Mon. Not. R. Astron. Soc. **386,** 577–592 (2008) doi:10.1111/j.1365-2966.2008.13075.x

The nearly universal merger rate of dark matter haloes in ΛCDM cosmology

Onsi Fakhouri⋆ and Chung-Pei Ma⋆
Department of Astronomy, 601 Campbell Hall, University of California, Berkeley, CA 94720, USA

Accepted 2008 February 5. Received 2008 February 4; in original form 2007 October 25

ABSTRACT

We construct merger trees from the largest data base of dark matter haloes to date provided by the Millennium Simulation to quantify the merger rates of haloes over a broad range of descendant halo mass ($10^{12} \lesssim M_0 \lesssim 10^{15}\,\mathrm{M}_\odot$), progenitor mass ratio ($10^{-3} \lesssim \xi \leqslant 1$), and redshift ($0 \leqslant z \lesssim 6$). We find the mean merger rate *per halo*, B/n, to have very simple dependence on M_0, ξ, and z, and propose a universal fitting form for B/n that is accurate to 10–20 per cent. Overall, B/n depends very weakly on the halo mass ($\propto M_0^{0.08}$) and scales as a power law in the progenitor mass ratio ($\propto \xi^{-2}$) for minor mergers ($\xi \lesssim 0.1$) with a mild upturn for major mergers. As a function of time, we find the merger rate per Gyr to evolve roughly as $(1+z)^{n_\mathrm{m}}$ with $n_\mathrm{m} = 2$–2.3, while the rate per unit redshift is nearly independent of z. Several tests are performed to assess how our merger rates are affected by e.g. the time interval between Millennium outputs, binary versus multiple progenitor mergers, and mass conservation and diffuse accretion during mergers. In particular, we find halo fragmentations to be a general issue in merger tree construction from *N*-body simulations and compare two methods for handling these events. We compare our results with predictions of two analytical models for halo mergers based on the extended Press–Schechter (EPS) model and the coagulation theory. We find that the EPS model overpredicts the major merger rates and underpredicts the minor merger rates by up to a factor of a few.

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

1 INTRODUCTION

In hierarchical cosmological models such as Lamdba cold dark matter (ΛCDM), galaxies' host dark matter haloes grow in mass and size primarily through mergers with other haloes. As the haloes merge, their more centrally concentrated baryonic components sink through dynamical friction and merge subsequently. The growth of stellar masses depends on both the amount of mass brought in by mergers and the star formation rates. Having an accurate description of the mergers of dark matter haloes is therefore a key first step in quantifying the mergers of galaxies and in understanding galaxy formation and growth.

Earlier theoretical studies of galaxy formation typically relied on merger trees generated from Monte Carlo realizations of the merger rates given by the analytical extended Press–Schechter (EPS; Bond et al. 1991; Lacey & Cole 1993) model (e.g. Kauffmann, White & Guiderdoni 1993; Somerville & Primack 1999; Cole et al. 2000). Some recent studies have chosen to bypass the uncertainties and inconsistencies in the EPS model by using halo merger trees from *N*-body simulations directly (Kauffmann et al. 1999; Benson et al. 2000; Helly et al. 2003; Kang et al. 2005; Springel et al. 2005). As we find in this paper, obtaining robust halo merger rates and merger trees requires rich halo statistics from very large cosmological simulations as well as careful treatments of systematic effects due to different algorithms used for e.g. assigning halo masses, constructing merger trees, removing halo fragmentation events and choosing time spacings between simulation outputs.

The aim of this paper is to determine the merger rates of dark matter haloes as a function of halo mass, merger mass ratio (i.e. minor versus major) and redshift, using numerical simulations of the ΛCDM cosmology. This basic quantity has not been thoroughly investigated until now mainly because large catalogues of haloes from finely spaced simulation outputs are required to provide sufficient merger event statistics for a reliable construction of merger trees over a wide dynamic range in time and mass. We achieve this goal by using the public data base of the Millennium Simulation (Springel et al. 2005), which follows the evolution of roughly 2×10^7 dark matter haloes from redshift $z = 127$ to 0. This data set allows us to determine the merger rates of dark matter haloes ranging from galaxy-mass scales of $\sim 10^{12}\,\mathrm{M}_\odot$ over redshifts $z = 0$ to ~ 6, to cluster-mass scales up to $\sim 10^{15}\,\mathrm{M}_\odot$ for $z = 0$ to a few. We are also able to quantify the merger rates as a function of the progenitor mass ratio ξ, from major mergers ($\xi \gtrsim 0.1$) down to minor mergers

⋆E-mail: onsi@berkeley.edu (OF); cpma@berkeley.edu (C-PM)

of $\xi \sim 0.03$ for galaxy haloes and down to $\xi \sim 3 \times 10^{-4}$ for cluster haloes.

The inputs needed for measuring merger rates in simulations include a catalogue of dark matter haloes and their masses at each redshift, and detailed information about their ancestry across redshifts, i.e. the merger tree. Unfortunately there is not a unique way to identify haloes, assign halo masses and construct merger trees. In this paper we primarily consider a halo mass definition based on the standard friends-of-friends (FOF) algorithm and briefly compare it with an alternative mass definition based on spherical overdensity.

For the merger trees, we investigate two possible algorithms for treating events in which the particles in a given progenitor halo end up in more than one descendant halo ('fragmentations'). We find that these events are common enough that a careful treatment is needed. In the conventional algorithm used in the literature, the progenitor halo is linked one-to-one to the descendant halo that has inherited the largest number of the progenitor's particles. The ancestry links to the other descendant haloes are severed (for this reason we call this scheme 'snipping'). We consider an alternative algorithm ('stitching') in this paper, in which fragmentations are assumed to be artefacts of the FOF halo identification scheme. We therefore choose to recombine the halo fragments and stitch them back into the original FOF halo.

Earlier theoretical papers on merger rates either relied on a small sample of main haloes to estimate the overall redshift evolution over a limited range of halo masses, or were primarily concerned with the mergers of *galaxies* or *subhaloes*. For halo mergers, for example, Governato et al. (1999) studied $z < 1$ major mergers of galaxy-sized haloes in an open CDM and a tilted $\Omega_m = 1$ CDM model using *N*-body simulations in a 100-Mpc box and 144^3 particles. Gottlöber, Klypin & Kravtsov (2001) used a sample of ∼4000 haloes to study the environmental dependence of the redshift evolution of the major merger rate at $z < 2$ in ΛCDM. Berrier et al. (2006) studied major mergers of subhaloes in *N*-body simulations in a 171-Mpc box with 512^3 particles and the connection to the observed close pair counts of galaxies. For galaxy merger rates, Murali et al. (2002) and Maller et al. (2006) are based on up to ∼500 galaxies formed in SPH simulations in ∼50-Mpc boxes with up to 144^3 gas particles, while Guo & White (2008) used the semi-analytical galaxy catalogue of De Lucia et al. (2006) based on the Millennium Simulation.

This paper is organized as follows. Section 2 describes the dark matter haloes in the Millennium Simulation (Section 2.1) and how we construct the merger trees (Section 2.2). We then discuss the issue of halo fragmentation and the two methods ('snipping' and 'stitching') used to treat these events in Section 2.3. The notation used in this paper is summarized in Section 2.4.

Section 3 describes how mergers are counted (Section 3.1) and presents four (related) statistical measures of the merger rate (Section 3.2). The relation between these merger rate statistics and the analytical merger rate based on the EPS model is derived in Section 3.3.

Our main results on the merger rates computed from the Millennium Simulation are presented in Section 4. We first discuss the $z \approx 0$ results and quantify the merger rates as a function of the descendant halo mass and the progenitor mass ratios using merger trees constructed from the stitching method (Section 4.1). The evolution of the merger rates with redshifts up to $z \sim 6$ is discussed in Section 4.2. We find a simple universal form for the merger rates and present an analytic fitting form that provides a good approximation (at the 10–20 per cent level) over a wide range of parameters (Section 4.3).

Section 5 compares the stitching and snipping merger rates (Section 5.1) and presents the key results from a number of tests that we have carried out to assess the robustness of our results. Among the tests are: time convergence and the dependence of the merger rates on the redshift spacing Δz between the Millennium outputs used to construct the merger tree (Section 5.2); how the counting of binary versus multiple progenitor mergers affects the merger rates (Section 5.3); mass non-conservation arising from 'diffuse' accretion in the form of unresolved haloes during mergers (Section 5.4) and how the definition of halo masses and the treatment of fragmentation events affect the resulting halo mass function (Section 5.5).

In Section 6, we discuss two theoretical frameworks that can be used to model halo mergers: EPS and coagulation. A direct comparison of our merger rates and the EPS predictions for the Millennium ΛCDM model shows significant differences over a large range of parameter space (Section 6.1). Section 6.2 discusses Smoluchowski's coagulation equation and the connection between our merger rates and the coagulation merger kernel.

The appendix compares a third merger tree (besides snipping and stitching) constructed from the Millennium catalogue by the Durham group (Helly et al. 2003; Bower et al. 2006; Harker et al. 2006). Two additional criteria are imposed on the subhaloes in this algorithm to reduce spurious linkings of FOF haloes. We find these criteria to result in reductions in both the major merger rates and the halo mass function.

The cosmology used throughout this paper is identical to that used in the Millennium Simulation: a Λ CDM model with $\Omega_m = 0.25, \Omega_b = 0.045, \Omega_\Lambda = 0.75, h = 0.73$, an initial power-law index $n = 1$ and $\sigma_8 = 0.9$ (Springel et al. 2005). Masses and lengths are quoted in units of $M_\odot$ and Mpc without the Hubble parameter *h*.

2 HALOES AND MERGER TREES IN THE MILLENNIUM SIMULATION

2.1 Dark matter haloes

The Millennium Simulation provides the largest data base to date for studying the merger histories of dark matter haloes in the ΛCDM cosmology. The simulation uses 2160^3 particles with a particle mass of $1.2 \times 10^9\,M_\odot$ in a 685-Mpc box and traces the evolution of roughly 2×10^7 dark matter haloes from redshift $z = 127$ to 0 (Springel et al. 2005).

The haloes in the simulation are identified by grouping the simulation particles using the standard FOF algorithm (Davis et al. 1985) with a linking length of $b = 0.2$. Each FOF halo (henceforth referred to as *FOF* or *halo*) is then broken into constituent subhaloes by the SUBFIND algorithm, which identifies dark matter substructure as locally overdense regions within each FOF and removes any remaining gravitationally unbound particles (Springel et al. 2001). The result is a list of disjoint subhaloes typically dominated by one large background host subhalo and a number of smaller satellite subhaloes.

Each subhalo in the catalogue is assigned a mass given by the number of particles bound to the subhalo; only subhaloes with more than 20 simulation particles are included in the data base. Each FOF halo is then given two definitions of mass: $M_{\rm FOF}$, which counts the number of particles associated with the FOF group, and M_{200}, which assumes the halo is spherical and computes the virial mass within the radius at which the average interior density of the halo is 200 times the mean density of the universe. $M_{\rm FOF}$ includes background particles that are unbound by the SUBFIND algorithm so it is generally larger than the sum of the subhalo masses. In this paper we mainly use $M_{\rm FOF}$ as it is found to be the more robust mass definition in our merger study. We discuss M_{200} and a number of mass conservation issues in Section 5.4.

2.2 Merger tree construction

Merger trees of dark matter haloes in the Millennium data base are constructed by connecting *subhaloes* (*not the FOF haloes*) across 64 snapshot outputs: a subhalo at a given output is taken to be the *descendant*[1] of a *progenitor* subhalo at a prior output (i.e. higher redshift) if it contains the largest number of bound particles in the progenitor subhalo. This procedure results in a merger tree in which each progenitor subhalo has a single descendant subhalo, even though in general, the particles in the progenitor do not necessarily all end up in the same descendant subhalo.

It is worth noting that merger trees in N-body simulations are typically constructed based on the FOF haloes and not on the subhaloes. The standard way of assigning the progenitor and descendant FOF haloes in those studies, however, is the same as the procedure applied to the *subhaloes* in Millennium discussed above; that is, the descendant halo is the halo that inherits the most number of bound particles of the progenitor. As will be elaborated on below, we call this the 'snipping' method.

The focus of this paper is on the merger history of the *FOF* haloes rather than the subhaloes, so we must process the subhalo merger tree available from the public data base to construct a consistent merger tree for the FOF haloes. We consider an FOF halo A to be a descendant of an earlier FOF halo B if B contains a subhalo whose descendant subhalo is in A. Progenitor FOF haloes are said to have merged when all their descendant subhaloes are identified with one descendant FOF. We illustrate this process in Fig. 1 with an actual merger tree taken from the Millennium data base. The upper left-hand corner, for example, shows three FOF haloes at $z = 0.24$ with masses $8.5 \times 10^{12}, 4 \times 10^{11}$ and $3.8 \times 10^{10}\,{\rm M}_{\odot}$ merging into a single FOF halo at the next Millennium output ($z = 0.21$). The largest FOF halo $z = 0.24$ has seven subhaloes (white circles) in addition to the host (sub)halo, while each of the two smaller FOF haloes has only one host (sub)halo. For clarity, the ancestral links between subhaloes are suppressed in Fig. 1.

