5$ and 7 from top to bottom. The merger trees (histograms) overpredict the number of haloes by more than an order of magnitude after many steps in redshift.

an excess of haloes compared to the extended Press–Schechter model. This excess becomes more and more severe as the number of steps increases. We show this for the conditional mass function for a halo with an initial mass $M_0 = 500 M_l$ in Fig. 2. This problem becomes quite serious when we combine the merger histories of a grid of haloes, with the appropriate Press–Schechter weighting for the parent at $z = 0$, to obtain the total mass function, shown in Fig. 3. The mass function reconstructed in this way should yield exact agreement with the original Press–Schechter expression for the universal mass function. With this method, the number of haloes is overpredicted by almost an order of magnitude by a redshift of $z \sim 6$. This is especially troublesome as the Press–Schechter model already gives an $\sim$30-50 per cent excess of small haloes compared to N-body simulations (Lacey &Cole 1994; Gross et al. 1998). The problem that we demonstrate here may explain why LC93 and Lacey & Cole (1994) find that a similar Monte Carlo method leads to halo formation times that are 40 per cent higher than the analytic predictions or the N-body results.

We believe that this problem is due to the simplifying assumption of binary mergers. Because the merger rate of very small haloes becomes effectively infinite for CDM-like power spectra, non-binary mergers, at least involving small haloes, cannot be neglected. This statement is complementary to our original premise regarding the importance of what we have called accreted mass. One might think that we have simply not chosen a small enough time-step, but if this were the case the results should improve steadily as we decrease the time-step. We do not observe this behaviour even for extreme reductions in the time-step.

Another way of stating the problem is that we have actually violated item (ii) in our list in Section 4. The first mass is chosen from the distribution $f(M)$, but the mass of the second progenitor is not; it is just assumed to be whatever mass is left over. This means that for every progenitor with mass M we always get a progenitor

with mass $M_0 - M$. It is easy to see that this will lead to inconsistencies with the mean distribution function $P(M)$.

6.2 Binary trees with accreted mass

We now attempt to cure the problem noted above by relaxing the simplifying assumption of binary mergers from the previous subsection. Namely, we postulate that mergers can involve at most two *progenitors* (haloes with $M > M_l$), but an arbitrary number of haloes with mass less than M_l. This, of course, amounts to allowing for accreted mass. At any branching we may have only accreted mass (zero progenitors), or alternatively one or two progenitors plus accreted mass. In addition, the progenitor masses must always be picked from the probability distribution $f(M)$. Leftover mass can contribute to accretion, but cannot be used for progenitors.

The new recipe is as follows. The algorithm is shown in flow-chart form in Fig. 4. Given the parent mass M_0 we compute the time-step $\Delta\omega$ as before. Using this time-step throughout the following steps, we proceed as follows.

(i) Pick a mass M from the mass-weighted probability distribution equation (22). This mass can be anywhere in the range $0 \leq M \leq M_0$. If $M < M_l$, we count it as accreted mass. If $M \gtrsim M_l$, we count it as a progenitor.

(ii) Compute the unallocated mass $\Delta M = M_0 - M$.

(iii) If the unallocated mass ΔM is larger than M_l, then it may or may not contain a progenitor. To determine this, pick another mass M from the distribution, but with the restriction $M < \Delta M$. Depending on its mass, count it as accreted mass or a progenitor as before. In either case, subtract M from the mass reservoir.

