Title
Mergers and Mass Assembly of Dark Matter Halos in a Lambda Cold Dark Matter Universe

Author
Fakhouri, Onsi Joe

Publication Date
2010

Peer reviewed|Thesis/dissertation

Mergers and Mass Assembly of Dark Matter Halos in a Λ-Cold Dark Matter Universe

by

Onsi Joe Fakhouri

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Astrophysics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Chung-Pei Ma, Chair

Professor Eliot Quataert

Professor Lawrence Hall

Fall 2010

Mergers and Mass Assembly of Dark Matter Halos
in a Λ-Cold Dark Matter Universe

Abstract

Mergers and Mass Assembly of Dark Matter Halos
in a Λ-Cold Dark Matter Universe

by

Onsi Joe Fakhouri

Doctor of Philosophy in Astrophysics

University of California, Berkeley

Professor Chung-Pei Ma, Chair

Using the combined results of the Millennium and Millennium II numerical simulations we compute a number of robust statistics for the growth of dark matter haloes in the ΛCDM concordance model. With an unprecedented 1.35 billion subhaloes we compute, and provide a fit to, the halo merger rate over five orders of magnitude in descendant halo mass ($10^{10} \leq M_0 \leq 10^{15} M_\odot$) and progenitor mass ratio ($10^{-5} \leq \xi \leq 1$) over a redshift range spanning $z = 0$ to 15. We also compute and categorize halo mass accretion histories and quantify the relative contributions made by the mergers of resolved progenitors and the accretion of unresolved diffuse material to the halo growth rate. In addition, we perform a detailed quantitative study of the environmental dependence of halo growth, finding a higher merger rate among halos in denser regions and, curiously, a negative correlation between environment density and the diffuse component of halo growth. Throughout, we pay special attention to the problem of extracting self-consistent merger trees from the Millennium simulations: numerical issues abound and the problem of halo fragmentation in particular is shown to be important at the 20% level, but tractable via a variety of post-processing algorithms which we describe. Finally, we apply the insight obtained from our analysis to the problem of generating self-consistent Monte Carlo merger trees in the extended Press-Schechter framework.

In the beginning was the Word, and the Word was with God, and the Word was God. He was with God in the beginning. Through Him all things were made; without Him nothing was made that Has been made. — John 1:1-3

For the Word that continues to speak today - in the merging of haloes, the birthing of life, the mystery of faith, the longing of hope, the joys of love, even the cries of the poor.

Contents

Acknowledgments

Many thanks to Chung-Pei for her passion and optimism - an excellent foil to my neurotic obsession with the minutia. I could not have asked for, or imagined, a better-matched advisor and have been blessed to work with her – you'll be missed! Thanks also to Jun Zhang for fun and lively conversations in the lounge, in his office, and on road trips to conferences.

Graduate school would have been impossible without the support of my friends at Veritas and the First Presbyterian Church of Berkeley. In particular the Physics Prayer Group and "Men's Group B" have been an invaluable spiritual resource, as has soon-to-be-Dr. Dan Garcia's listening ear and twitchy trigger finger. And no acknowledgements section would be complete without a shout-out to the indefatigable Carrie Bare who's humor and friendship brought much joy in the "dark" years.

I am acutely aware that graduate school is a privilege that is, all too often, reserved for those members of our society who come from a relatively privileged and stable background. I am no different and am thankful for, and indebted to, the love and support of my parents and my family - again, I am blessed beyond words.

I leave graduate school successful and (mostly) emotionally unscathed. I have two people to thank for that: my beautiful, loving, wife Hannah who has graciously and joyfully come alongside me again and again - your words are life giving and I love you! And Jesus Christ, Lord and savior, in whom "all things were created: things in heaven and on earth... He is before all things" (even dark matter haloes!) "and in Him all things hold together." Simply put: His resurrection makes all this madness meaningful.

And finally, the politics (for which I am thankful): the work in this thesis was supported in part by NSF grant AST 0407351 and every drop of data came directly from the brilliant work that went into the Millennium and Millennium II simulations. There would be no halos to analyze, no merger trees to clean up, and no bizzare halo trajectories to follow if it weren't for this impressive computational achievement, and so I quote: The Millennium Simulation databases used in this thesis and the web application providing online access to them were constructed as part of the activities of the German Astrophysical Virtual Observatory.

Part I

Introduction

Dark matter haloes, the hosts of luminous galaxies like our Milky Way, are understood to form hierarchically in modern cosmological theory. This means that large, massive, haloes do not simply condense out of the primordial dark matter distribution; rather, haloes grow in mass and size primarily by merging with other haloes. As these mergers take place the centrally concentrated baryonic components - the galaxies - sink via dynamical friction and subsequently merge. As a result, the stellar components of galaxies also grow via mergers: both via the combination of stellar populations in galaxy-galaxy mergers and the uptick in star formation rate observed when gaseous systems merge.

While the mergers of galaxies and the mergers of dark matter haloes are not identical processes, the two are closely related. In particular, the merger history of a dark matter halo affects many important theoretical and observational properties of galaxies including: star formation rates, color and morphology transformations, formation redshifts, galaxy occupation numbers, dynamical states of stellar disks, and galaxy mass and luminosity functions. To understand the observed distributions of these properties several theoretical models have been constructed that require, as a primary input ingredient, the growth rates and merger rates of dark matter haloes.

In this thesis we quantify these rates in detail, providing fitting forms for both the merger rate and the mass accretion history ($M(z)$) of dark matter haloes that are valid over several orders of magnitude in mass and mass ratio, and out to redshifts as high as $z \sim 15$. We accomplish this by analyzing the Millennium [Springel et al., 2005] and Millennium-II [Boylan-Kolchin et al., 2009b] simulations. These are among the largest n-body cosmological simulations ever run: the Millennium simulation uses 2160^3 particles with a particle mass of $1.2 \times 10^9 M_\odot$ to follow the evolution of 2×10^7 dark matter haloes from $z = 127$ to $z = 0$ in a cosmological box of size 685 Mpc. The Millennium-II simulation uses the same number of particles but has 125 times better mass resolution. When combined the two simulations provide us with $\sim 750,000$ dark matter haloes at $z = 0$ with > 1000 particles and over 1.35 billion subhalo progenitors with which to produce merger trees rooted at these haloes out to $z \sim 15$.