2.3 Halo fragmentation

Even though each subhalo in the Millennium tree, by construction, is identified with a single descendant subhalo (see last subsection), the resulting FOF tree can have fragmentation events in which an FOF halo is split into two (or more) descendant FOF haloes. The red circles in Fig. 1 at $z =$ (0.12:0.09) and (0.06:0.04) illustrate two such events: the subhaloes of the progenitor FOF halo end up in different descendant FOF haloes. It is important to emphasize that this fragmentation issue is not unique to the use of subhaloes in the Millennium Simulation, but rather occurs in general in any merger tree construction where groups of particles at two different redshifts must be connected. This is because particles in a progenitor halo rarely end up in exactly one descendant halo; a decision must therefore be made to select a unique descendant. There is not a unique way to do this, and we explore below two methods that we name *snipping* and *stitching* to handle these fragmentation events.

[1] It is common practice in the literature to call the descendant halo the *parent* halo even though the parent is formed later and, hence, is younger than the progenitor. We avoid this confusing notation throughout.

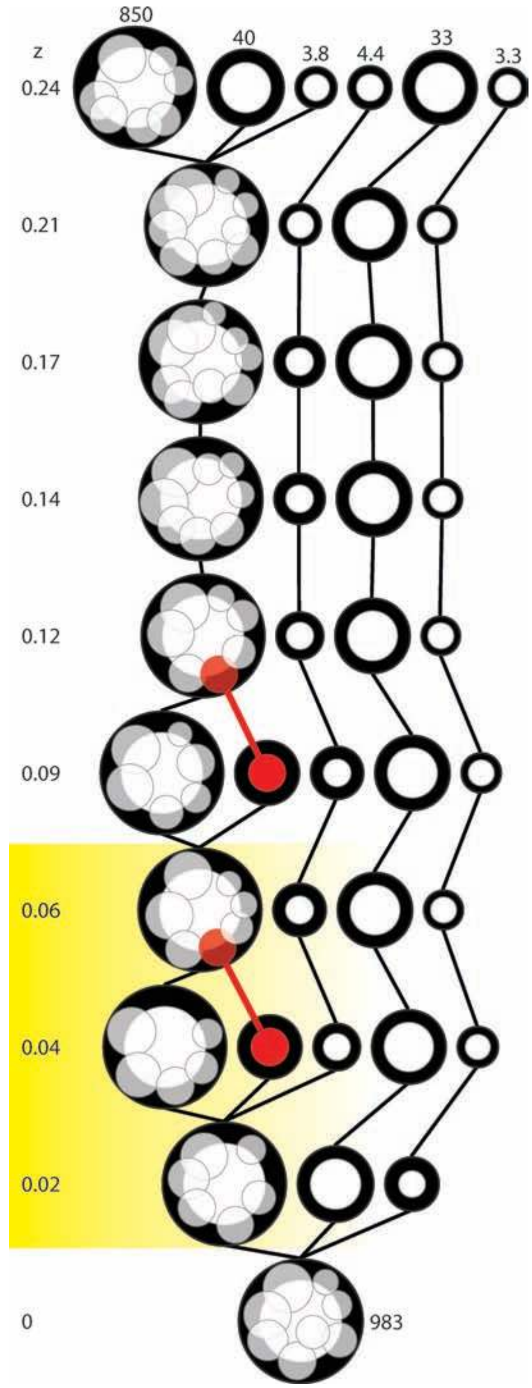

Figure 1. Example of a typical FOF merger tree extracted from the Millennium data base. Black circles denote FOF haloes; white circles within black circles denote subhaloes. The radius of each circle is proportional to the log of the mass of the object; the black circles are further scaled up by a factor of 1.5 for clarity. (The locations of the white circles within their parent FOF haloes are drawn randomly.) Red circles denote fragmenting subhaloes. The highlighted (yellow) fragmentation event is studied in Fig. 2. The numbers above the haloes at $z = 0.24$ and to the right-hand side of the final descendant FOF at $z = 0$ correspond to the FOF masses (in units of $10^{10}\,{\rm M}_{\odot}$).

Fig. 2 illustrates these two methods for the fragmentation event shown in the highlighted (yellow) region of Fig. 1. The snipping method is commonly used in the literature (e.g. Sheth & Tormen 2002), presumably for its simplicity. Fragmentation events are

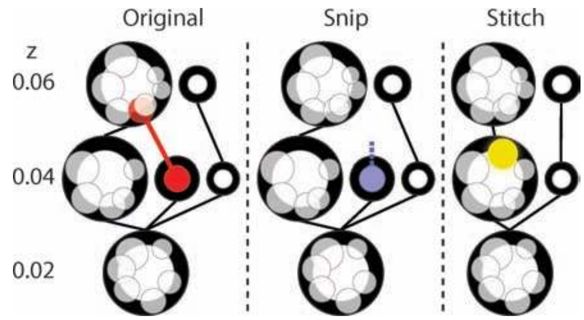

Figure 2. Left-hand panel: A close-up of the highlighted (yellow region) fragmentation event in Fig. 1. The middle and right-hand panels illustrate how the snipping and stitching methods handle fragmentation in order to assign a unique descendant halo. The blue circle (centre panel) shows the snipped orphan subhalo, and the yellow circle (right-hand panel) shows how that subhalo is stitched. The black, white and red circles are the same as in Fig. 1.

removed by 'snipping' the link between the smaller descendant halo and its progenitor FOF halo, as shown in the middle panel of Fig. 2. The fragmenting progenitor FOF halo then has only one descendant FOF halo. We note that this method can result in a number of progenitorless *orphan* FOF haloes (e.g. the blue subhalo in Fig. 2).

In this paper we investigate an alternative method that we name 'stitchingí. This method is motivated by our observation that about half of the fragmented haloes in the Millennium Simulation remerge within the following two to three outputs (see below). The two fragmentation events in Fig. 1 both belong to this category: the fragmented haloes at $z = 0.09$ and 0.04 (red circles) are seen to have remerged by the following output time ($z = 0.06$ and 0.02). This behaviour is not too surprising because merging haloes oscillate in and out of their respective virial radii before dynamical friction brings them into virial equilibrium (typically on time-scales of a few Gyr; see e.g. Boylan-Kolchin, Ma & Quataert 2008). During this merging phase, the FOF halo finder can repeatedly disassociate and associate the progenitor haloes, leading to spurious fragmentation and remerger events and inflating the merger rate. This behaviour needs to be taken into account before a robust merger rate can be obtained.

We therefore do not count remerging fragments as merger events in the 'stitching' method. Specifically, we group the fragmented haloes into two categories: those that remerge within three outputs after fragmentation occurs, and those that do not. The fragmented haloes that remerge are stitched into a single FOF descendant (e.g. the yellow subhalo in the right-hand panel of Fig. 2); those that do not remerge are snipped and become orphan haloes. Often the fragment subhaloes have become members of a new FOF group that is otherwise unrelated to the original FOF. In such instances they are removed from that group and stitched into the main FOF descendant.[2] A further test of the dependence of our results on the choice of three outputs is described in Section 5.1.

As can be seen in Fig. 2, the snipping method will yield a higher merger rate than stitching due to the remerger events. We quantify the relative importance of these events in Fig. 3, where the ratio of fragmentation events to merger events is seen to peak at 40 per cent for major fragmentation events (defined to be fragmentations where the fragment subhalo carries 10 per cent or more of the halo mass) at low z and falls off at high z where Δz is large. For the fragmentation events occurring at a given redshift z_f in Fig. 3 (filled circles), the drop of each curve with decreasing z tracks how many of them have remerged by that redshift. As noted above, we find that about half of the fragmented haloes remerge within two to three outputs (corresponding to a fixed $\Delta z/(1 + z)$ as the outputs are log-spaced). Given a fragmentation-to-merger ratio of 40 per cent, and a remerger rate of 50 per cent, the remerging fragments can impact the merger rate measurements inflating them at the ∼20 per cent level.

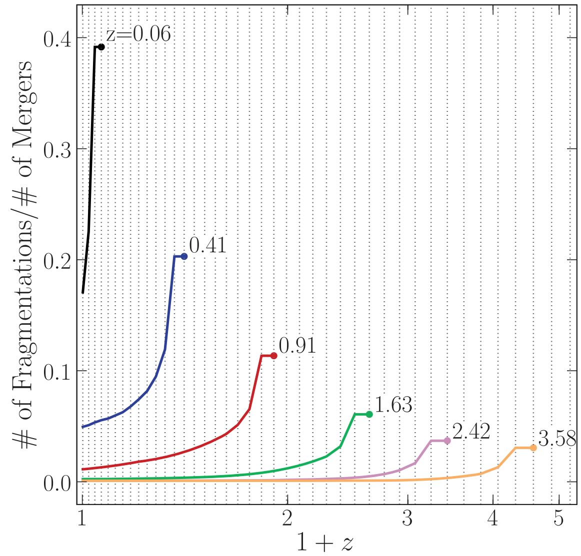

Figure 3. Distribution of the ratio of fragmentation to merger events as a function of redshift. The dotted vertical lines correspond to the redshifts of the Millennium outputs. We choose six redshifts (labelled) for illustrative purposes and plot the ratio of the number of fragmentations to the number of mergers (filled circles) at each redshift. A mass ratio cut-off is applied: both the fragments and mergers must have mass ratios exceeding 10 per cent. The line emanating from each circle then traces the evolution of the number of fragmentation events (the number of mergers being held fixed), which drops as subhalo fragments remerge with their original FOF halo. We note that about half of the subhalo fragments remerge within two to three simulation outputs. Finally, the six filled circles decrease with increasing redshift, reaching ∼40 per cent at $z = 0$ but dropping to ∼5 per cent at high z – this is primarily due to the increasing Δz between Millennium outputs.

Moreover, we find that this effect is more severe for fragmentations where the mass of the fragment is small relative to the mass of the original parent halo (we call these minor fragmentations). If we consider fragmentations in which the subhalo fragments carry between 1 and 10 per cent of the original FOF mass, the fragmentation-to-merger ratio at $z = 0$ ($z = 1.6$) jumps to 57 per cent (13.3 per cent) versus 39 per cent (6 per cent) for major fragmentations. For very minor fragmentations (subhalo fragments that carry less than 1 per cent of the total mass) the fragmentation-to-merger ratios are 85 and 28 per cent at $z = 0$ and 1.6, respectively. Thus we anticipate that fragmentation events will more severely pollute the minor-merger regime of the merger rate statistics.

2.4 Notation

We apply both the stitching and snipping methods and produce FOF merger trees from the 46 Millennium outputs that span $z = 0$ and

[2] There is, however, one exceptional case: if a subhalo fragment becomes the largest subhalo of an FOF, all subhaloes in that FOF are stitched into the fragment's original FOF.

6.2. From these trees we connect different outputs and generate a catalogue of descendant FOF haloes at the low-z (z_D) output and their associated progenitor FOF haloes at the high-z (z_P) output. We refer to this as the $z_P : z_D$ catalogue and produce a number of catalogues for a variety of output spacings. The redshift spacing is denoted by $\Delta z = z_P - z_D$. The Millennium outputs are logarithmically distributed, providing fine Δz down to 0.02 (corresponding to ∼260 Myr) near $z = 0$ and larger Δz at high redshifts, e.g. $\Delta z \approx 0.1$ at $z \approx 1$ and ≈0.5 at $z \approx 6$. Specifically, the lowest 10 redshift outputs are at 0.0, 0.02, 0.04, 0.06, 0.09, 0.12, 0.14, 0.17, 0.21 and 0.24.

For a given FOF-descendant halo in a $z_P : z_D$ catalogue, we use M_0 to denote its $M_{\rm FOF}$ mass, $N_{\rm p}$ to denote the number of progenitor haloes, and M_i with $i \in (1, 2, \ldots, N_{\rm p})$ to denote the rank-ordered $M_{\rm FOF}$ mass of the progenitors, i.e. $M_1 \geqslant M_2 \geqslant \cdots M_{N_{\rm p}}$. We impose a minimum mass cut-off of $M_0 \geqslant 2 \times 10^{12}\,{\rm M}_\odot$ on the descendant FOF halo and a cut-off of $M_i \geqslant 4.8 \times 10^{10}\,{\rm M}_\odot$ on the progenitors, which corresponds to 40 particles and is twice the minimum halo mass in the Millennium data base.

For certain results reported below, we make use of three large mass bins: $2 \times 10^{12} \leqslant M_0 < 3 \times 10^{13}\,{\rm M}_\odot$, $3 \times 10^{13} \leqslant M_0 < 10^{14}\,{\rm M}_\odot$ and $10^{14}\,{\rm M}_\odot \leqslant M_0$, referred to as the *galaxy-scale*, *group-scale* and *cluster-scale* bins, respectively.

3 MERGER STATISTICS AND CONNECTION TO EPS

3.1 Counting many-to-one mergers

Despite the fine time spacing between Millennium's outputs, a non-negligible number of the descendant FOF haloes have more than two progenitors listed in the merger tree (i.e. $N_{\rm p} > 2$). For completeness, we list in Table 1 the actual number of merger events in the Millennium Simulation available to us after we construct the FOF merger trees. Statistics at five representative redshifts are shown: $z \approx 0$, 0.5, 1, 2 and 3. At each z, we list separately the number of FOF haloes that have $N_{\rm p} = 1, 2$ and >2 progenitor haloes, for three separate descendant mass bins. As expected for hierarchical cosmological models, the halo numbers drop with increasing z and increasing M_0.