(iv) Repeat this process until either

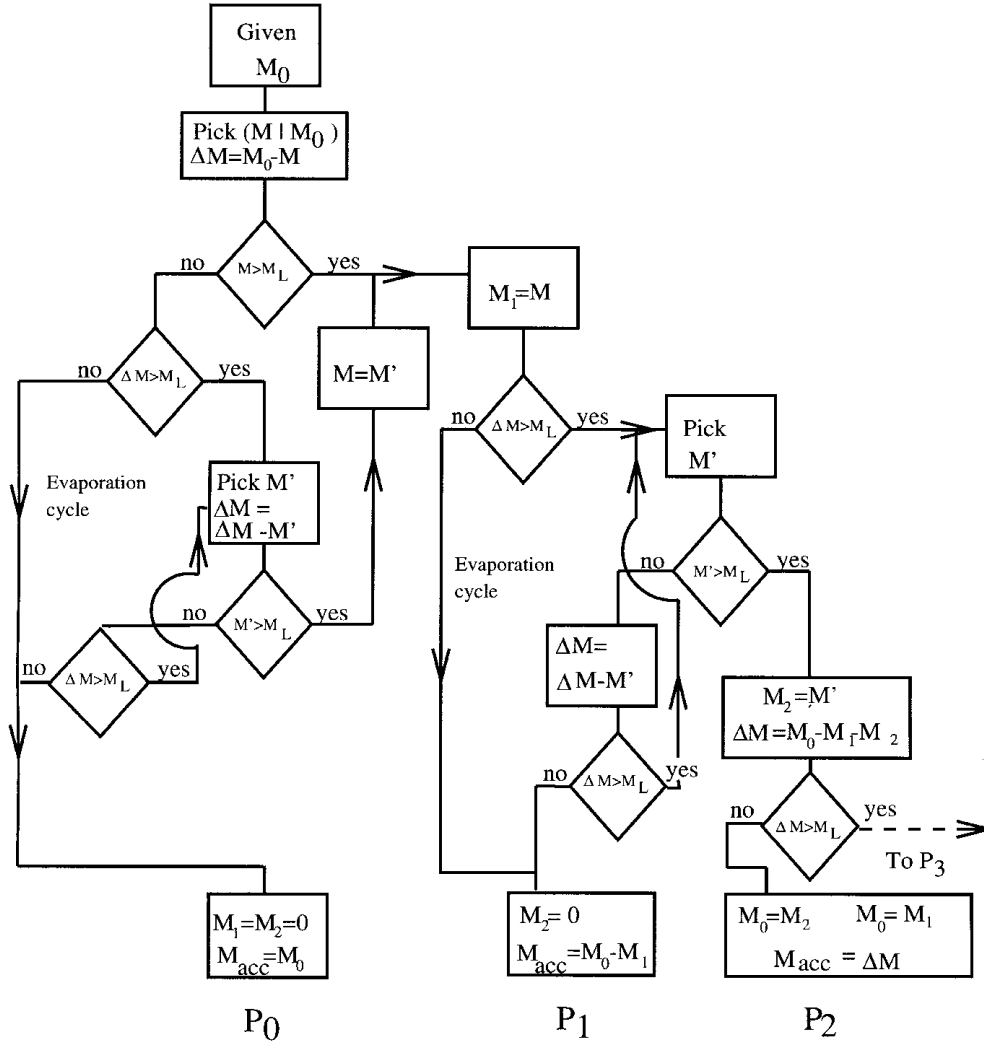(a) the mass reservoir ΔM falls below the minimum halo

Figure 4. A flow-chart for one redshift step of the disintegration of a halo in a merging tree. The end-points ($\mathcal{P}_1$, $\mathcal{P}_2$, etc.) lead to identical flow-charts for the subsequent time-step, with a new parent mass M_0. A detailed discussion of the algorithm and this flow-chart is given in Section 6.2.

mass M_l, in which case it must abandon any aspirations of harbouring a real progenitor and must be accreted mass, or

(b) we have found a total of two progenitors ($M > M_l$), in which case the remaining mass is considered to be accreted mass in accord with our ansatz.

(v) Each progenitor now becomes a parent, we calculate a new time-step, and repeat the whole process.

In the flow-chart, branches leading to the outcome of zero, one and two progenitors are labelled $\mathcal{P}_0$, $\mathcal{P}_1$, and $\mathcal{P}_2$, in connection with the formalism developed in the previous section.

Note that this procedure does not strictly fulfil requirement (ii) of Section 4 of equal treatment of progenitors, regardless of the order in which they are picked. This inequality is necessary due to the mass conservation requirement.

The results for the conditional mass function of a single halo are shown in Fig. 5. The discrepancy is now in the opposite direction: the number of haloes is underpredicted relative to the extended Press–Schechter prediction. Apparently this procedure now over-estimates the accreted mass. This is not too surprising, since we allowed large amounts of mass to be designated as accreted mass simply because two progenitors had already been found. It appears that even in the limit of small time-steps and for large-mass haloes,

mergers between more than two haloes cannot be neglected. In some ways this is not surprising either, because after all the division into $M > M_l$ and $M < M_l$ is arbitrary and has no physical basis.

6.3 The successful method: *N*-branch trees with accretion

It is trivial to generalize our previous recipe to allow an unrestricted number of progenitors. We now continue picking progenitor masses until the unallocated mass ΔM is less than M_l. This is indicated on the flow-chart by the dashed line labelled 'To $\mathcal{P}_3$'. Note that the total accreted mass can still exceed M_l, because some of the attempts to pick progenitors yield haloes with $M < M_l$ and contribute to the accreted mass. We still pick the time-step so that the number of progenitors cannot get too large (we find that we never exceed 10 progenitors per time-step even for cluster mass haloes ($M_0 = 5 \times 10^4 M_l$)). We find a good compromise between efficiency and accuracy if we introduce an additional scaling in the expression for $\Delta\omega$ from equation (23), of the form $b + a\log_{10}(M/M_l)$, where we have used the parameters $a = 0.3$ and $b = 0.8$ for the results shown here. This optimal scaling would change for a different power spectrum shape.