By contrast, earlier halo growth studies have either depended on Monte Carlo merger tree realizations based on the analytical extended Press-Schecter (EPS; Lacey and Cole 1994, Bond et al. 1991) model (e.g. Kauffmann et al. 1993, Somerville and Primack 1999, Springel et al. 2001a) or on relatively low-resolution N-body simulations. For example: Governato et al. [1999] studied $z < 1$ major mergers in an N-body simulation with 144^3 particles, Gottlöber et al. [2001] used a sample of ~ 4000 haloes to study the environmental dependence of the major merger rate at $z < 2$, Berrier et al. [2006] studied major mergers in a 171Mpc box with 512^3 particles, Murali et al. [2002] and Maller et al. [2006] studied galaxy merger rates using ~ 500 galaxies in SPH simulations. Similarly, early studies of halo mass accretion histories have only been performed on low resolution simulations: the work of Wechsler et al. [2002] analyzed only ~ 900 haloes, Tasitsiomi et al. [2004] studies 14 haloes, Cohn and White [2005] studies ~ 1500 cluster-sized haloes.

The rich statistics afforded by the Millennium simulations allow us to push beyond first-order measurements of halo growth and merger rates. We also quantify the second-order

effects on these growth rates due to halo environment. Here, again, there is an abundance of early low-resolution work [Gottlöber et al., 2001, Sheth and Tormen, 2004, Gao et al., 2005, Harker et al., 2006, Wechsler et al., 2006, Jing et al., 2007, Wang et al., 2007, Gao and White, 2007, Maulbetsch et al., 2007] though much of this work used halo clustering, an unintuitive statistical quantity, to compute environmental effects. The superior statistics of the Millennium simulations allow us to employ robust, intuitive, local overdensity measures for environment instead.

Unfortunately, the rich merger trees obtained from the Millennium simulations are not completely free of numerical inconsistencies and ambiguities. Two issues, in particular, are of primary importance in computing robust and reliable merger and growth rates and we exert much effort in this thesis to resolve them. The first is the issue of mass definition: we find, in agreement with White [2001] that spherical over density mass measures are inferior, in the merger context, to those derived by the friends-of-friends Davis et al. [1985]. Moreover, we agree with Genel et al. [2009] that the use of the bound mass associated with subhaloes, instead of the total FOF mass is better suited for measuring robust halo growth and merger rates.

The second numerical hurdle we face is the issue of halo fragmentation. These are events in which the particles in a given progenitor halo end up in more than one descendant halo. We find that these events are common enough and severe enough to warrant special attention: in order to compute self-consistent merger rates these fragmentation events must be filtered out using a post-processing algorithm. We define a variety of these algorithms and compare and contrast their effects on our results throughout the thesis.

This thesis consists of six chapters, each containing a paper published in the Monthly Notices of the Royal Astronomical Society between 2008 and 2010. The organization is as follows:

In part II we compute and present fits to the merger rates of dark matter haloes. Chapter 1 presents results for the Millennium simulation along with a suite of tests to ensure the robustness of the resulting merger rates. We introduce and quantify the problem of halo fragmentation and present the snipping and stitching post-processing algorithms to deal with it.

Chapter 2 extends the work in chapter 1 to the Millennium-II simulation and provides our most statistically well-founded fit for the dark matter halo merger rate. The results in this chapter are valid over five orders of magnitude in descendant halo mass ($10^{10} \lesssim M_0 \lesssim 10^{15} M_\odot$), five orders of magnitude in mass ratio ($10^{-5} \lesssim \xi \leq 1$) and a wide range of redshift ($0 \leq z \lesssim 15$).

Part III (chapter 3) presents a detailed study of the mass accretion histories (MAH) of dark matter haloes. In this chapter we categorize the qualitative behavior of halo growth trajectories ($M(z)$) and propose a two-parameter fitting scheme that faithfully reproduces the majority of MAHs and outperform earlier fitting models. We also provide fits to the joint probability distribution of these two fitting parameters thereby enabling others to compute Monte Carlo distributions of MAHs that statistically match the Millennium simulation.

In part IV we focus on the environmental dependence of halo growth. Chapter 4 intro-

duces and discusses δ, the overdensity parameter we use to quantify the local environment of each halo in the Millennium simulation. We also present the dependence of the halo merger rate on halo environment (we find a positive correlation: halos in denser regions experience higher merger rates) and discuss the sensitivity of these measurements to our choice of post-processing algorithm.

Chapter 5 quantifies the environmental dependence of the halo growth rate. This rate is sensitive to both the merger rate quantified in chapter 4 and the rate of diffuse accretion. We find that the diffuse component is, surprisingly, *negatively* correlated with environment and note that this explains the curious correlation between formation redshift and halo clustering observed in the literature. We also discuss potential candidates that might make up this diffuse growth component.

Finally, part V (chapter 6) connects the insight obtained from the Millennium simulation merger rates to the problem of generating self-consistent Monte Carlo merger trees in the EPS framework. A number of algorithms already exist to generate these merger trees. We analyze and compare these algorithms in detail and outline their deficiencies. We then discuss the necessary properties that an algorithm must possess to generate a self-consistent suite of merger trees and show that EPS does not predict a unique algorithm by proposing three new, distinct, self-consistent algorithms. We compare the higher-moment statistics generated by these three algorithms to the results obtained from the Millennium simulation and show that the analytic EPS model cannot match all the merger statistics in N-body simulations simultaneously; modification to the model is therefore required.

The cosmology used throughout this thesis is identical to that used in the Millennium simulation: a ΛCDM 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.