Fig. 4 shows the distribution of the number of progenitors, $f(N_{\rm p})$, for the $z = 0.06{:}0$ merger tree for the same three mass bins. Only the stitching method is shown; the snipping method has a similar distribution. We find that (62, 22, 16) per cent of the haloes have $N_{\rm p} = (1, 2, > 2)$ identifiable progenitors at $z = 0.06$; more than half of the FOF haloes at $z = 0$ therefore have only one progenitor at $z = 0.06$ and did not experience a merger during this redshift interval. When separated into different descendant mass bins, the

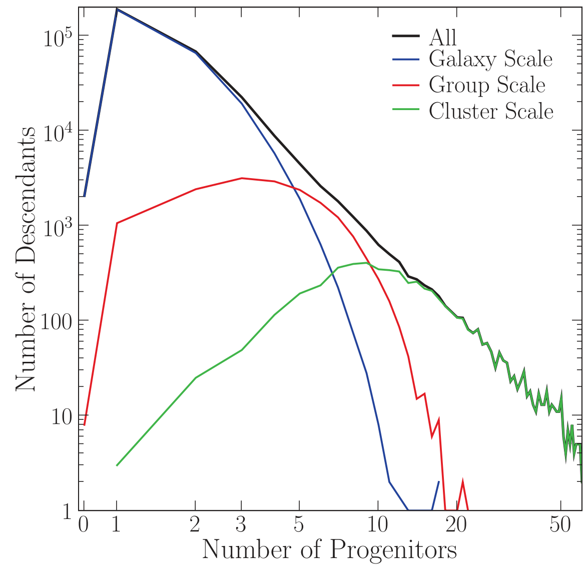

Figure 4. Distribution of the number of progenitors, $N_{\rm p}$, for the $z = 0.06{:}0$ merger tree. There are ∼300 000 descendant FOF haloes at redshift 0 (black) with $M_0 \geqslant 2 \times 10^{12}\,{\rm M}_\odot$. Of these ∼280 000 have $2 \times 10^{12} \leqslant M_0 < 3 \times 10^{13}\,{\rm M}_\odot$ (galaxy-scale; dark blue), ∼16 000 have $3 \times 10^{13} \leqslant M_0 < 10^{14}\,{\rm M}_\odot$ (group scale; red) and ∼5400 have $M_0 \geqslant 10^{14}\,{\rm M}_\odot$ (cluster-scale; green).

peak of $f(N_{\rm p})$ moves to higher $N_{\rm p}$ for more massive haloes. For a fixed (z_P, z_D), clusters therefore tend to have more progenitors, and unlike galaxy-mass haloes, very few of the cluster haloes are single-progenitor events (i.e. $N_{\rm p} = 1$).

For completeness, we include *all* the progenitors (above our minimum mass cut-off of 40 particles) in our merger rate statistics. Since we have no information about the order in which the multiple progenitors merge with one another, we assume that each progenitor halo M_i with $i \geqslant 2$ merges with M_1, the most massive progenitor, at some stage between the two outputs. Thus a descendant halo with $N_{\rm p}$ progenitors is assumed to be the result of a sequence of $(N_{\rm p} - 1)$ binary merger events, where each merger event is assigned a mass ratio

$$\xi \equiv \frac{M_i}{M_1}, \quad i = 2, \ldots, N_{\rm p} \tag{1}$$

which by construction satisfies $\xi \leqslant 1$. This assumption ignores the possibility that two smaller progenitor FOF haloes merge together before merging with the most massive progenitor. Section 5.3 describes how we have tested the validity of this assumption and found negligible effects as long as a sufficiently small Δz is used.

Table 1. The number of merger events in the Millennium Simulation that we use to determine the merger rates. Merger trees at five representative redshifts are shown: $z \approx 0$, 0.5, 1, 2 and 3. At each z, we list the number of FOF haloes that have a single progenitor halo ($N_{\rm p} = 1$, i.e. no mergers), two progenitors ($N_{\rm p} = 2$, i.e. binary mergers) and multiple progenitors ($N_{\rm p} > 2$), for three separate descendant mass bins: $2 \times 10^{12} \leqslant M_0 < 3 \times 10^{13}\,{\rm M}_\odot$ (galaxy), $3 \times 10^{13} \leqslant M_0 < 10^{14}\,{\rm M}_\odot$ (group) and $M_0 \geqslant 10^{14}\,{\rm M}_\odot$ (cluster). Only progenitor haloes with mass $>4.8 \times 10^{10}\,{\rm M}_\odot$ (40 simulation particles) are counted.

$z_P : z_D$		Galaxy scale			Group scale			Cluster scale	
	$N_{\rm p} = 1$	$N_{\rm p} = 2$	$N_{\rm p} > 2$	$N_{\rm p} = 1$	$N_{\rm p} = 2$	$N_{\rm p} > 2$	$N_{\rm p} = 1$	$N_{\rm p} = 2$	$N_{\rm p} > 2$
0.06 : 0	188 400	65 711	27 939	1063	2418	13 256	3	25	5356
0.56 : 0.51	189 351	61 718	22 031	1212	2468	9374	6	18	3014
1.08 : 0.99	145 779	68 467	35 426	325	878	7630	0	2	1308
2.07 : 1.91	76 298	52 525	39 097	31	77	2225	0	0	129
3.06 : 2.83	30 641	26 675	25 072	0	4	343	0	0	4

3.2 Definitions of merger rates

In this subsection we define four related quantities that will be used to measure the merger rates of dark matter haloes. Merger rates can be measured in either per Gyr or per unit redshift; the two sets of quantities are related by a factor of d t/d z. We will present most of our results in units of per redshift since, as we will show below, the merger rates have a particularly simple form in those units.

As a starting point, we consider the symmetric merger rate

$$B_{MM'}(M, M', z_{\rm P}{:}z_{\rm D})\,{\rm d}M\,{\rm d}M', \tag{2}$$

which measures the mean merger rate (i.e. the number of mergers per unit redshift) per unit volume between progenitor FOF haloes in the mass range (M, M + dM) and (M', M' + dM'). We compute this quantity using merger trees constructed between the progenitor output redshift $z_{\rm P}$ and the descendant output redshift $z_{\rm D}$. Note that $B_{MM'}(M, M')$ has units of [number of mergers $\times\,(\Delta z)^{-1}\,{\rm Mpc}^{-3}\,{\rm M}_\odot^{-2}$] and generally depends on both $z_{\rm P}$ and $z_{\rm D}$.

Instead of the individual progenitor masses M and M', it is often useful to express merger rates as a function of the *descendant* FOF mass and the mass ratio of the progenitors. We do this by transforming $B_{MM'}(M, M')\,{\rm d}M\,{\rm d}M'$ to

$$B(M_0, \xi, z_{\rm P}{:}z_{\rm D})\,{\rm d}M_0\,{\rm d}\xi, \tag{3}$$

which measures the mean merger rate (per volume) for descendant FOF haloes in the mass range (M_0, M_0 + dM_0) at redshift $z_{\rm D}$ that have progenitor FOF haloes at $z_{\rm P}$ with mass ratio in the range of (ξ, ξ + dξ), where $\xi = M_i/M_1$, $i \geqslant 2$ as discussed in Section 3.1. The quantity $B(M_0, \xi)$ therefore has units of [number of mergers $\times\,\Delta z^{-1}\,{\rm Mpc}^{-3}\,{\rm M}_\odot^{-1}\,{\rm d}\xi^{-1}$]. In the mass-conserving binary limit of $M_0 = M + M'$ and $\xi = M'/M$ (where $M' < M$), $B_{MM'}$ and B in equations (2) and (3) are related by a simple transformation. In practice, the relation between the two quantities is complicated by multiple mergers and imperfect merger mass conservation.

Since the halo abundance in a ΛCDM universe decreases with increasing halo mass, many more haloes contribute to the merger rates in equations (2) and (3) in the lower mass bins of M, M' or M_0. It is useful to normalize out this effect and calculate the mean merger rates *per halo*. To do this, we divide out the number density of the descendant FOF haloes from the merger rate B and define

$$\frac{B}{n} \equiv \frac{B(M_0, \xi, z_{\rm P}{:}z_{\rm D})}{n(M_0, z_{\rm D})}, \tag{4}$$

which measures the mean number of mergers *per halo* per unit redshift for a descendant halo of mass M_0 with progenitor mass ratio ξ; the units are [number of mergers/number of descendants $\times(\Delta z)^{-1}$ $({\rm d}\xi)^{-1}$], which is dimensionless. The mass function $n(M_0, z)\,{\rm d}M_0$ gives the number density of the descendant FOF haloes with mass in the range of (M_0, M_0 + dM_0).

The differential merger rates defined above can be integrated over ξ and M_0 to give the mean merger rate over a certain range of merger mass ratios for haloes in a given mass range. Explicitly, the mean rate of mergers for descendant haloes in mass range $M_0 \in [m, M]$ with progenitor mass ratios in the range $\xi \in (x, X)$,

$$\frac{{\rm d}\bar{N}_{\rm merge}}{{\rm d}z}\,([m, M], [x, X], z_{\rm P}{:}z_{\rm D}), \tag{5}$$

is simply an integral over $B(M_0, \xi, z_{\rm P}{:}z_{\rm D})$:

$$\frac{{\rm d}\bar{N}_{\rm merge}}{{\rm d}z} \equiv \frac{1}{N}\int_m^M\!\!\int_x^X B(M_0, \xi, z_{\rm P}{:}z_{\rm D})\,{\rm d}\xi\,{\rm d}M_0, \tag{6}$$

where

$$N \equiv \int_m^M n(M_0, z_{\rm D})\,{\rm d}M_0 \tag{7}$$

is the total number of descendant haloes in the relevant mass range. For sufficiently small $(M - m)$, ${\rm d}\bar{N}_{\rm merge}/{\rm d}z$ is simply related to the merger rate per halo, B/n, by

$$\frac{{\rm d}\bar{N}_{\rm merge}}{{\rm d}z} \sim \int_x^X \frac{B}{n}\,{\rm d}\xi. \tag{8}$$

3.3 Connection to EPS

The merger rates determined from the Millennium Simulation can be compared to the analytic predictions of the EPS formalism (Bond et al. 1991; Lacey & Cole 1993). To relate our per halo merger rate B/n to EPS, we begin with equation (2.18) of Lacey & Cole (1993) for

$$\frac{{\rm d}^2 p}{{\rm d}\ln\Delta M^{\rm LC}\,{\rm d}t}\left(M_1^{\rm LC} \rightarrow M_2^{\rm LC}\middle|t\right), \tag{9}$$

the probability that a halo of mass $M_1^{\rm LC}$ will merge with another halo of mass $\Delta M^{\rm LC} = M_2^{\rm LC} - M_1^{\rm LC}$ in time interval d t. Their notation (which we denote with superscripts 'LC') is related to ours by $M_2^{\rm LC} \rightarrow {\rm M}_0$, with $M_1^{\rm LC}$ and $\Delta M^{\rm LC}$ mapped to our progenitor masses M_1 and M_2. As we will see below, the order is ambiguous due to an inconsistency in the EPS model that stems from the assumption of binary mergers. To relate ${\rm d}^2p/{\rm d}\ln\Delta M^{\rm LC}/{\rm d}t$ to B/n, we first multiply it by $n(M_1^{\rm LC})$, then convert the variables to (M_0, ξ) (see below), and finally divide by $n(M_0)$.

Before presenting the actual equation relating the two rates, we note two caveats. First, in order to compute an analytical merger rate from EPS we must assume that mergers are binary and perfectly mass conserving, i.e. $M_0 = M_1 + M_2$ in our notation. Neither assumption is strictly true in numerical simulations, e.g. Table 1 and Fig. 4 show the distributions of the progenitor multiplicity $N_{\rm p}$. We defer to Section 5 for a detailed discussion of the tests that we have performed to quantify the binary nature and the degree of merger mass conservation in the Millennium Simulation.

Second, the EPS rate in equation (9) is not symmetric in the progenitor masses $M_1^{\rm LC}$ and $\Delta M^{\rm LC}$, in contrast to our merger rate $B_{MM'}$ in equation (2), which is constructed to be symmetric in the progenitor masses M and M'. We will therefore get different EPS rates depending on if $M_1^{\rm LC}$ is chosen to be the bigger or smaller progenitor. We will examine both options below: (A) $\xi = \Delta M^{\rm LC}/M_1^{\rm LC} \leqslant 1$ and (B) $\xi = M_1^{\rm LC}/\Delta M^{\rm LC} \leqslant 1$.

With these caveats in mind, we find that the per halo merger rate B/n corresponds to the following expression in the EPS model:

$$\frac{B(M_0, \xi, z)}{n(M_0, z)} \leftrightarrow \sqrt{\frac{2}{\pi}}\,\frac{{\rm d}\delta_{\rm c}}{{\rm d}z}\,\frac{1}{\sigma(M')}\left|\frac{{\rm d}\ln\sigma}{{\rm d}\ln M}\right|_{M'}\left[1 - \frac{\sigma^2(M_0)}{\sigma^2(M')}\right]^{-3/2}, \tag{10}$$

where M' can be the smaller progenitor, i.e. $M' = M_2 = M_0\xi/(1 + \xi)$ (option A), or the larger progenitor, i.e. $M' = M_1 = M_0/(1 + \xi)$ (option B). The variable $\sigma^2(M)$ is the variance of the linear density field smoothed with a window function containing mass M, and $\delta_{\rm c}(z) \propto 1/D(z)$ is the standard density threshold, with $D(z)$ being the linear growth factor. Note that the exponential dependence at the high-mass end of the halo mass function has cancelled out on the right-hand side of equation (10). Also note that both sides of equation (10) are for merger rates *per redshift* and not per time.