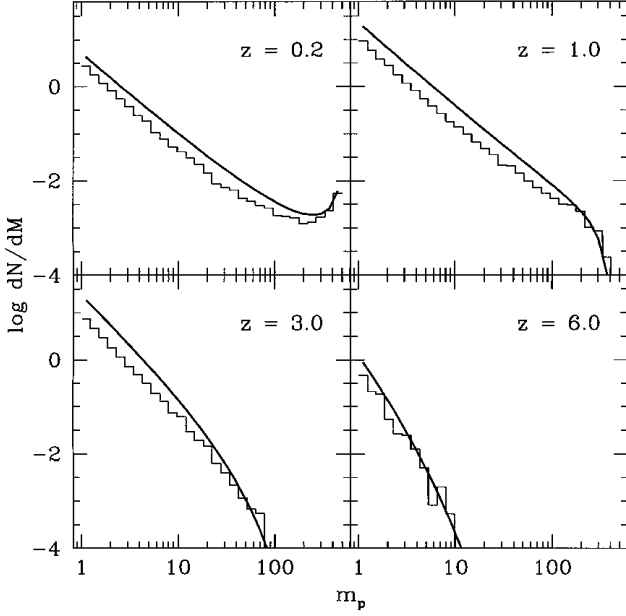This recipe gives good results for the conditional mass function

Figure 5. The number of progenitors with mass M, for a halo with $M_0 = 500$, using the Binary Tree Method with accretion. Here mergers may involve only two haloes with mass greater than M_l, but an indefinite number of mergers with haloes with mass smaller than M_l (accretion). As usual, the solid line is the extended Press–Schechter prediction. The trees (histograms) now underproduce haloes compared to the Press–Schechter model.

for parent haloes with a wide range of masses. We show this in Fig. 6 for parent haloes with $M_0 = 5\,M_l$, $500\,M_l$ and $5 \times 10^4\,M_l$. The agreement is poorest for haloes with $M_0 \lesssim 10\,M_l$. If we require strict mass conservation, this is an unavoidable problem due to the shape of the conditional mass function for haloes of this size. This should be kept in mind when setting the value of M_l — it should be chosen such that only objects larger than $\sim 10 M_l$ correspond to observable galaxies. We also check the mass-weighted quantity $\bar{f}_p$, the fraction of mass in progenitors as a function of redshift. This is shown in Fig. 7. This quantity shows good agreement for the smallest halo, $M_0 = 5\,M_l$, which shows that the accreted mass is being treated properly, so that we do not need to worry about the less than perfect agreement in the conditional mass function, as long as the condition on M_l mentioned above is satisfied. Note that the agreement of the mass function can be improved by adjusting the time-step, but at the expense of $\bar{f}_p$. We adjust the time-step to achieve the best possible agreement for both the number- and mass-weighted quantities, over the entire mass range. We compute the universal mass function from the weighted grid of merging histories constructed using our new scheme, and plot this in Fig. 8. We now find very good agreement with the prediction of the standard Press–Schechter theory.

We therefore conclude that although this method is not rigorous, it produces acceptable agreement with the mean quantities that we can check with the Press–Schechter model.

7 THE NUMBER–MASS DISTRIBUTION OF PROGENITORS

We have now developed a convenient and efficient method for constructing merger trees. The averages derived from an ensemble of these trees agree with the important mean quantities predicted by the extended Press–Schechter theory. However, part of the
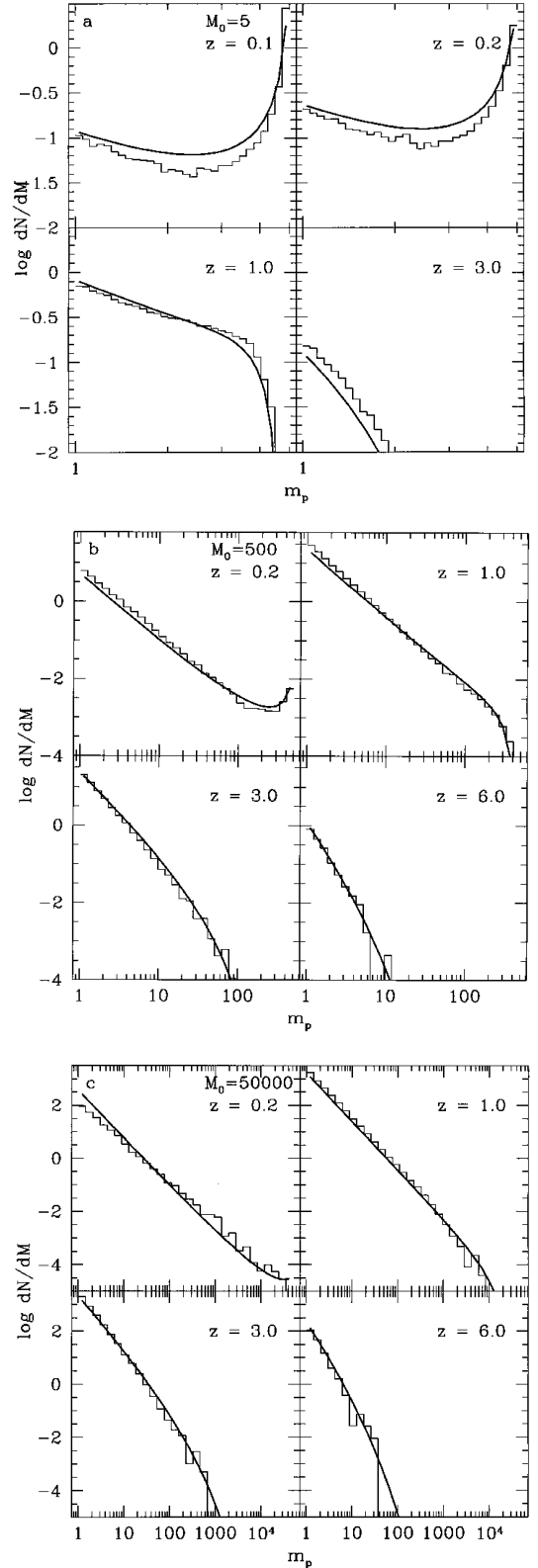