Part II

The Merger Rates of Dark Matter Haloes

Chapter 1

The Nearly Universal Merger Rate of Dark Matter Haloes in ΛCDM Cosmology

Fakhouri, Onsi; Ma, Chung-Pei — May 2008
Monthly Notices of the Royal Astronomical Society Volume 386, Issue 2, pp. 577-592

We construct merger trees from the largest database 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 ($10^{12} \lesssim M_0 \lesssim 10^{15} M_\odot$), progenitor mass ratio ($10^{-3} \lesssim \xi \leq 1$), and redshift ($0 \leq z \lesssim 6$). 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%. 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 $(1+z)^{n_m}$ with $n_m = 2 - 2.3$, 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 vs 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.

1.1 Introduction

In hierarchical cosmological models such as 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 realisations of the merger rates given by the analytical extended Press-Schechter (EPS; Lacey and Cole 1993, Bond et al. 1991) model (e.g. Kauffmann et al. 1993, Somerville and Primack 1999, Springel et al. 2001a). 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 vs 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 database 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 $z = 0$. This dataset allows us to determine the merger rates of dark matter haloes ranging from galaxy-mass scales of $\sim 10^{12} M_\odot$ over redshifts $z = 0$ to ~ 6, to cluster-mass scales up to $\sim 10^{15} M_\odot$ 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 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, that is, 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 et al. [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 and White [2008] used the semi-analytical galaxy catalogue of De Lucia et al. [2006] based on the Millennium simulation.

This paper is organised as follows. Section 1.2 describes the dark matter haloes in the Millennium simulation (§1.2.1) and how we construct the merger trees (§1.2.2) . We then discuss the issue of halo fragmentation and the two methods ('snipping' and 'stitching') used to treat these events in §1.2.3. The notation used in this paper is summarised in §1.2.4.

Section 1.3 describes how mergers are counted (§1.3.1) and presents four (related) statistical measures of the merger rate (§1.3.2). The relation between these merger rate statistics and the analytical merger rate based on the Extended Press-Schechter (EPS) model is derived in Section 1.3.3.

Our main results on the merger rates computed from the Millennium simulation are presented in Section 1.4. 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 (§1.4.1). The evolution of the merger rates with redshifts up to $z \sim 6$ is discussed in Section 1.4.2. 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% level) over a wide range of parameters (§1.4.3).

Section 1.5 compares the stitching and snipping merger rates (§1.5.1) 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 (§1.5.2); how the counting of binary vs multiple progenitor mergers affects the merger rates (§1.5.3); mass non-conservation arising from 'diffuse' accretion in the form of unresolved haloes during mergers (§1.5.4); and how the definition of halo masses and the treatment of fragmentation events affect the resulting halo mass function (§1.5.5).

In Section 1.6, 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 (§1.6.1). Section 1.6.2 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 [Bower et al., 2006, Harker et al., 2006, Helly et al., 2003]. 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 ΛCDM 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.

1.2 Haloes and Merger Trees in the Millennium Simulation

1.2.1 Dark Matter Haloes

The Millennium simulation provides the largest database 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 $z = 0$ [Springel et al., 2005].

The haloes in the simulation are identified by grouping the simulation particles using the standard friends-of-friends algorithm (FOF: 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., 2001a]. 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 database. Each FOF halo is then given two definitions of mass: M_{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_{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_{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 1.5.4.

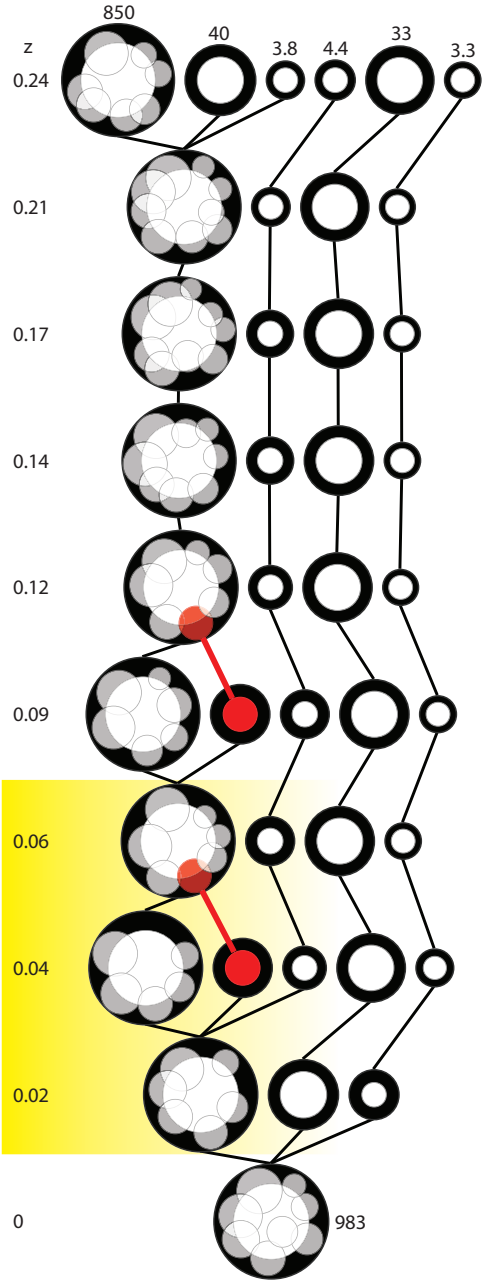

Figure 1.1: Example of a typical FOF merger tree extracted from the Millennium database. Black circles denote FOF haloes; white circles within black circles denote subhaloes. The radius of each circle is proportional to the log of the mass of the object; the black circles are further scaled up by a factor 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. 1.2. The numbers above the haloes at $z = 0.24$ and to the right of the final descendant FOF at $z = 0$ correspond to the FOF masses (in units of $10^{10} M_{\odot}$).