We present our results for the merger rates determined from the Millennium Simulation in the next section and compare them to the two EPS predictions in Section 6.1.

4 RESULTS

Throughout this section, we report our results from the Millennium merger tree where the fragmented haloes are handled with the stitching method. We find the merger rates given by the snipping method to agree with the stitching results to within 25 per cent. Details of the comparison are discussed in Section 5.1.

4.1 Merger rates at $z \approx 0$

Fig. 5 is a contour plot of the symmetric merger rate in equation (2), $B_{MM'}(M, M', z_P{:}z_D)$, calculated using the stitching merger tree constructed from the $z = 0.06{:}0$ Millennium outputs. Darker (bluer) regions denote higher merger rates, which are concentrated in the lower (M, M') corner because there are more low-mass haloes. Minor mergers (off-diagonal) are more common than major mergers (along the diagonal). The lower left-hand corner is blank due to our lower cut-off on the descendant FOF mass ($\sim$1000 particles; $M_0 \gtrsim 2 \times 10^{12}\,\mathrm{M}_\odot$). The noisy nature of the upper right-hand corner is due to limited merger statistics at $\sim 10^{15}\,\mathrm{M}_\odot$.

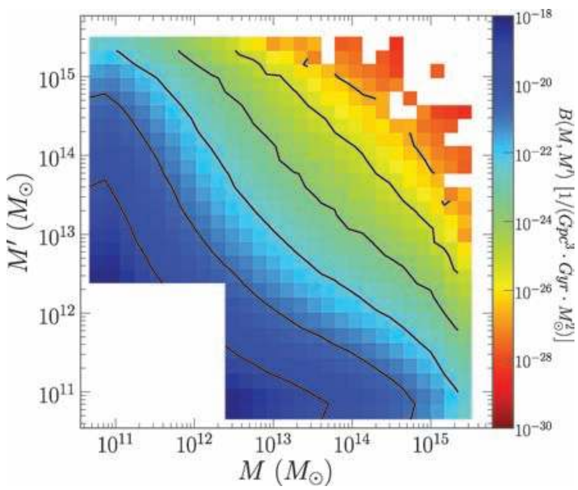

Figure 5. Symmetric merger rate $B_{MM'}$ of equation (2) as a function of progenitor masses M and M' computed from the $z = 0.06{:}0$ Millennium merger tree. The merger rates decrease from blue to red; the overlaid black lines are contours of constant merger rates.

As we discussed in Section 3.2, instead of progenitor masses M and M', it is often more illuminating to study merger rates as a function of the descendant FOF halo mass M_0 and the mass ratio ξ of the progenitors. This is shown in Fig. 6 for the same data set as in Fig. 5. The left-hand panel plots the merger rate $B(M_0, \xi, 0.06{:}0)$ of equation (3) against the progenitor mass ratio ξ for fixed bins of descendant FOF mass M_0. We observe that the merger rate $B(M_0, \xi)$ is a power law in the progenitor mass ratio ξ when $\xi \lesssim 0.1$ and shows an upturn in the major merger regime. The power-law index is close to -2 and is nearly independent of the descendant mass M_0. More precise values are given in the fitting form in equation (12) and Table 2 below.

The main quantity we study in this paper is the mean merger rate *per descendant halo*, B/n, of equation (4), shown in the right-hand panel of Fig. 6. The rising amplitude of B with decreasing M_0 is remarkably largely removed when B/n is plotted: the curves in the left-hand panel for different M_0 mass bins collapse on to nearly a single curve in the right-hand panel. This behaviour indicates that the merger rate per halo is nearly independent of the descendant halo mass. This weak mass dependence is further illustrated in Fig. 7 and is also reported in Guo & White (2008). As we will quantify in Section 4.3 below, the dependence on M_0 is approximately $\propto M_0^{0.08}$.

Our lower cut-off of 40 particles for the progenitor FOF halo mass implies a lower cut-off in the mass ratio of $\xi \geqslant 4.8 \times 10^{10}\,\mathrm{M}_\odot/M_0$. This resolution cut-off is seen in the left-hand panel of Fig. 6, where we have sufficient halo statistics to measure the merger rates for the higher mass haloes (lower curves) down to very minor mergers, e.g. $\xi < 10^{-3}$ for $M_0 > 5 \times 10^{13}\,\mathrm{M}_\odot$; whereas the dynamic range is smaller for galaxy-size haloes, e.g. $\xi > 0.01$ for $M_0 < 5 \times 10^{12}\,\mathrm{M}_\odot$.

The present-day merger rates shown in Fig. 6 are all obtained from the $z = 0.06{:}0$ merger tree. The low-redshift outputs available from the Millennium data base in fact have a smaller spacing of

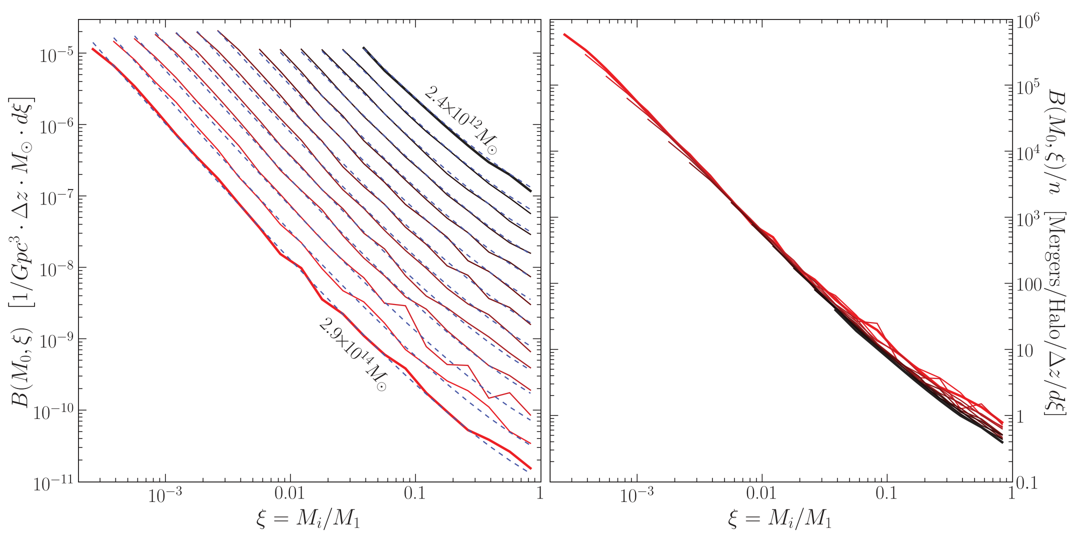

Figure 6. Left-hand panel: Mean merger rate $B(M_0, \xi)$ of equation (3) for the $z = 0.06{:}0$ merger tree as a function of the mass ratio of the progenitors, ξ, for bins of fixed descendant halo mass M_0 (colour coded from black to red for increasing M_0). The overlaid dashed blue lines are from our fitting formula in equation (12). Note that the presence of a fixed minimum mass resolution ($4.7 \times 10^{10}\,\mathrm{M}_\odot$) corresponds to a minimum mass ratio ξ that decreases as M_0 increases. Right-hand panel: Mean merger rate *per halo*, $B(M_0, \xi)/n(M_0)$, of equation (4) for the same tree. Dividing out the halo number density $n(M_0)$ brings the curves on the left-hand panel to nearly a single curve, indicating B/n has very weak dependence on M_0.

Table 2. Best-fitting parameters for equation (12).

Method	A	$\tilde{\xi}$	α	β	γ	η	χ^2_ν
Snip	0.0101	0.017	0.089	−2.17	0.316	0.325	1.86
Stitch	0.0289	0.098	0.083	−2.01	0.409	0.371	1.05

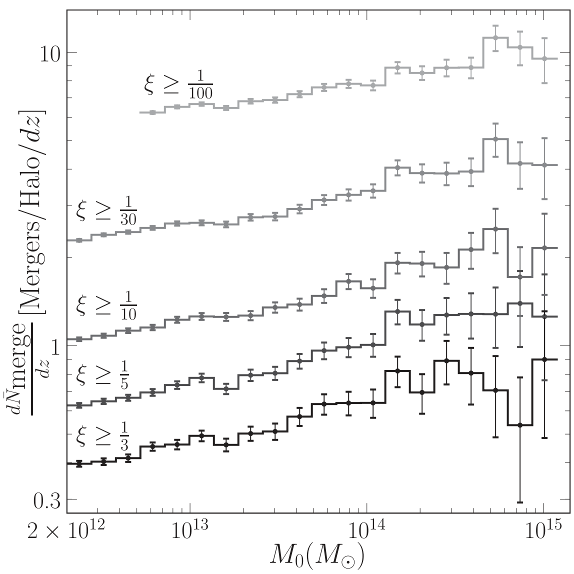

Figure 7. Mean merger rate per halo (per unit z), $\mathrm{d}N_\mathrm{m}/\mathrm{d}z$, as a function of descendant mass, M_0, for various ranges of the progenitor mass ratio ξ. The upper curves include increasingly more minor mergers. The $z = 0.06$:0 merger tree is used. Note the weak mass dependence over three decades of mass. The error bars are computed assuming Poisson counting statistics in both the number of mergers and the number of haloes.

$\Delta z \sim 0.02$. We use the 0.06:0 merger tree to avoid any edge effects arising from our stitching criterion that only subhalo fragments that remerge within three outputs are stitched together (see Section 2.3). In practice, this precaution is not critical and we find little difference between the 0.06:0 and 0.02:0 results.

4.2 Merger rates at higher redshift

Figs 6 and 7 summarize our results for the $z = 0$ merger rates. At higher redshifts, the Millennium data base provides sufficient halo statistics for us to measure merger rates up to $z \sim 6$. The results are shown in Fig. 8, where we plot the merger rate per unit time (upper panel), $\mathrm{d}\bar{N}_\mathrm{merge}/\mathrm{d}t$, and per unit redshift (lower panel), $\mathrm{d}\bar{N}_\mathrm{merge}/\mathrm{d}z$, as a function of redshift for three ranges of descendant masses (galaxy, group, cluster) and four ranges of progenitor mass ratios ($\xi \geqslant 1/3$, 1/10, 1/30 and 1/100). Errors are computed assuming Poisson statistics for the number of mergers and haloes. We have suppressed merger rates with poor merger statistics (and, therefore large error bars) to keep the plots legible.

The mean merger rate per Gyr (upper panel) is seen to increase at higher z. We have fit power laws to each M_0 and ξ range (dotted curves) of the form

$$\frac{\mathrm{d}\bar{N}_\mathrm{merge}}{\mathrm{d}t} \propto (1+z)^{n_\mathrm{m}} \tag{11}$$

and find $n_\mathrm{m} \sim 2$–2.3 for the ranges of M_0 and ξ shown. The Millennium merger rates are seen to flatten out slightly at low z and deviate from a power law when the cosmological constant starts to dominate the energy density of the universe.

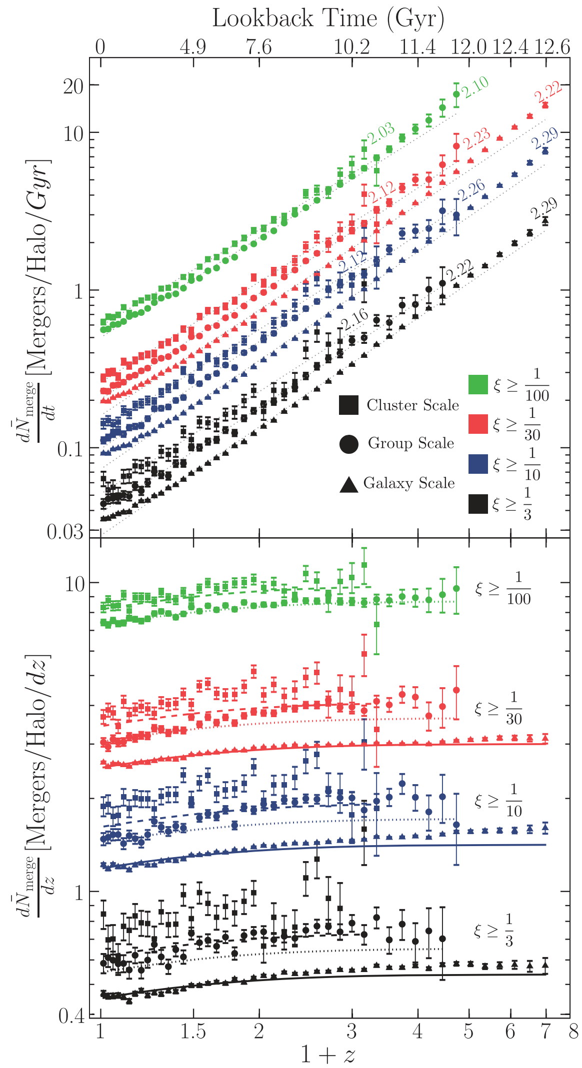

Figure 8. Upper panel: Mean merger rate per halo (per Gyr), $\mathrm{d}N_\mathrm{m}/\mathrm{d}t$, as a function of redshift for three bins of descendant mass M_0 and four ranges of progenitor mass ratio ξ from the Millennium Simulation (using the stitching tree). The overlaid lines plot the best-fitting power laws, $(1+z)^{n_\mathrm{m}}$, with n_m ranging from 2.03 to 2.29 (labelled). Note that power laws are reasonable fits at $z \gtrsim 0.3$ but underpredict the Millennium rates at lower z. Lower panel: Same as the upper panel but showing the merger rate $\mathrm{d}\,N_\mathrm{m}/\mathrm{d}\,z$ per unit z instead of per Gyr. The dotted grey lines here show our fitting formula in equation (12), which is tuned to provide close fits at low z. In both panels, the error bars are computed assuming Poisson counting statistics in both the number of mergers and the number of haloes, and the curve for galaxy-scale haloes (triangles) with $\xi \geqslant 1/100$ (green) is suppressed because such minor mergers fall below the simulation resolution limit.