Figure 6. The number of progenitors with mass M, for a halo with initial mass M_0, using the N-Branch Tree Method with accretion. This is the same as the binary tree method with accretion, except that an arbitrary number of progenitors is allowed at each branching. The solid line is the extended Press–Schechter prediction. (a) $M_0 = 5M_l$; (b) $M_0 = 500M_l$; (c) $M_0 = 5 \times 10^4 M_l$. The merger trees (histograms) are in reasonably good agreement with the extended Press–Schechter model.
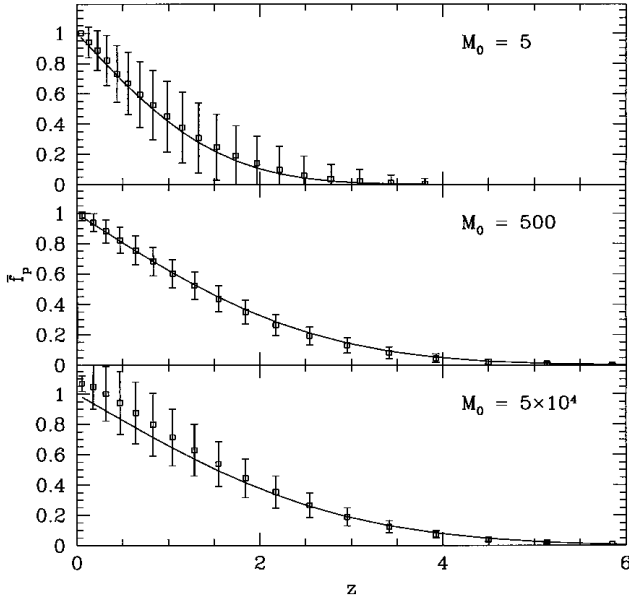
Figure 7. The average fraction of the original mass M_0 contained in progenitors (haloes with $M > M_l$) at a redshift z. The solid line shows the prediction of the Press–Schechter model. The square symbols show the average given by the merger trees (N-Branch Tree Method with accretion) Error bars show the standard deviation (1σ) over many ensembles.

Figure 9. The probability distribution of the number of progenitors, for a parent halo with mass M_0 at several redshifts. The histogram shows the results from the Monte Carlo merging trees (N-Branch Tree Method with accretion). The stars show the results of the semi-analytic model developed in Section 5. Left panel: $M_0 = 5M_l$. Right Panel: $M_0 = 50M_l$.

motivation for developing this new method was to ensure that the ensemble obeys the true joint probability distribution. We have not yet shown this to be the case, and we have already mentioned the lack of any information from the standard Press–Schechter formalism which would allow us to evaluate the Monte Carlo trees against analytic predictions.

We do have the predictions of the model we developed in Section 5, but we do not trust them for reasons discussed in that section. However, out of curiosity we compare the predictions of the

semi-analytic accretion model for the probability distribution of the number of progenitors, $\mathcal{P}_N$, with the results of the Monte Carlo merger trees. The integrals were performed numerically using a recursive, adaptive step-size Runga–Kutta algorithm. Fig. 9 shows the $\mathcal{P}_N$ distributions for haloes with mass $M_0 = 5$ and 50 at redshifts of 0.1, 0.2, 1 and 3. The results agree rather well for redshift steps of $\Delta z = 0.2$ or smaller. For larger steps in redshift, the semi-analytic results do not agree as well, probably due to the breakdown of the accretion model.