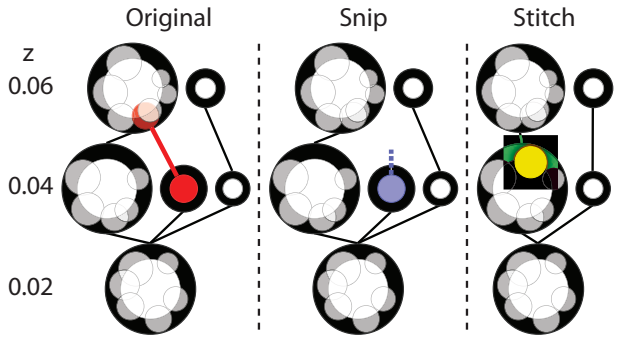

Figure 1.2: Left: A closeup of the highlighted (yellow region) fragmentation event in Fig. 1.1. The middle and right panels illustrate how the snipping and stitching methods handle fragmentation in order to assign a unique descendant halo. The blue circle (centre panel) shows the snipped orphan subhalo, and the yellow circle (right panel) shows how that subhalo is stitched. The black, white, and red circles are the same as in Fig. 1.1.

1.2.2 Merger Tree Construction

Merger trees of dark matter haloes in the Millennium database 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 database 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. 1.1 with an actual merger tree taken from the Millennium database. The upper left corner, for example, shows three FOF haloes at $z = 0.24$ with masses 8.5×10^{12}, 4×10^{11}, and $3.8 \times 10^{10} M_{\odot}$ merging into a single FOF halo at the next Millennium output ($z = 0.21$). The largest FOF halo

[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.

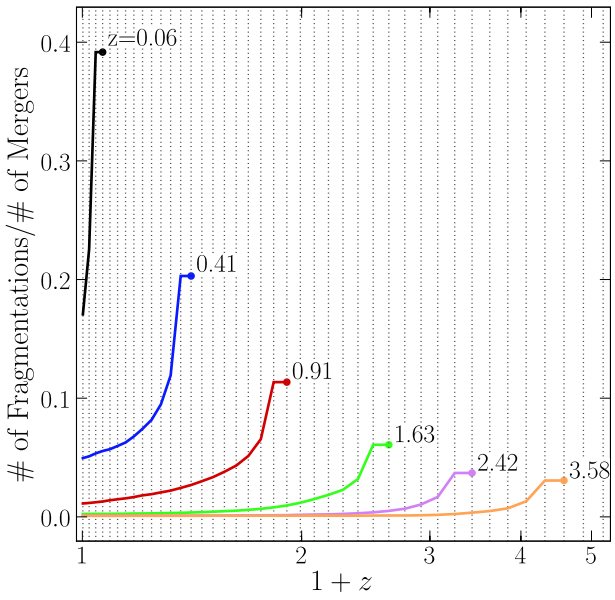

Figure 1.3: Distribution of the ratio of fragmentation to merger events as a function of redshift. The dotted vertical lines correspond to the redshifts of the Millennium outputs. We choose 6 redshifts (labelled) for illustrative purposes and plot the ratio of the number of fragmentations to the number of mergers (filled circles) at each redshift. A mass ratio cutoff is applied: both the fragments and mergers must have mass ratios exceeding 10%. 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 2-3 simulation outputs. Finally, the six filled circles decrease with increasing redshift, reaching $\sim 40\%$ at $z = 0$ but dropping to $\sim 5\%$ at high z – this is primarily due to the increasing Δz between Millennium outputs.

$z = 0.24$ has 7 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. 1.1.

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. 1.1 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 emphasise 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.

Fig. 1.2 illustrates these two methods for the fragmentation event shown in the highlighted (yellow) region of Fig. 1.1. The snipping method is commonly used in the literature (e.g. Sheth and Tormen 2002), presumably for its simplicity. Fragmentation events are removed by 'snipping' the link between the smaller descendant halo and its progenitor FOF halo, as shown in the middle panel of Fig. 1.2. The fragmenting progenitor FOF halo then has only one descendant FOF halo. We note that this method can result in a number of progenitor-less *orphan* FOF haloes (e.g., the blue subhalo in Fig. 1.2).

In this paper we investigate an alternative method that we name 'stitching.' This method is motivated by our observation that about half of the fragmented haloes in the Millennium simulation remerge within the following 2-3 outputs (see below). The two fragmentation events in Fig. 1.1 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 timescales of a few Gyrs; see, e.g., Boylan-Kolchin et al. 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 3 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 panel of Fig. 1.2); 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 3 outputs is described in Section 1.5.1.

As can be seen in Fig. 1.2, 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. 1.3, where the ratio of fragmentation events to merger events is seen to peak at 40% for major fragmentation events (defined to be fragmentations where the fragment subhalo carries 10% 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. 1.3 (filled circles), the drop of each curve with decreasing z tracks how many of them have remerged by that redshift. As

[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.

z_P:z_D	Galaxy-Scale			Group-Scale			Cluster-Scale		
	$N_p = 1$	$N_p = 2$	$N_p > 2$	$N_p = 1$	$N_p = 2$	$N_p > 2$	$N_p = 1$	$N_p = 2$	$N_p > 2$
0.06:0	188,400	65,711	27,939	1,063	2,418	13,256	3	25	5,356
0.56:0.51	189,351	61,718	22,031	1,212	2,468	9,374	6	18	3,014
1.08:0.99	145,779	68,467	35,426	325	878	7,630	0	2	1,308
2.07:1.91	76,298	52,525	39,097	31	77	2,225	0	0	129
3.06:2.83	30,641	26,675	25,072	0	4	343	0	0	4

Table 1.1: 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 ($N_p = 1$, i.e., no mergers), two progenitors ($N_p = 2$, i.e. binary mergers), and multiple progenitors ($N_p > 2$), for three separate descendant mass bins: $2 \times 10^{12} \leq M_0 < 3 \times 10^{13} M_\odot$ (galaxy), $3 \times 10^{13} \leq M_0 < 10^{14} M_\odot$ (group), and $M_0 \geq 10^{14} M_\odot$ (cluster). Only progenitor haloes with mass $> 4.8 \times 10^{10} M_\odot$ (40 simulation particles) are counted.

noted above, we find that about half of the fragmented haloes remerge within 2-3 outputs (corresponding to a fixed $\Delta z/(1+z)$ as the outputs are log-spaced). Given a fragmentation-to-merger ratio of 40%, and a remerger rate of 50%, the remerging fragments can impact the merger rate measurements inflating them at the ∼20% level.