A large number of merger rate statistics can be easily read off of Fig. 8. For example, at around $z = 2$ ($z = 4$) every FOF halo on average experiences ~2–4 (10) minor mergers ($\xi \lesssim 1/30$) per Gyr, and about 10–20 per cent (70–90 per cent) of FOF haloes experience a major merger ($\xi \gtrsim 1/3$) every Gyr.

Unlike the rising $\mathrm{d}\bar{N}_\mathrm{merge}/\mathrm{d}t$, the merger rate *per unit redshift*, $\mathrm{d}\bar{N}_\mathrm{merge}/\mathrm{d}z$, shows a remarkably weak dependence on z in Fig. 8 (lower panel), increasing only slightly between $z = 0$ and 1 and staying nearly constant for $z \gtrsim 1$ for all ranges of M_0 and ξ shown.

The overlaid curves are computed by integrating over the fitting form for *B*/*n* to be discussed below (Section 4.3).

At $z > 0$, Fig. 8 shows that the dependence of $\mathrm{d}\bar{N}_{\rm merge}/\mathrm{d}z$ on progenitor ratio ξ and descendant mass M_0 is similar to the $z = 0$ merger rates shown in Fig. 6: minor mergers occur more frequently than major mergers, and the dependence on M_0 is weak, with galaxy-scale haloes (triangles) on average experiencing fewer mergers (per halo) than cluster-size haloes (squares).

4.3 A universal fitting form

We now propose a fitting form that can be used to approximate the halo merger rates in the Millennium Simulation discussed in the last two subsections to an accuracy of 10–20 per cent. The key feature we will use to simplify the fit is the nearly universal form of the merger rate (per halo) $B(M_0, \xi)/n$ shown in the right-hand panel of Fig. 6, and the weak redshift dependence shown in the bottom panel of Fig. 8. We find that the following functional form works well:

$$\frac{B(M_0,\xi,z)}{n(M_0,z)} = A\left(\frac{M_0}{\tilde{M}}\right)^{\alpha}\xi^{\beta}\exp\left[\left(\frac{\xi}{\tilde{\xi}}\right)^{\gamma}\right]\left(\frac{\mathrm{d}\delta_{\rm c}}{\mathrm{d}z}\right)^{\eta}, \qquad (12)$$

where $\tilde{M} = 1.2 \times 10^{12}\,\mathrm{M_{\odot}}$ is a constant and $\delta_{\rm c}(z) \propto 1/D(z)$ is the standard density threshold normalized to $\delta_{\rm c} = 1.686$ at $z = 0$, with $D(z)$ being the linear growth factor. Note that equation (12) is separable with respect to the three major variables M_0, ξ and z.

The form of the redshift dependence in equation (12) is chosen so that $\eta = 1$ corresponds to the EPS prediction in equation (10). In addition, this form has weak z dependence at $z \gtrsim 1$ since the growth factor approaches that of the Einstein–de Sitter model, $\delta_{\rm c}(z) = 1.68(1 + z)$, and $\mathrm{d}\delta_{\rm c}/\mathrm{d}z$ approaches a constant. This behaviour matches the weak redshift dependence seen in the Millennium merger rate (bottom panel of Fig. 8).

To determine the parameters in equation (12), we fit simultaneously to all redshifts $z < 1$, mass ratios $\xi > 10^{-3}$, and masses $10^{12} \lesssim M_0 \lesssim 10^{14}\,\mathrm{M_{\odot}}$. The *B*/*n* data points are weighted using their Poisson distributed errors. The resulting fits are plotted as dotted curves in Figs 6 and 8, and the fitting parameters are given in Table 2, along with the overall reduced χ^2_ν obtained by fitting to all redshifts $z < 1$ simultaneously. In addition to computing a global χ^2_ν we also compute a local $\chi^2_\nu(z)$ at each redshift and find relatively good convergence across the $z < 1$ redshift range: $\chi^2_\nu(z)$ remains below 1.5 for stitching and below 2–3 for snipping.

We note that the fitting form of equation (12) does not appear symmetric in the progenitor masses M_1 and M_2 because by construction, $\xi \equiv M_2/M_1 < 1$. However, for any pair of progenitors, we identify M_1 with the more massive and M_2 with the less massive progenitor and then compute $\xi = M_2/M_1 < 1$. This procedure yields the same ξ and therefore the same *B*/*n* regardless of the order of the input progenitors, in contrast to the EPS model discussed in Section 3.3, which is intrinsically asymmetric in M_1 and M_2.

5 TESTS

5.1 Snipping versus stitching trees

Fig. 9 shows the ratio of the $z = 0$ per halo merger rate *B*/*n* from the snipping and stitching methods. Overall, the merger rates given by the two methods differ by no more than 25 per cent over two to three orders of magnitude in both the progenitor mass ratio ξ and the descendant mass M_0. Within this difference, however, Fig. 9 and Table 2 show that the snipping method systematically

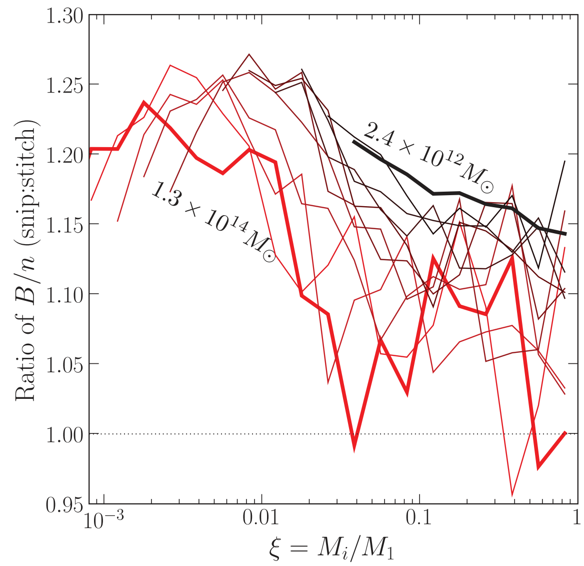

Figure 9. The ratio of the snipping and stitching *B*/*n* as a function of mass ratio ξ computed using the 0.06:0 catalogue for a variety of mass bins in the range $2.4 \times 10^{12}\,\mathrm{M_{\odot}}$ (black) $\leqslant M \leqslant 1.3 \times 10^{14}$ (red). We find differences at the 25 per cent level at low ξ with the snipping method consistently predicting a higher merger rate at all ξ. We attribute this to the population of remerging orphan haloes.

yields a higher merger rate and a steeper slope in the ξ dependence than the stitching method. These additional merger events come from the orphaned subhaloes that are first snipped and subsequently remerge (see Fig. 2). Moreover, as discussed in Section 2.3, the fragmentation-to-merger ratio is higher for more minor subhalo fragments (those with low fragment-to-FOF mass ratios). There are therefore more remerging orphan haloes with lower ξ, leading to the larger difference between snipping and stitching at low ξ seen in Fig. 9.

A remaining issue is our choice of the stitching criterion: as described in Section 2.3, we stitch only FOF fragments that are observed to remerge within the next three outputs. This choice is motivated by the fact that about half of the halo fragments at a given output will have remerged within three outputs (see Fig. 3), and that such a small Δz criterion will allow us to effectively compute instantaneous merger rates. We have tested this criterion further by implementing a more aggressive stitching algorithm that stitches *all* fragments, regardless of whether they eventually remerge. We call this ∞-stitching. This algorithm represents the opposite limit to the snipping method and may err on the side of *underestimating* the merger rates since it would stitch together close encounter fly-by events that do not result in actual mergers within a Hubble time. We find the amplitude of *B*/*n* from ∞-stitching to be lower than that from the 3-stitching by up to ∼25 per cent, similar in magnitude but opposite in sign to the difference between snipping versus 3-stitching shown in Fig. 9. The fitting form in equation (12) works well for ∞-stitching, where the best-fitting parameters are $A = 0.0344$, $\tilde{\xi} = 0.125$, $\alpha = 0.118$, $\beta = -1.921$, $\gamma = 0.399$ and $\eta = 0.853$. This algorithm shows excellent convergence properties (see Section 5.2) and excellent mass conservation properties (Section 5.4) but alters the FOF mass function by a few per cent.

Since the snipping algorithm tends to inflate the merger rate and the ∞-stitching algorithm tends to underestimate it, we believe the 3-stitching used in all the results in Section 4 should be a fairly robust scheme.

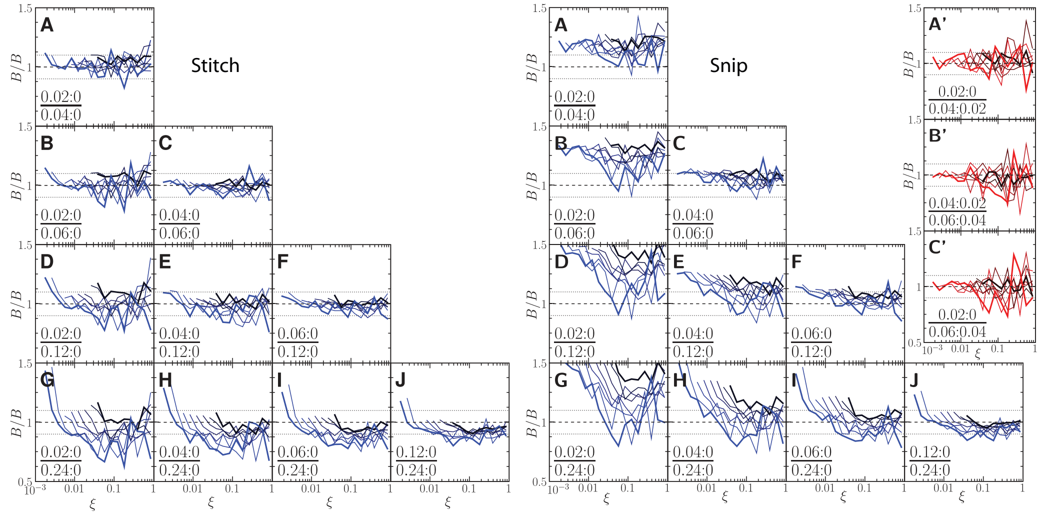

Figure 10. Δz convergence matrix (stitching, left-hand side; snipping, right-hand side). Each subplot is the ratio of B/n for two different catalogues (labelled). The dashed lines denote equality and the dotted lines are the 10 per cent deviation levels. The ratios are presented for a variety of mass bins with the high-mass bin highlighted in thick blue (or red) and the low-mass bin highlighted in thick black.

5.2 Convergence with respect to Δz

We have performed a number of tests to quantify the dependence of our merger rate results on the choice of Δz between the Millennium outputs used to construct the merger trees. It is not a priori clear which value of Δz is optimal: small Δz can result in poor merger statistics since most haloes would not have had time to merge; whereas large Δz does not have the time resolution to track individual merger events accurately and also runs the risk of smearing out real redshift-dependent effects. The optimal Δz may also vary with redshift.

Our first test focuses on $z \approx 0$ mergers and quantifies how B/n varies with the Δz used to construct the trees. Fig. 10 shows the ratios of B/n for five pairs of progenitor and descendant redshifts: $(z_{\rm P}, z_{\rm D}) =$ (0.02:0), (0.04:0), (0.06:0), (0.12:0) and (0.24:0), corresponding to a time interval of $\Delta t = 0.26$, 0.54, 0.83, 1.44 and 2.77 Gyr, respectively. For the stitching method (left-hand panel), there is excellent convergence for $\Delta z \lesssim 0.12$ (panels A–F), where the ratios of B/n are centred around 1 per cent and rarely deviate beyond the 10 level (dotted line). For $\Delta z = 0.24$ (panels G–J), the ratios start to drop below unity. This is consistent with the slowly rising merger rates with increasing z shown in Fig. 8. Thus, the stitching method yields merger trees with robust Δz convergence properties near $z = 0$, and we have chosen $\Delta z = 0.06$ to compute the merger rates in earlier sections.

The snipping method (right-hand panel) shows inferior Δz convergence. The B/n computed with smaller Δz consistently show higher merger rates than those computed with larger Δz. Moving up the left-hand column (panels G, D, B, A), we observe only some degree of convergence. Better convergence is seen along the main diagonal (panels A, C, F, J) in order of increasing Δz. In particular, panels C and F show excellent convergence properties (to the 10 per cent level) centred around (0.06:0). To emphasize that the problem is with Δz and not with a particular output (say, any possible edge effects at $z = 0$ or 0.02), we show in panels A′, B′, C′ the ratios of B/n computed using three merger trees with the same $\Delta z = 0.02$ but centred at progressively higher z:$z =$ (0.02:0), (0.04:0.02) and (0.06:0.04). The agreement is excellent, in striking contrast to panel B. Based on these tests, we have chosen to use $\Delta z = 0.06$ for the snipping method.