Fig. 10 demonstrates the importance of taking into account the joint distribution in the progenitor number – progenitor mass space. The figure shows strong correlations between the two variables. Some of the correlations are obvious: in Fig. 10(a), we show the distribution for a parent halo of $M_0 = 5\,M_l$. In the first redshift step ($z_0 = 0$, $z_1 = 0.2$) the highest probability is obtained for having a single progenitor, and this progenitor naturally contains a large fraction of the mass M_0. As we progress in redshift, this correspondence is not preserved. Accreted mass starts to be more and more significant, and the unseen accreted mass complements low-mass progenitors so that the sum may reach M_0. At the formation epoch of M_0, all $\mathcal{P}_N$ are populated, and in earlier stages most of the haloes go below M_l, when the dominant process is of single progenitors that accumulate accreted mass. It is interesting to notice that the 'formation epoch' (the earliest time when the largest progenitor has mass greater than $M_0/2$) is not dominated by mergers of equal-mass haloes, but rather approached via slow accretion. This general picture also remains valid for higher M_0 (Figs 10b and c): an increase in the number of progenitors occurs towards an intermediate redshift, and it then declines towards containing most of the mass in the accreted component. However, it should be noted that the highest mass considered here would be comparable to a galactic halo ($5 \times 10^{11} M_\odot$). For much larger mass haloes (comparable to group or cluster mass haloes), accretion is less important relative to the aggregation of roughly equal-mass progenitors.