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% of the original FOF mass, the fragmentation-to-merger ratio at $z = 0$ ($z = 1.6$) jumps to 57% (13.3%) vs 39% (6%) for major fragmentations. For very minor fragmentations (subhalo fragments that carry less than 1% of the total mass) the fragmentation-to-merger ratios are 85% and 28% at $z = 0$ and $z = 1.6$ respectively. Thus we anticipate that fragmentation events will more severely pollute the minor-merger regime of the merger rate statistics.

1.2.4 Notation

We apply both the stitching and snipping methods and produce FOF merger trees from the 46 Millennium outputs that span $z = 0$ and $z = 6.2$. 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 Myrs) 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_{FOF}

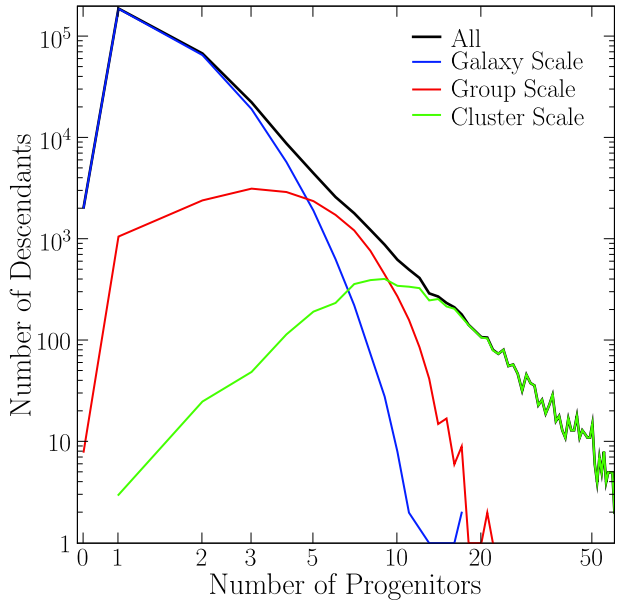

Figure 1.4: Distribution of the number of progenitors, N_p, for the $z = 0.06{:}0$ merger tree. There are $\sim 300,000$ descendant FOF haloes at redshift 0 (black) with $M_0 \geq 2\times10^{12}M_\odot$. Of these $\sim 280,000$ have $2 \times 10^{12} \leq M_0 < 3 \times 10^{13}M_\odot$ (galaxy-scale; dark blue), $\sim 16,000$ have $3 \times 10^{13} \leq M_0 < 10^{14}M_\odot$ (group scale; red), and $\sim 5,400$ have $M_0 \geq 10^{14}M_\odot$ (cluster-scale; green).

mass, N_p to denote the number of progenitor haloes, and M_i with $i \in (1, 2, \ldots, N_p)$ to denote the rank-ordered M_{FOF} mass of the progenitors, i.e. $M_1 \geq M_2 \geq \ldots M_{N_p}$. We impose a minimum mass cutoff of $M_0 \geq 2 \times 10^{12}M_\odot$ on the descendant FOF halo and a cutoff of $M_i \geq 4.8 \times 10^{10}M_\odot$ on the progenitors, which corresponds to 40 particles and is twice the minimum halo mass in the Millennium database.

For certain results reported below, we make use of three large mass bins: $2\times10^{12} \leq M_0 < 3 \times 10^{13}M_\odot$, $3 \times 10^{13} \leq M_0 < 10^{14}M_\odot$, and $10^{14}M_\odot \leq M_0$, referred to as the *galaxy-scale*, *group-scale*, and *cluster-scale* bins, respectively.

1.3 Merger Statistics and Connection to EPS

1.3.1 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. $N_p > 2$). For completeness, we list in Table 1.1 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 $N_p = 1, 2$ 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. 1.4 shows the distribution of the number of progenitors, $f(N_p)$, 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 $(62, 22, 16)\%$ of the haloes have $N_p = (1, 2, > 2)$ 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 peak of $f(N_p)$ moves to higher N_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. $N_p = 1$)

For completeness, we include *all* the progenitors (above our minimum mass cutoff 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 $i \geq 2$ merges with M_1, the most massive progenitor, at some stage between the two outputs. Thus a descendant halo with N_p progenitors is assumed to be the result of a sequence of $(N_p - 1)$ binary merger events, where each merger event is assigned a mass ratio

$$\xi \equiv \frac{M_i}{M_1}, \quad i = 2, ..., N_p \tag{1.1}$$

which by construction satisfies $\xi \leq 1$. This assumption ignores the possibility that two smaller progenitor FOF haloes merge together before merging with the most massive progenitor. Section 1.5.3 describes how we have tested the validity of this assumption and found negligible effects as long as a sufficiently small Δz is used.

1.3.2 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 dt/dz. 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

$$B_{MM'}(M, M', z_P{:}z_D) dM dM', \tag{1.2}$$

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_P and the descendant output redshift z_D. Note that $B_{MM'}(M, M')$ has units of $\left[\text{number of mergers} \times (\Delta z)^{-1}\,\text{Mpc}^{-3}\,M_\odot^{-2}\right]$ and generally depends on both z_P and z_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 $B_{MM'}(M, M')dMdM'$ to

$$B(M_0, \xi, z_P{:}z_D)dM_0 d\xi \,, \tag{1.3}$$

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_D that have progenitor FOF haloes at z_P with mass ratio in the range of $(\xi,\ \xi + d\xi)$, where $\xi = M_i/M_1, i \geq 2$ as discussed in Section 1.3.1. The quantity $B(M_0, \xi)$ therefore has units of

$$\left[\text{number of mergers} \times\ \Delta z^{-1}\,\text{Mpc}^{-3}\,M_\odot^{-1}\,d\xi^{-1}\right]. \tag{1.4}$$