We believe that the snipping method has inferior Δz convergence properties because of the remerging orphan subhaloes (see Section 2.3 and Fig. 2). These fragmentation events are sewn together in the stitching scheme and therefore do not contribute to the merger rates. In the snipping scheme, however, the snipped events provide a fresh supply of haloes, many of which remerge in the next few outputs. This effect artificially boosts the merger rate across small Δz.

Our second Δz convergence test is performed at all redshifts. We test three types of spacings: (1) *adjacent* spacing uses adjacent catalogues, e.g. at low z, it uses (0.02:0), (0.04:0.02), (0.06:0.04); (2) *skip 1* spacing skips an output, e.g. (0.04:0), (0.06:0.02), ... and (3) *skip 2* spacing skips two outputs, e.g. (0.06:0), (0.09:0.02) and so on. Fig. 11 shows $\mathrm{d}\bar{N}_{\rm merge}/\mathrm{d}z$ computed using these three Δz for galaxy-mass haloes. We again see excellent Δz convergence for the stitching method at $z \lesssim 1.5$ (left-hand panel) and worse Δz

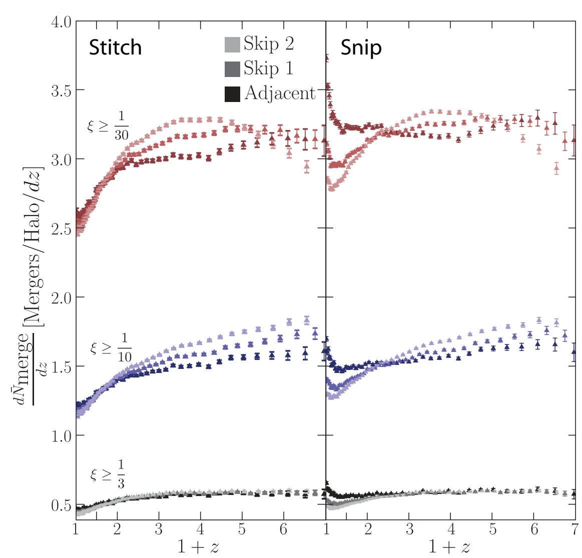

Figure 11. Merger rate $\mathrm{d}\bar{N}_{\rm merge}/\mathrm{d}z$ computed using three types of redshift spacings: *adjacent*, *skip 1* and *skip 2* (see text). For clarity, only the galaxy-mass haloes are shown; the group and cluster haloes behave similarly.

convergence for the snipping method (right-hand panel). The latter follows the behaviour seen in Fig. 10, with adjacent spacing ($\Delta z = 0.02$ at $z = 0$) overpredicting the merger rate.

At higher redshifts ($z \gtrsim 1.5$), Fig. 11 shows that the merger rates in the minor merger regime differ by up to $\sim$15 per cent depending on which of the three types of spacing is used. This difference is not likely to be due to the fragmentation events since as Fig. 3 shows, the ratio of fragmentation to merger events is 40 per cent near $z = 0$ but drops to $\lesssim$10 per cent for $z \gtrsim 3$. Rather, we believe that the inferior Δz convergence at high z is due to the increasing Δz between Millennium outputs (e.g. the smallest Δz is $\sim$0.5 at $z \sim 6$ versus $\Delta z = 0.02$ at $z \approx 0$) and the inaccuracy of the multiple counting ordering assumption for large Δz (see Section 5.3). At high z, we therefore advocate using the finest output spacings available in the Millennium data base, noting the good time convergence for major mergers ($\xi \gtrsim 1/3$) but $\sim$15 per cent variations in the minor merger rates.

5.3 Multiple versus binary counting

As discussed in Section 3.1, for descendant FOF haloes with more than two progenitor haloes, we include all progenitors when we calculate the merger rates for completeness. Since mergers are often assumed to be binary events, we have tested the difference between our multiple counting results and those obtained by counting only the two most massive progenitors of a given descendant halo. Fig. 12 compares the merger rates (per halo), $B(M_0, \xi)/n$, for these two counting methods (dashed: multiple; solid: binary) as a function of the progenitor mass ratio ξ for three descendant masses M_0 (increasing from left- to right-hand side). In each panel, results from four merger trees using increasing Δz of 0.02, 0.06, 0.12 and 0.24 are shown.

The most notable trend in Fig. 12 is that the multiple counting method gives similar merger rates regardless of Δz, indicating good time convergence in the results (as we have discussed in detail in Section 5.2). The binary counting rates, on the other hand, deviate increasingly from the multiple rates when larger Δz are used because the binary assumption becomes less valid for larger Δz. As a function of ξ, the binary and multiple merger rates match well in the major merger regime but deviate significantly for small ξ. This occurs because binary counting counts only the two most massive progenitors and ignores the additional (typically low-mass) progenitors. It therefore closely approximates the major-merger rates but underestimates the minor-merger regime of the multiple counting result.

Fig. 12 suggests that for a given minimum mass resolution (i.e. a minimum ξ), there is a corresponding Δz for which the binary counting method is a good approximation. For example, for $6 \times 10^{13}\ \mathrm{M}_\odot$ haloes (centre panel), the binary and multiple merger rates are similar down to $\xi \approx 0.05$ for $\Delta z \sim 0.12$, and down to $\xi \approx 0.005$ when Δz is decreased to 0.02. Thus the multiple counting *B*/*n* can be thought of as the small-Δz limit of the binary *B*/*n*.

Another test we have performed is on the ordering of mergers assumed in the multiple counting method described in Section 3.1. There, we assumed that the less massive progenitors $M_2, M_3, \ldots$ each merged with the most massive progenitor M_1 and not with one another. This assumption is motivated by the fact that satellite haloes in *N*-body cosmological simulations are typically seen to accrete on to a much more massive host halo as a minor merger event instead of merging with another satellite halo. We have quantified the validity of this assumption by taking large Δz in the Millennium outputs, applying this ordering, and checking against the actual merging order among the progenitors when finer Δz is used. (Of course, we cannot do this for the minimum $\Delta z = 0.02$ available in the data base.) The fraction of misordering naturally rises with increasing Δz due to the degraded time resolution, but for $\Delta z \lesssim 0.06$, we find the fraction of progenitors to have merged with a progenitor other than M_1 to be $\lesssim$10 per cent. Most of the mergers among multiple progenitors, therefore, do occur between the most massive progenitor and a less massive progenitor, as we have assumed.

5.4 Mass conservation and 'diffuse' accretion

Thus far we have analysed mergers in terms of the progenitor halo mass M_i and the descendant halo mass M_0. Mergers are, however, messy events, and the sum of M_i does not necessarily equal M_0. To quantify this effect, we define a 'diffuse' component, ΔM, for a given descendant halo:

$$M_0 = \sum_{i=1}^{N_\mathrm{p}} M_i + \Delta M, \tag{13}$$

where ΔM is diffuse in the sense that it is not resolved as distinct haloes in the simulation. A non-zero ΔM can be due to physical processes such as tidal stripping and diffuse mass accretion that cause a net loss or gain in halo mass after a merger event. In simulations,

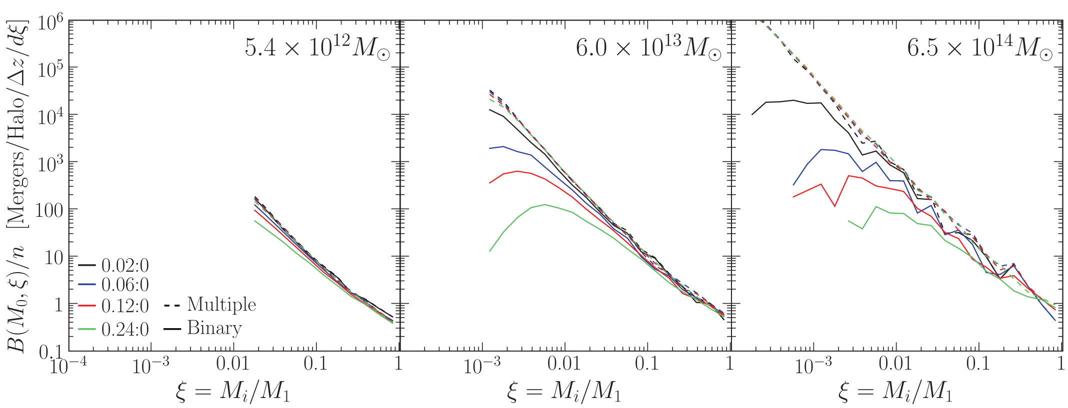

Figure 12. Comparisons of merger rate per halo, *B*/*n*, computed via multiple counting (dashed lines) and binary counting (solid lines) for four merger trees with increasing Δz. Three descendant mass bins are shown (from left- to right-hand side). We find the multiple counting rate to be in excellent agreement regardless of Δz of the tree, whereas the binary counting *B*/*n* curves fall off from the observed power-law behaviour towards lower ξ.

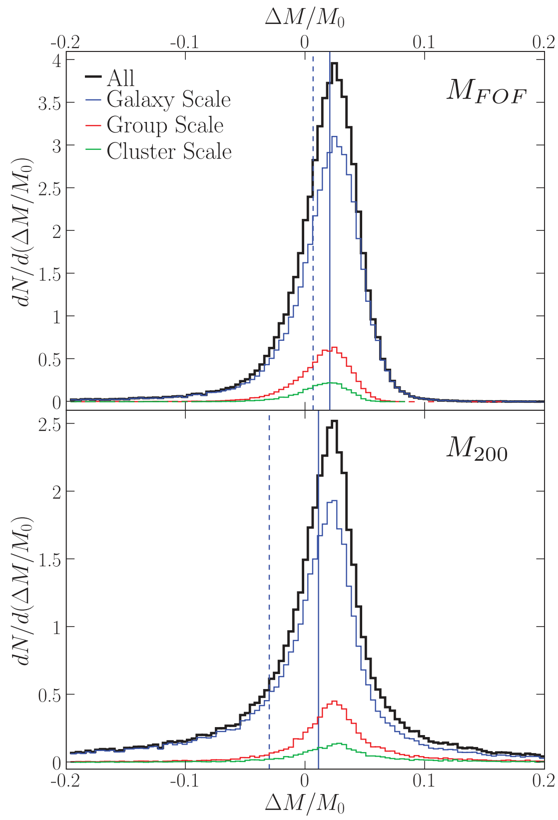

Figure 13. Distributions of $\Delta M/M_0$ from the $z = 0.06$:0 Millennium merger tree (using stitching; snipping is nearly identical) computed using the $M_{\rm FOF}$ mass (top panel) versus M_{200} virial mass (bottom panel). The solid vertical line is the median of the distribution for the galaxy-mass bin and the dashed line is the mean. We note a longer negative ΔM tail for the M_{200} tree when compared to the $M_{\rm FOF}$ tree. Note, however, that the peaks of the two distributions are in good agreement ($\Delta M/M_0 \sim 2.5$ per cent).

additional numerical factors also contribute to ΔM due to different algorithms used in, for example, defining halo mass (FOF versus spherical overdensity). ΔM therefore does not necessarily have to be positive in every merger event.

Fig. 13 shows the distribution of $\Delta M/M_0$ for the $z = (0.06$:0) Millennium merger tree. Only haloes that have experienced mergers between these two redshifts (i.e. those with more than one progenitor) are plotted. The snipping tree (not shown) gives a very similar distribution as the stitching tree shown here. For comparison, we have computed $\Delta M/M_0$ using the two different halo mass definitions $M_{\rm FOF}$ and M_{200}. The distribution shows a prominent peak at $\Delta M/M_0 \sim 2.5$ per cent for both $M_{\rm FOF}$ and M_{200}, indicating that in the majority of the merger events between $z = 0.06$ and 0.0, ∼97.5 per cent of the mass of the descendant halo comes from resolvable progenitor haloes.

Even though the two distributions in Fig. 13 have similar peaks, the M_{200} mass definition produces longer $\Delta M/M_0$ tails than the $M_{\rm FOF}$ mass definition, and the mean of the distribution (dotted vertical line) is negative for M_{200}. We believe this is because mass definitions based on the assumption of spherical symmetry (as M_{200} does) have difficulties assigning accurate mass to non-spherical FOF haloes and tend to underestimate the halo mass (see e.g. White 2001). $M_{\rm FOF}$, on the other hand, can account for all the mass in a given FOF object that is identified as 'merged' by the FOF halo finder well before virialization. As discussed in Section 2.1, we have been using the $M_{\rm FOF}$ mass thus far.

Our main results on merger rates in Section 4 were determined for numerically resolved dark matter haloes; they are therefore not affected directly by the fact that ΔM is generally non-zero for merger events. We find, however, that $\Delta M/M_0$ increases with Δz, and this diffuse accretion component makes an important contribution to the mass growth of a halo over its lifetime. We will explore the growth of haloes in further detail in subsequent papers.

5.5 Halo mass function

The mass function of dark matter haloes in principle depends on both the definition of halo mass and the algorithm used to treat fragmentation events. As illustrated in Fig. 2, the snipping method by construction preserves the original FOF mass function, while the stitching scheme modifies it slightly as it rearranges fragmented subhaloes between FOFs. We find that the impact on the mass function is negligible (less than ∼0.25 per cent) so will use the stitching result below.