The probability for the number of progenitors spans substantial

Figure 8. The total mass function obtained from the *N*-Branch Tree Method with accretion (the successful method). The solid lines show the prediction of the Press–Schechter theory, at $z = 0.16$, 2.5, 3.6, 5 and 7 from top to bottom. Broken histograms show the mass function from the merger trees, constructed as described in Fig. 3.
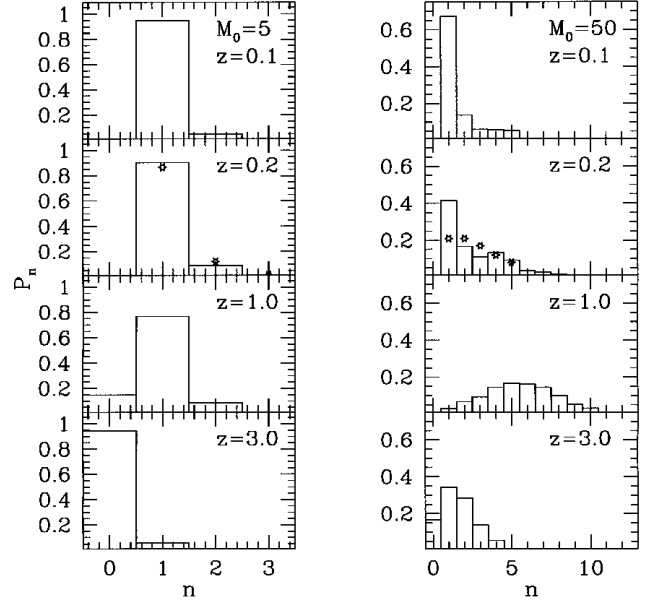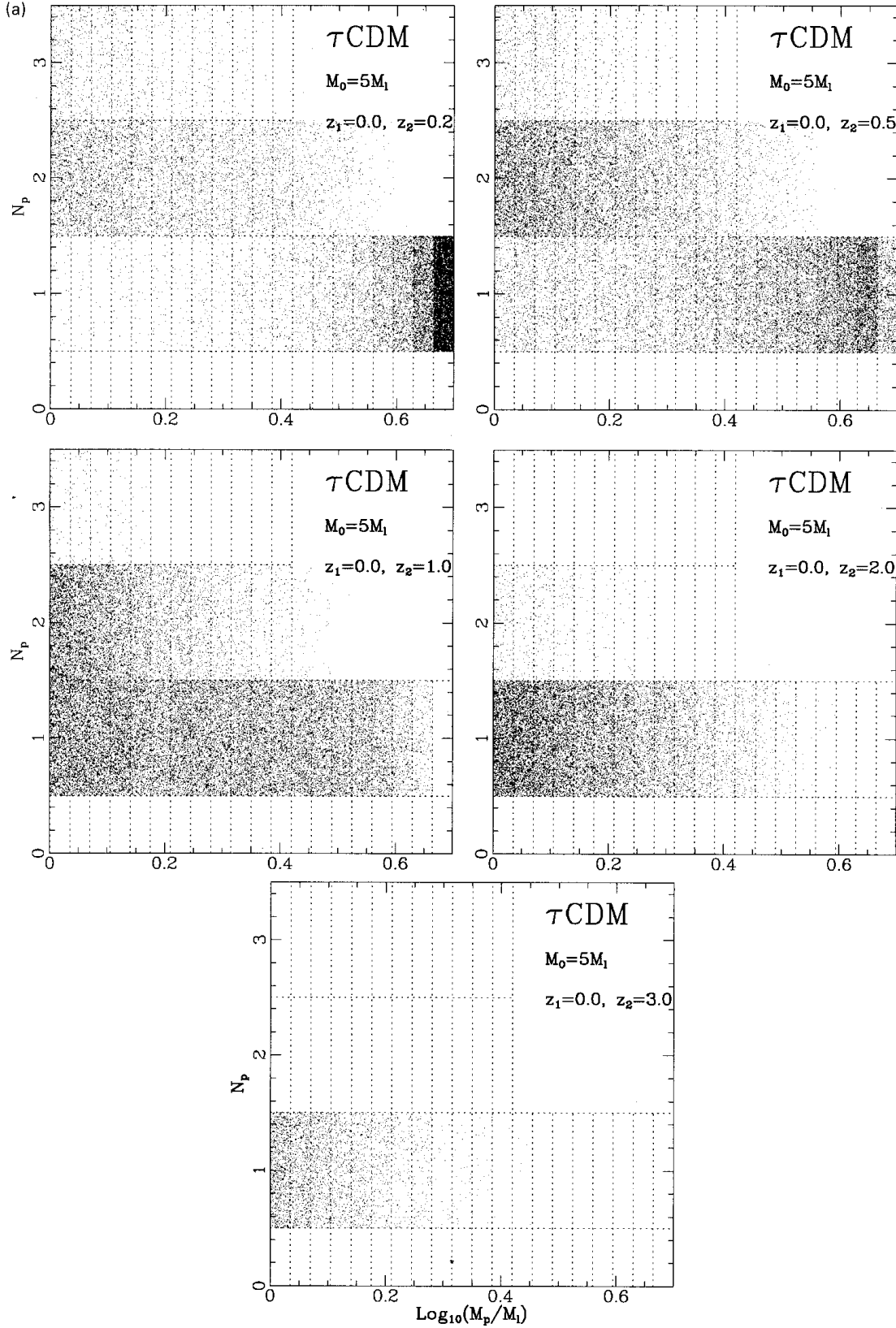