In the mass-conserving binary limit of $M_0 = M + M'$ and $\xi = M'/M$ (where $M' < M$), $B_{MM'}$ and B in equations (1.2) and (1.3) 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 (1.2) and (1.3) in the lower mass bins of M, M', or M_0. It is useful to normalise 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:

$$\frac{B}{n} \equiv \frac{B(M_0, \xi, z_P{:}z_D)}{n(M_0, z_D)} \,, \tag{1.5}$$

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

$$\left[\text{number of mergers/number of descendants} \times\ (\Delta z)^{-1}\,(d\xi)^{-1}\right], \tag{1.6}$$

which is dimensionless. The mass function $n(M_0, z)dM_0$ 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 $\xi \in (x, X)$,

$$\frac{d\bar{N}_{\text{merge}}}{dz}\left([m, M], [x, X], z_P{:}z_D\right) \,, \tag{1.7}$$

is simply an integral over $B(M_0, \xi, z_P{:}z_D)$:

$$\frac{d\bar{N}_{\text{merge}}}{dz} \equiv \frac{1}{N}\int_m^M\!\!\int_x^X B(M_0, \xi, z_P{:}z_D)\, d\xi\, dM_0 \,, \tag{1.8}$$

where

$$N \equiv \int_m^M n(M_0, z_D)\, dM_0 \tag{1.9}$$

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

$$\frac{d\bar{N}_{\rm merge}}{dz} \sim \int_x^X \frac{B}{n}\, d\xi\,. \tag{1.10}$$

1.3.3 Connection to EPS

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

$$\frac{d^2p}{d\ln \Delta M^{LC}\, dt}(M_1^{LC} \to M_2^{LC}|t)\,, \tag{1.11}$$

the probability that a halo of mass M_1^{LC} will merge with another halo of mass $\Delta M^{LC} = M_2^{LC} - M_1^{LC}$ in time interval dt. Their notation (which we denote with superscripts 'LC') is related to ours by $M_2^{LC} \to M_0$, with M_1^{LC} and ΔM^{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 $d^2p/d\ln\Delta M^{LC}/dt$ to B/n, we first multiply it by $n(M_1^{LC})$, 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 1.1 and Fig. 1.4 show the distributions of the progenitor multiplicity N_p. We defer to Section 1.5 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 (1.11) is not symmetric in the progenitor masses M_1^{LC} and ΔM^{LC}, in contrast to our merger rate $B_{MM'}$ in equation (1.2), 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^{LC} is chosen to be the bigger or smaller progenitor. We will examine both options below: (A) $\xi = \Delta M^{LC}/M_1^{LC} \leq 1$ and (B) $\xi = M_1^{LC}/\Delta M^{LC} \leq 1$.

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

$$\frac{B(M_0, \xi, z)}{n(M_0, z)} \leftrightarrow \sqrt{\frac{2}{\pi}} \frac{d\delta_c}{dz} \frac{1}{\sigma(M')} \left|\frac{d\ln\sigma}{d\ln M}\right|_{M'} \left[1 - \frac{\sigma^2(M_0)}{\sigma^2(M')}\right]^{-3/2} \tag{1.12}$$

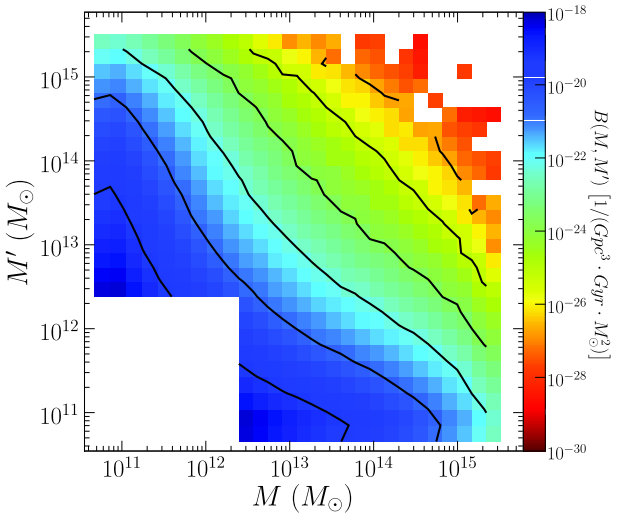

Figure 1.5: Symmetric merger rate $B_{MM'}$ of equation (1.2) as a function of progenitor masses M and M' computed from the $z = 0.06$:0 Millennium merger tree. The merger rates decrease from blue to red; the overlaid black lines are contours of constant merger rates.

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 $\delta_c(z) \propto 1/D(z)$ 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 (1.12). Also note that both sides of equation (1.12) 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 1.6.1.

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%. Details of the comparison are discussed in Section 1.5.1.

1.4.1 Merger Rates at $z \approx 0$

Fig. 1.5 is a contour plot of the symmetric merger rate in equation (1.2), $B_{MM'}(M, M', z_P{:}\ z_D)$, 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)