Fig. 14 shows the ratio of the Millennium halo mass function to the fits by Jenkins et al. (2001) for $M_{\rm FOF}$ at four redshifts: $z \approx 0$, 1, 2 and 3. The fit of Jenkins et al. is accurate to better than 10 per cent for low redshift ($z \lesssim 1$), but it underestimates the Millennium halo abundance by $\gtrsim$25 per cent at the high-mass end for $z > 1$. This discrepancy is present but not obvious on the log–log plot in fig. 2 of Springel et al. (2005). Lukic et al. (2007) also noted this difference. Since the stitching and snipping mass functions are virtually identical, this appears to be a discrepancy between the Millennium FOF catalogue and the fit of Jenkins et al. (2001).

6 THEORETICAL MODELS FOR MERGER RATES

6.1 Extended Press–Schechter model

In Section 3.3 we discussed how our merger rates are related to the conditional probabilities in the EPS model and obtained equation (10), where there are two choices for the definition of the

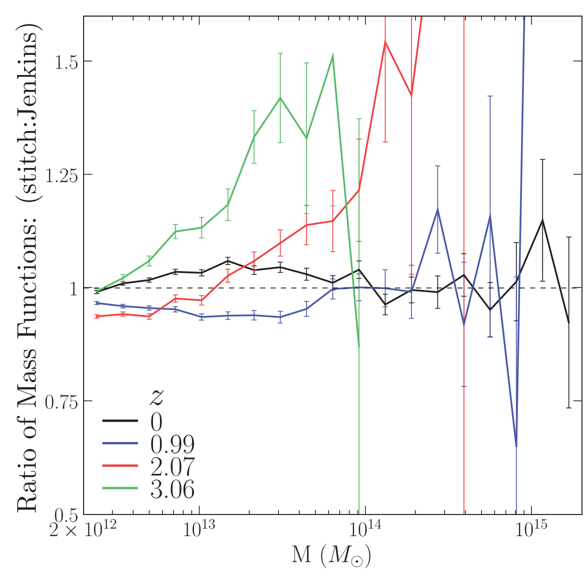

Figure 14. Ratios of the Millennium halo mass function (computed from the stitching trees) to the fit of Jenkins et al. (2001) using the $M_{\rm FOF}$ mass. The results for the snipping trees are virtually identical. We note a significant deviation of ∼25 per cent at $z \sim 3$ between the Jenkins fit and the Millennium mass function.

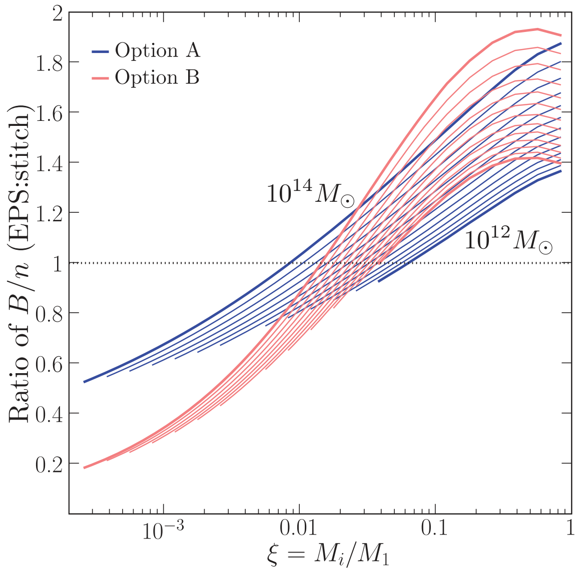

Figure 15. Comparison between our Millennium merger rate (from the fit) and the two predictions of the EPS model. The ratio of *B*/*n* from EPS to Millennium is plotted. Blue and red label the two options in assigning progenitor masses in the EPS model (see text); within each colour, the set of curves from bottom to top denotes increasing M_0 bins, from $\sim 10^{12}$ to $\sim 3 \times 10^{14}$ M$_\odot$. The EPS model is seen to overpredict the major merger rate by up to a factor of ∼2 and underpredicts the minor merger rate by up to a factor of ∼5.

progenitor mass M' since the EPS model is not symmetric in the two progenitor masses. In Fig. 15 we show the ratio of the EPS prediction to our Millennium *B*/*n* at $z = 0$, where we have computed the EPS rates given by the right-hand side of equation (10) using the same cosmological parameters as for the Millennium Simulation. We compute the variance of the smoothed linear density field, $\sigma^2(M)$, in the ΛCDM cosmology using the power spectrum fit provided in Eisenstein & Hu (1999).

Fig. 15 shows that EPS *underpredicts* the $z = 0$ rate for minor mergers by up to a factor of ∼5, and *overpredicts* the rate at $\xi \gtrsim 0.05$, indicating that the dependence of the EPS merger rate on ξ is shallower than our $B/n \sim \xi^\beta$, where the best-fitting β is −2.17 and −2.01 for the snipping and stitching methods, respectively (see Table 2). In terms of the descendant mass M_0, the dependence of the EPS rate is too steep compared to our *B*/*n*, leading to the spread in each bundle of curves in Fig. 15. The two choices of M' in EPS are seen to lead to different predictions. Assigning M' to be the smaller progenitor (option A) results in a somewhat smaller discrepancy than option B.

Fig. 15 compares the rates at $z = 0$. At higher redshifts, we find the Millennium merger rate to evolve as $\propto (\mathrm{d}\delta_c/\mathrm{d}z)^\eta$, where $\eta \approx 0.37$ (see equation 12 and Table 2) and is shallower than the EPS prediction of $\eta = 1$ in equation (10). Since the functional forms of both our fit for *B*/*n* and the EPS expression are separable with respect to M_0, ξ and z, the $z = 0$ curves in Fig. 15 will maintain the same shape at higher z, but the amplitude of the ratio will increase. For instance, the ratio shown in Fig. 15 will be increased by a factor of 1.26, 1.32, 1.34 and 1.35 at $z = 1$, 2, 4 and 6, respectively. The discrepancy between the Millennium results and the EPS predictions is therefore even worse at higher *z*.

Given that the Press–Schechter mass function is known not to match the halo abundances in simulations very closely, it is not particularly surprising that the EPS merger rates in Fig. 15 do not match the Millennium results closely. The substantial discrepancy, however, does highlight the limitation of the EPS model and provides the motivation to build more accurate merger rates based on improved PS mass functions. We address this issue in separate papers (Zhang, Ma & Fakhouri 2008), in which we investigate a moving density-barrier algorithm to generate merger trees that produces a better match to simulation results than the constant barrier of the PS model.

6.2 Halo coagulation

The merging of dark matter haloes is, in principle, a coagulation process. Coagulation is often modelled by the Smoluchowski coagulation equation (Smoluchowski 1916), which governs the time evolution of the mass function $n(M, t)$ of the objects of interest with a coagulation kernel. In the absence of fragmentations, the time change of n is given by

$$\frac{\mathrm{d}n(M)}{\mathrm{d}t} = \frac{1}{2}\int_0^M A(M', M-M')n(M')n(M-M')\mathrm{d}M' - \int_0^\infty A(M, M')n(M')n(M)\,\mathrm{d}M', \quad (14)$$

where the first term on the right-hand side is a source term due to mergers of two smaller haloes of mass M' and $M - M'$, while the second term is a sink term due to haloes in the mass bin of interest merging with another halo of mass M', forming a halo of higher mass $M + M'$. When applied to hierarchical structure formation, $A(M, M')$, the symmetric coagulation kernel (in units of volume/time), tracks the probability for a halo of mass M to merge with a halo of mass M'. Our merger rate per halo, *B*/*n*, can be simply related to *A* by

$$A(M, M') \leftrightarrow \frac{B(M, M')}{n(M)n(M')}. \quad (15)$$

We note, however, that the coagulation equation in the form of equation (14) is valid only for mass-conserving binary mergers. As seen throughout this paper, these assumptions are not strictly true in numerical simulations, and modifications are required to account for the issues that have been discussed thus far, such as net mass gain or loss (i.e. $\Delta M \neq 0$), multiple merger events, and halo fragmentation. While the relative errors may be small when integrated over a small time interval, repeated application of equation (14) using equation (15) may not yield robust results.

Assuming that $n(M)$ is the Press–Schechter mass function, Benson, Kamionkowski & Hassani (2005) have developed numerical techniques to construct the coagulation kernel for self-similar cosmological models with initial power-law power spectrum $P(k) \propto k^n$. Their technique is underconstrained and does not yield a unique expression for $A(M, M')$. In order to pick out a particular solution, a regularization condition was applied to force $A(M, M')$ to vary smoothly. We have transformed the coordinates of their fits to $A(M, M')$ to compare their results with our merger rate *B*/*n*. Fig. 16 shows the ratio of their fitting formula to the Millennium ΛCDM merger rate for spectral indices $n = -1$ and -2 as a function of progenitor mass ratio ξ for various descendant halo mass bins. The difference can be up to a factor of several.

7 CONCLUSIONS AND DISCUSSIONS

In this paper we have computed the merger rates of dark matter FOF haloes as a function of descendant halo mass M_0, progenitor mass ratio ξ, and redshift z using the merger trees that we constructed

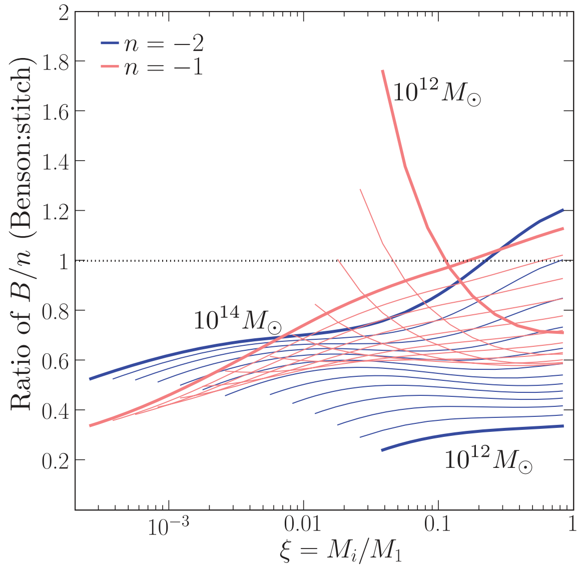

Figure 16. Same as Fig. 15, only now comparing the merger rates from Benson et al. (2005) to the Millennium *B/n* for initial power spectrum index $n = -2$ (blue) and $n = -1$ (red).

from the halo catalogue of the Millennium Simulation. Our main results are presented in Figs 6–8, which show very simple and nearly separable dependence on M_0, ξ and z. The mean merger rate per descendant FOF halo, *B/n*, is seen to depend very weakly on the halo mass M_0 (Fig. 6, right-hand panel and Fig. 7). As a function of redshift z, the per halo merger rate in units of per Gyr increases as $(1+z)^\alpha$, where $\alpha \sim 2$ to 2.3 (top panel of Fig. 8), but when expressed in units of per redshift, the merger rate depends very weakly on z (bottom panel of Fig. 8). Regardless of M_0 and z, the dependence of *B/n* on the progenitor mass ratio, $\xi = M_i/M_1$, is a power law to a good approximation in the minor merger regime ($\xi \lesssim 0.1$) and shows an upturn in the major merger regime (Fig. 6). These simple behaviours have allowed us to propose a universal fitting formula in equation (12) that is valid for $10^{12} \leqslant M_0 \lesssim 10^{15}\,\mathrm{M}_\odot$, $\xi \gtrsim 10^{-3}$ and up to $z \sim 6$.

Throughout the paper we have emphasized and quantified the effects on the merger rates due to events in which a progenitor halo fragments into multiple descendant haloes. We have shown that the method commonly used to remove these fragmented haloes in merger trees – the snipping method – has relatively poor Δz convergence (Figs 10 and 11). Our alternative approach – the stitching method – performs well with regards to this issue without drastically modifying the mass conservation properties or the mass function of the Millennium FOF catalogue (Figs 13 and 14).

We have computed the two predictions for merger rates from the analytical EPS model for the same ΛCDM model used in the Millennium Simulation. At $z = 0$, we find the EPS major merger rates to be too high by 50–100 per cent (depending on halo mass) and the minor merger rates to be too low by up to a factor of 2–5 (Fig. 15). The discrepancy increases at higher z.

The coagulation equation offers an alternative theoretical framework for modelling the mergers of dark matter haloes. We have discussed how our merger rate is related to the coagulation merger kernel in theory. In practice, however, we find that mergers in simulations are not always mass-conserving binary events, as assumed in the standard coagulation form given by equation (14). Equation (14) will therefore have to be modified before it can be used to model mergers in simulations.

Gottlöber et al. (2001) studied the rate of major mergers (defined to be $\xi \geqslant 1/3$ in our notation) in *N*-body simulations and found a steeper power-law dependence of $\propto (1+z)^3$ (at $z \lesssim 2$) for the merger rate per Gyr than ours. Their simulations did not have sufficient mass resolution to determine the rate at $z \gtrsim 2$. It is important to note, however, that our *B/n* at redshift z measures the instantaneous rate of mergers during a small Δz interval at that redshift. By contrast, they studied the merging history of *present-day* haloes and measured only the rate of major mergers for the most massive progenitor at redshift z of a $z = 0$ halo (see their paragraph 4, Section 2). A detailed comparison is outside the interest of this paper.