Figure 10. The evolution with redshift of the two-dimensional distribution of the number of progenitors and their masses for a halo of mass M_0, obtained from Monte Carlo realizations of the *N*-Branch Merger Trees with accretion (the successful method). The mass of the parent halo is (a) $M_0 = 5M_l$, (b) $M_0 = 50M_l$, (c) $M_0 = 500M_l$.
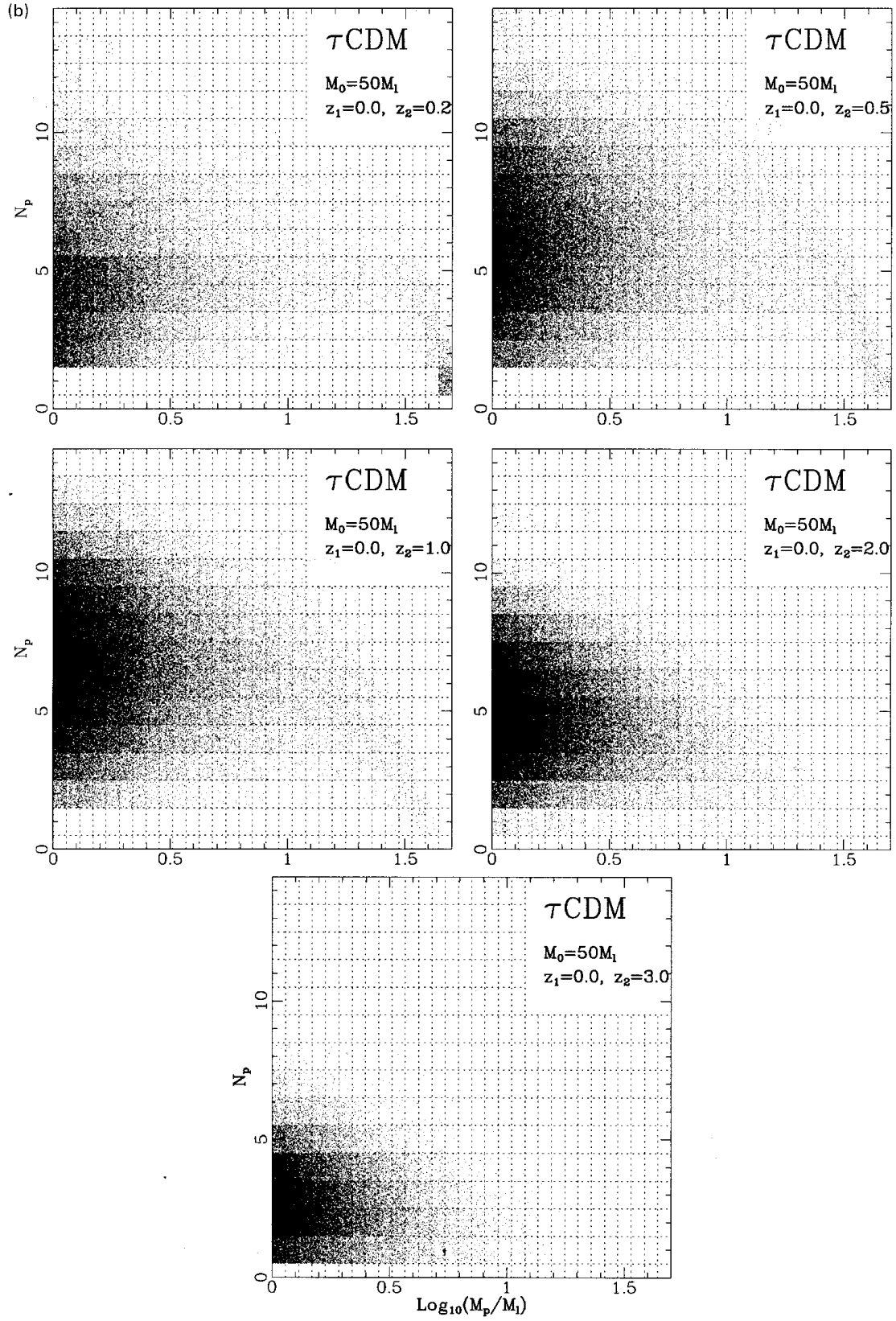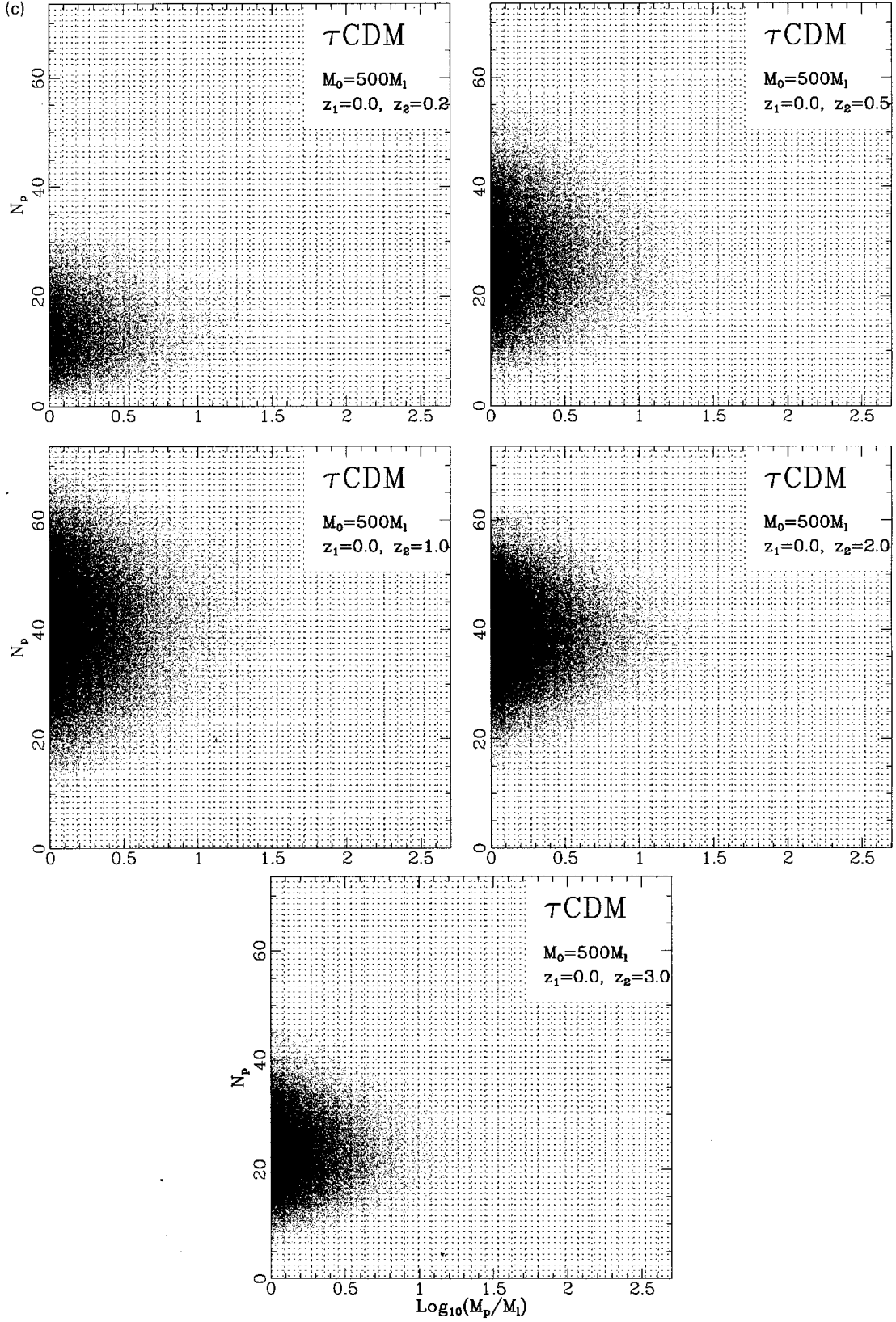
Figure 10 – *continued.*