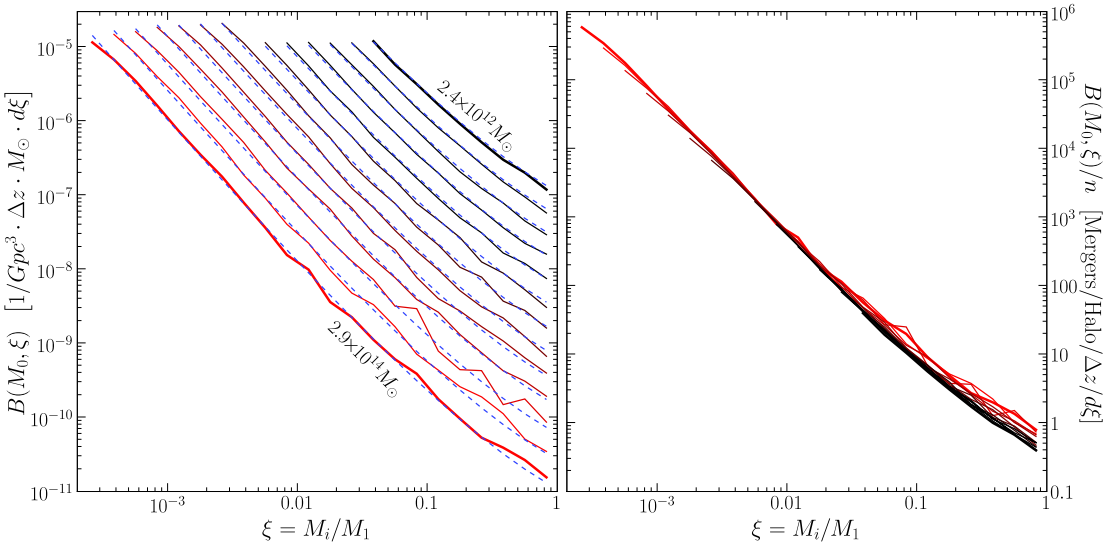

Figure 1.6: Left panel: Mean merger rate $B(M_0, \xi)$ of equation (1.3) for the $z = 0.06$:0 merger tree as a function of the mass ratio of the progenitors, ξ, for bins of fixed descendant halo mass M_0 (colour coded from black to red for increasing M_0). The overlaid dashed blue lines are from our fitting formula in equation (1.14). Note that the presence of a fixed minimum mass resolution ($4.7 \times 10^{10} M_\odot$) corresponds to a minimum mass ratio ξ that decreases as M_0 increases. Right panel: Mean merger rate *per halo*, $B(M_0, \xi)/n(M_0)$, of equation (1.5) for the same tree. Dividing out the halo number density $n(M_0)$ brings the curves on the left panel to nearly a single curve, indicating B/n has very weak dependence on M_0.

are more common than major mergers (along the diagonal). The lower left corner is blank due to our lower cutoff on the descendant FOF mass (~ 1000 particles; $M_0 \gtrsim 2 \times 10^{12} M_\odot$). The noisy nature of the upper right corner is due to limited merger statistics at $\sim 10^{15} M_\odot$.

As we discussed in Section 1.3.2, 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. 1.6 for the same dataset as in Fig. 1.5. The left panel plots the merger rate $B(M_0, \xi, 0.06\!:\!0)$ of equation (1.3) 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 (1.14) and Table 1.2 below.

The main quantity we study in this paper is the mean merger rate *per descendant halo*, B/n, of equation (1.5), shown in the right panel of Fig. 1.6. The rising amplitude of B with decreasing M_0 is remarkably largely removed when B/n is plotted: the curves in the left panel for different M_0 mass bins collapse onto nearly a single curve in the right 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. 1.7 and is also reported in Guo and White [2008]. As we will quantify in Section 1.4.3 below, the dependence on M_0 is approximately $\propto M_0^{0.08}$.

Our lower cutoff of 40 particles for the progenitor FOF halo mass implies a lower cutoff in the mass ratio of $\xi \geq 4.8 \times 10^{10} M_\odot / M_0$. This resolution cutoff is seen in the left panel of Fig. 1.6, 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 $M_0 > 5 \times 10^{13} M_\odot$; whereas the dynamic range is smaller for galaxy-size haloes, e.g., $\xi > 0.01$ for $M_0 < 5 \times 10^{12} M_\odot$.

The present-day merger rates shown in Fig. 1.6 are all obtained from the $z = 0.06$:0 merger tree. The low-redshift outputs available from the Millennium database in fact have a smaller spacing of $\Delta z \sim 0.02$. 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 Sec. 1.2.3). In practice, this precaution is not critical and we find little difference between the 0.06:0 and 0.02:0 results.

1.4.2 Merger Rates at Higher Redshift

Figs. 1.6 and 1.7 summarise our results for the $z = 0$ merger rates. At higher redshifts, the Millennium database provides sufficient halo statistics for us to measure merger rates up to $z \sim 6$. The results are shown in Fig. 1.8, where we plot the merger rate per unit time (upper panel), $d\bar{N}_{\rm merge}/dt$, and per unit redshift (lower panel), $d\bar{N}_{\rm merge}/dz$, as a function of redshift for three ranges of descendant masses (galaxy, group, cluster) and four ranges of progenitor mass ratios ($\xi \geq 1/3$, $1/10$, $1/30$, and $1/100$). Errors are computed assuming Poisson statistics for the number of mergers and haloes. We have suppressed merger rates

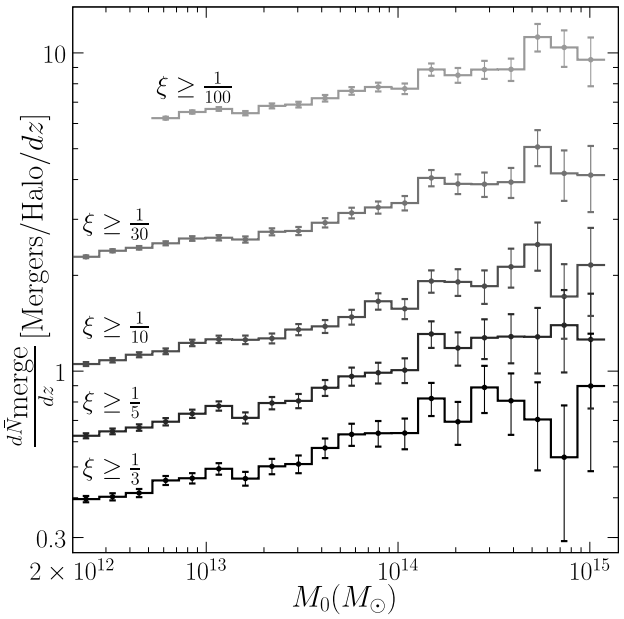

Figure 1.7: Mean merger rate per halo (per unit z), dN_m/dz, as a function of descendant mass, M_0, for various ranges of the progenitor mass ratio ξ. The upper curves include increasingly more minor mergers. The z=0.06:0 merger tree is used. Note the weak mass dependence over three decades of mass. The error bars are computed assuming Poisson counting statistics in both the number of mergers and the number of haloes.