Mergers of dark matter haloes are related to but not identical to mergers of galaxies. It typically takes the stellar component of an infalling galaxy extra time to merge with a central galaxy in a group or cluster after their respective dark matter haloes have been tagged as merged by the FOF algorithm. This time delay is governed by the dynamical friction time-scale for the galaxies to lose orbital energy and momentum, and it depends on the mass ratios of the galaxies and the orbital parameters (Boylan-Kolchin et al. 2008, and references therein). In addition to this difference in merger time-scale, the growth in the stellar mass of a galaxy is not always proportional to the growth in its dark matter halo mass. A recent analysis of the galaxy catalogue in the Millennium Simulation (Guo & White 2008) finds galaxy growth via major mergers to depend strongly on stellar mass, where mergers are more important in the build-up of stellar masses in massive galaxies while star formation is more important in galaxies smaller than the Milky Way. Extending the analysis of this paper to the mergers of *subhaloes* in the Millennium Simulation will provide the essential link between their and our results.

For similar reasons, our results for the evolution of the dark matter halo merger rate per Gyr [$(1+z)^{n_\mathrm{m}}$ with $n_\mathrm{m} \sim$ 2–2.3] cannot be trivially connected to the observed merger rate of *galaxies*. It is nonetheless interesting to note that a broad disagreement persists in the observational literature of galaxy merger rates. The reported power-law indices n_m have ranged from 0 to 5 (see e.g. Burkey et al. 1994; Carlberg, Pritchet & Infante 1994; Woods, Fahlman & Richer 1995; Yee & Ellingson 1995; Patton et al. 1997; Le Fèvre et al. 2000; Patton et al. 2002; Conselice et al. 2003; Bundy et al. 2004; Lavery et al. 2004; Lin et al. 2004). Berrier et al. (2006) followed the redshift evolution of subhalo mergers in *N*-body simulations and provided a more detailed comparison with recent observations by e.g. Lin et al. (2004) that find $n_\mathrm{m} < 1$. They attributed such a weak redshift evolution in the number of close companions per galaxy to the fact that the high merger rate per halo at early times is counteracted by a decrease in the number of haloes massive enough to host a galaxy pair.

The merger rates in this paper are global averages over all halo environments. The rich statistics in the Millennium Simulation allow for an in-depth analysis of the environmental dependence of dark matter halo merger rates, which we will report in a subsequent paper (Fakhouri & Ma, in preparation).

ACKNOWLEDGMENTS

We have enjoyed enlightening discussions with Mike Boylan-Kolchin, Liang Gao, Ari Laor, Simon White, Andrew Zentner and Jun Zhang. This work is supported in part by NSF grant AST 0407351. The Millennium Simulation data bases used in this paper and the web application providing online access to them were

constructed as part of the activities of the German Astrophysical Virtual Observatory.

REFERENCES

Benson A. J., Cole S., Frenk C. S., Baugh C. M., Lacey C. G., 2000, MNRAS, 311, 793
Benson A. J., Kamionkowski M., Hassani S. H., 2005, MNRAS, 357, 847
Berrier J. C., Bullock J. S., Barton E. J., Guenther H. D., Zentner A. R., Wechsler R. H., 2006, ApJ, 652, 56
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Boylan-Kolchin M., Ma C.-P., Quataert E., 2008, MNRAS, 383, 93
Bundy K., Fukugita M., Ellis R. S., Kodama T., Conselice C. J., 2004, ApJ, 601, L123
Burkey J. M., Keel W. C., Windhorst R. A., Franklin B. E., 1994, ApJ, 429, L13
Carlberg R. G., Pritchet C. J., Infante L., 1994, ApJ, 435, 540
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Conselice C. J., Bershady M. A., Dickinson M., Papovich C., 2003, AJ, 126, 1183
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499
Eisenstein D. J., Hu W., 1999, ApJ, 511, 5
Gottlöber S., Klypin A., Kravtsov A. V., 2001, ApJ, 546, 223
Governato F., Gardner J. P., Stadel J., Quinn T., Lake G., 1999, AJ, 117, 1651
Guo Q., White S. D. M., 2008, MNRAS, 384, 2
Harker G., Cole S., Helly J., Frenk C., Jenkins A., 2006, MNRAS, 367, 1039
Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., 2003, MNRAS, 338, 903
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Kang X., Jing Y. P., Mo H. J., Börner G., 2005, ApJ, 631, 21
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lavery R. J., Remijan A., Charmandaris V., Hayes R. D., Ring A. A., 2004, ApJ, 612, 679
Le Fèvre O. et al., 2000, MNRAS, 311, 565
Lin L. et al., 2004, ApJ, 617, L9
Lukic Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, ApJ, 671, 1160
Maller A. H., Katz N., Kereš D., Davé R., Weinberg D. H., 2006, ApJ, 647, 763
Murali C., Katz N., Hernquist L., Weinberg D. H., Davé R., 2002, ApJ, 571, 1
Patton D. R., Pritchet C. J., Yee H. K. C., Ellingson E., Carlberg R. G., 1997, ApJ, 475, 29
Patton D. R. et al., 2002, ApJ, 565, 208
Sheth R. K., Tormen G., 2002, MNRAS, 329, 61
Smoluchowski M., 1916, Phys. Zeit., 17, 557
Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V. et al., 2005, Nat, 435, 629
White M., 2001, A&A, 367, 27
Woods D., Fahlman G. G., Richer H. B., 1995, ApJ, 454, 32
Yee H. K. C., Ellingson E., 1995, ApJ, 445, 37
Zhang J., Ma C.-P., Fakhouri O., 2008, preprint (arXiv:0801.3459)

APPENDIX A: THE DURHAM TREE

In this paper we have used two methods to handle fragmentation events in the Millennium FOF merger trees: snipping and stitching. Here we discuss and compare a third method used by the Durham group (Helly et al. 2003; Bower et al. 2006; Harker et al. 2006).

The Durham algorithm is designed to reduce spurious linkings of FOF haloes in low-density regions. Before constructing the FOF merger tree, they filter through the Millennium FOF and subhalo data base, and split up a subhalo from its FOF halo if (1) the subhalo's centre is outside twice the half mass radius of the FOF halo or (2) the subhalo has retained more than 75 per cent of the mass it had at the last output time at which it was an independent halo (Harker et al. 2006). Condition (1) is effectively a spatial cut, while (2) is based on the argument that less massive subhaloes are expected to undergo significant stripping as they merge with more massive haloes. This algorithm then *discards* the subhaloes that are split off from FOF groups at $z = 0$, along with any associated progenitor subhaloes. Around 15 per cent of the original FOF haloes are split in this algorithm.

The Durham algorithm tends to reduce the number of fragmented haloes in the resulting trees, but it does not eliminate all such events. A method much like our snipping method is used to treat the remaining fragmentation events. The resulting FOF tree is available at the Millennium public data base along with the original Millennium tree.

To compare with our stitching and snipping trees, we have repeated all of our merger rate calculations and tests using the Durham tree. Fig. A1 shows the ratio of the resulting Durham merger rate, *B*/*n*, to that from our stitching tree at $z = 0$. The Durham rate is generally lower than our rate for minor mergers (by up to ∼30 per cent), and it drops precipitously for major mergers ($\xi \gtrsim 0.3$). The two additional conditions applied to split up subhaloes in the Durham algorithm therefore appear to have eliminated most of the major merger events.

Moreover, these splitting conditions in the Durham algorithm also modify the halo mass function. Fig. A2 shows the ratio of the Durham mass function to the fit of Jenkins et al. (2001) at $z = 0$, 0.5 and 1 (thick solid curves with error bars). The ratio of our stitching mass function to the same fit is overlaid for compari-

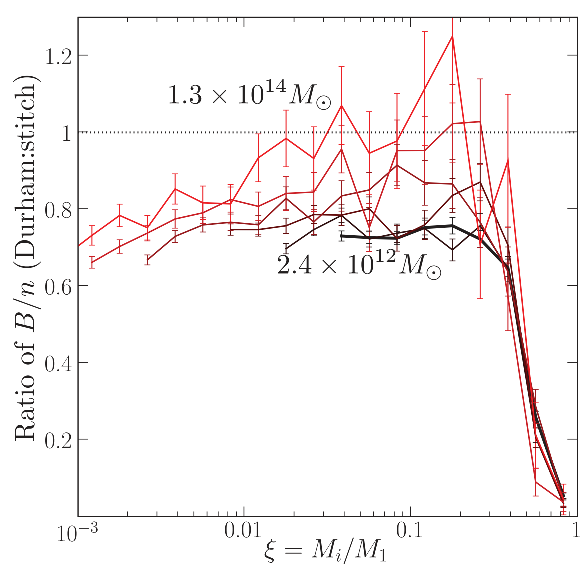

Figure A1. The ratio of the Durham merger rate *B*/*n* to our stitching rate *B*/*n* (Section 4.3) as a function of progenitor mass ratio ξ for a number of descendant mass bins ranging from $\sim 2 \times 10^{12}$ M$_\odot$ (black) to $\sim 10^{14}$ M$_\odot$ (red). The Durham merger rate tends to be lower than the stitching merger rate, and suffers a sudden drop in the major merger regime ($\xi \gtrsim 0.3$).

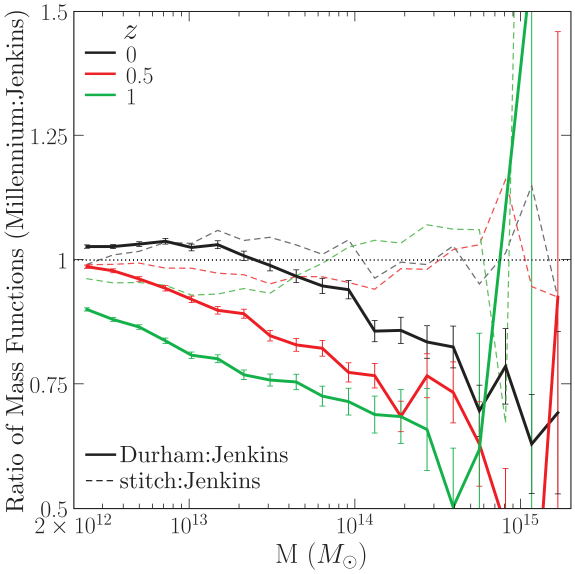

Figure A2. The ratio of the Durham halo mass function to the fit of Jenkins et al. (2001) at redshifts 0, 0.5 and 1 (thick solid curves with error bars). The ratio of the halo mass function from our stitching method to the same fit is shown for comparison (thin dashed curves). The Durham algorithm tends to reduce the masses of massive haloes, leading to a deficit that grows to $\sim$50 per cent at $\sim 10^{15}\,\mathrm{M}_\odot$ and at higher z.

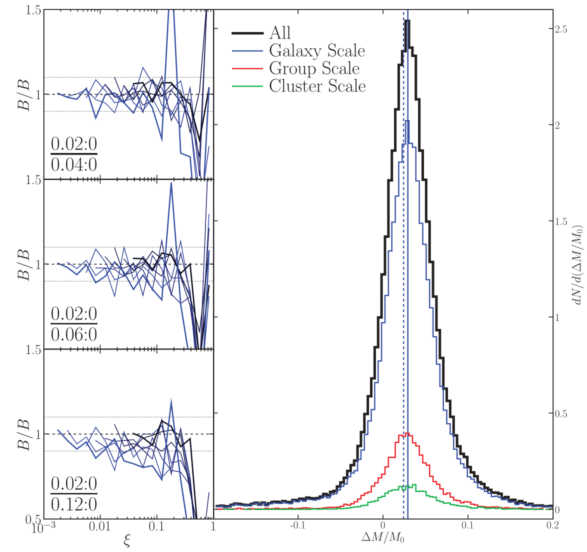

Figure A3. Left-hand panels: A subset of the Δz convergence matrix presented in Fig. 10, now computed using the Durham tree. Note the poor convergence properties of the major merger end ($\xi \gtrsim 0.3$). This corresponds to the region of the largest difference between the stitching and Durham merger rates (see Fig. 17). Right-hand panel: The distribution of $\Delta M/M_0$ for the $z = 0.06$:0 Durham catalogue (similar to Fig. 13).

son (thin dotted curves). The Durham mass function is systematically lower: the number of $z = 0$ haloes with $M \gtrsim 10^{14}\,\mathrm{M}_\odot$ is $\sim$25 per cent lower, and the difference increases at $z \sim 1$, affecting the halo mass function even at $M \sim 2 \times 10^{12}\,\mathrm{M}_\odot$.

We believe that the deficit of major merger events and massive haloes in the Durham catalogue is partially due to their second criterion that splits off subhaloes that have retained 75 per cent of their original mass. This condition may indeed remove spurious FOF linkings in the minor merger regime, but major merger events tend to preserve much of the original progenitor masses and have been systematically split by the Durham algorithm.

Finally, Fig. A3 (right-hand panel) shows that the Durham tree has similar mass conservation properties as our stitching tree in Fig. 13. The distribution of the mass in the 'diffuse' component, $\Delta M/M_0$, has a very similar peak of $\sim$3 per cent, although the negative ΔM events have been suppressed. For Δz convergence (left-hand panels; cf. Section 5.2), the Durham tree performs well in the minor merger regime but is consistently poor for major mergers, again probably due to the splitting condition (2) above.

The Durham algorithm is tuned to address questions of galaxy evolution. The issues we have uncovered regarding this algorithm are specifically for the mergers of dark matter haloes, the subject of this paper; issues with the mergers of galaxies will require a separate study.

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