Figure 10 – *continued.*

parts of its permitted range even for $M_0 = 500$. It is therefore clear why an infinitesimal step is needed for the two-progenitor scheme to work. As soon as we consider a finite time-step, the probability for $\mathcal{P}_{N>2}$ is no longer negligible.

The details of the redshift sequence represent the characteristics of the specific power spectrum and cosmology we used for the Monte Carlo realization. The formation time as a function of halo mass and redshift is determined by the cosmology and the power spectrum. The qualitative trend of this sequence, however, is similar for all cosmologies and all hierarchical power spectra.

Fig. 10 suggests that the progenitor mass and number distribution functions are an interesting avenue to pursue in the study of structure formation via merger trees. More importantly, it points at the existing interplay between the accreted mass, the progenitor mass and the number of progenitors, an interplay in which none of the three can be treated separately from the other.

8   SUMMARY AND CONCLUSIONS

We have presented a new method for constructing the merging history of dark matter haloes in a semi-analytic way. We have highlighted the need to impose an arbitrary mass cut-off for practical reasons, which leads us to distinguish between haloes above and below this threshold as 'progenitors' or 'accreted mass', respectively. The scheme we have proposed and implemented treats accreted mass and progenitors in a self-consistent way, and produces good agreement with the average quantities predicted by the underlying Press–Schechter theory, such as the conditional and universal mass function of haloes and the mean mass in progenitors as a function of redshift.

Our method is an improvement on the method proposed by Lacey & Cole (1993), which, after many steps in redshift, substantially overproduces haloes relative to the Press–Schechter mass function. Our work suggests that it is not possible to simultaneously conserve mass exactly and retain the exact agreement with the conditional mass function from the extended Press–Schechter model. The method of Kauffmann & White (1993) reproduces the conditional mass function exactly and conserves mass approximately. Our method conserves mass exactly and reproduces the conditional mass function approximately. This seems to be a necessary trade-off. Our method does have certain practical advantages: it does not require the creation and storage of a large number of ensembles, it is numerically robust, it does not require the imposition of a grid in mass or redshift, and it will work for any power spectrum.

We have pointed out the necessity of investigating the full probability distribution of the number of progenitors and their masses. This cannot be tested within the boundaries of the existing theory, and so must be examined by comparisons with *N*-body simulations. However, the simulations have their own problems and complications, such as the limitations of mass and spatial resolution and the ambiguities of defining haloes, so they should not be regarded as necessarily representing the absolute truth. In addition, the agreement between the simulations and the Press–Schechter model is only approximate, even for the mean quantities such as the mass function. It would therefore be desirable to have a reliable theoretical means of addressing this problem. We have attempted to

reformulate the extended Press–Schechter theory to obtain the full probability distribution for the number of progenitor haloes $\mathcal{P}_N$. Although this model gives qualitatively reasonable results for certain cases, some ingredients remain ad hoc.

For the moment, this leaves us with no recourse but to appeal to *N*-body simulations. In a companion paper (Somerville et al. 1999) we will compare the results we have obtained here with numerical simulations. This comparison has two goals: (a) to determine the quality of agreement of the analytic and Monte Carlo results with the *N*-body simulation results, and (b) to study the full distribution of the various quantities, and determine whether the Monte Carlo method developed here reproduces these results. The merger trees will then be used as the framework for the development of full semi-analytic galaxy formation models, also used to compare with a variety of observations (Somerville 1997; Somerville & Primack 1999).

REFERENCES

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bower R., 1991, MNRAS, 248, 332
Cole S., 1991, ApJ, 367, 45
Cole S., Kaiser N., 1988, MNRAS, 233, 637
Cole S., Aragón-Salamanca A., Frenk C., Navarro J., Zepf S., 1994, MNRAS, 271, 781
Efstathiou G., Frenk C. S., White S. D. M., Davis M., 1988, MNRAS, 235, 715
Efstathiou G., Bond J., White S., 1992, MNRAS, 258, 1
Eisenstein D., Loeb A., 1996, ApJ, 459, 432
Gelb J. M., Bertschinger E., 1994, ApJ, 436, 467
Gross M. A. K., Somerville R. S., Primack J. R., Holtzman J., Klypin A. A., 1998, MNRAS, 301, 81
Kauffmann G., 1996, MNRAS, 281, 487
Kauffmann G., White S., 1993, MNRAS, 261, 921
Kauffmann G., White S., Guiderdoni B., 1993, MNRAS, 264, 201
Lacey C., Cole S., 1993, MNRAS, 262, 627 (LC93)
Lacey C., Cole S., 1994, MNRAS, 271, 676
Ma C., 1996, ApJ, 471, 13
Press W., Schechter P., 1974, ApJ, 187, 425
Somerville R., 1997, PhD thesis, Univ. California, Santa Cruz, http://www.fiz.huji.ac.il/˜rachels/thesis.html
Somerville R., Lemson G., Kolatt T., Dekel A., 1999, MNRAS, submitted
Somerville R. S., Primack J. R., 1999, MNRAS, submitted
White S., Frenk C., 1991, ApJ, 379, 52

This paper has been typeset from a T$_E$X/L^{A}T$_E$X file prepared by the author.