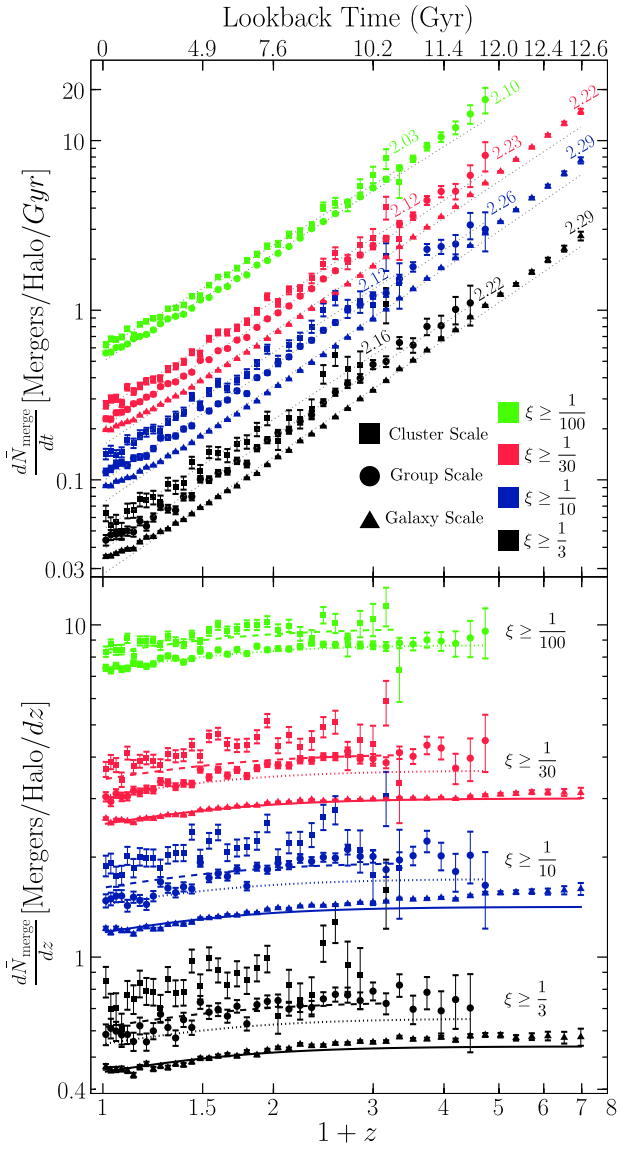

Figure 1.8: Upper panel: Mean merger rate per halo (per Gyr), dN_m/dt, as a function of redshift for three bins of descendant mass M_0 and four ranges of progenitor mass ratio ξ from the Millennium simulation (using the stitching tree). The overlaid lines plot the best-fit power laws, $(1+z)^{n_m}$, 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 dN_m/dz per unit z instead of per Gyr. The dotted grey lines here show our fitting formula in equation (1.14), 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 $\xi \geq \frac{1}{100}$ (green) is suppressed because such minor mergers fall below the simulation resolution limit.

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

Table 1.2: Best fit parameters for equation (1.14).

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

$$\frac{d\bar{N}_{\rm merge}}{dt} \propto (1+z)^{n_m} \tag{1.13}$$

and find $n_m \sim 2-2.3$ 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.

A large number of merger rate statistics can be easily read off of Fig. 1.8. For example, at around $z = 2$ ($z = 4$) every FOF halo on average experiences $\sim$ 2-4 (10) minor mergers ($\xi \lesssim 1/30$) per Gyr, and about 10-20% (70-90%) of FOF haloes experience a major merger ($\xi \gtrsim 1/3$) every Gyr.

Unlike the rising $d\bar{N}_{\rm merge}/dt$, the merger rate *per unit redshift*, $d\bar{N}_{\rm merge}/dz$, shows a remarkably weak dependence on z in Fig. 1.8 (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 (Sec. 1.4.3).

At $z > 0$, Fig. 1.8 shows that the dependence of $d\bar{N}_{\rm merge}/dz$ on progenitor ratio ξ and descendant mass M_0 is similar to the $z = 0$ merger rates shown in Fig. 1.6: 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).

1.4.3 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%. 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 panel of Fig. 1.6, and the weak redshift dependence shown in the bottom panel of Fig. 1.8. We find that the following functional form works well:

$$\frac{B(M_0,\xi,z)}{n(M_0,z)} = A\left(\frac{M_0}{\tilde{M}}\right)^{\alpha} \xi^{\beta} \exp\left[\left(\frac{\xi}{\tilde{\xi}}\right)^{\gamma}\right] \left(\frac{d\delta_c}{dz}\right)^{\eta}, \tag{1.14}$$

where $\tilde{M} = 1.2 \times 10^{12} M_\odot$ is a constant and $\delta_c(z) \propto 1/D(z)$ is the standard density threshold normalised to $\delta_c = 1.686$ at $z = 0$, with $D(z)$ being the linear growth factor. Note that equation (1.14) is separable with respect to the three major variables M_0, ξ, and z.

The form of the redshift dependence in equation (1.14) is chosen so that $\eta = 1$ corresponds to the EPS prediction in equation (1.12). In addition, this form has weak z dependence at $z \gtrsim 1$ since the growth factor approaches that of the Einstein-de Sitter model, $\delta_c(z) = 1.68(1 + z)$, and $d\delta_c/dz$ approaches a constant. This behaviour matches the weak redshift dependence seen in the Millennium merger rate (bottom panel of Fig. 1.8).

To determine the parameters in equation (1.14), we fit simultaneously to all redshifts $z < 1$, mass ratios $\xi > 10^{-3}$, and masses $10^{12} \lesssim M_0 \lesssim 10^{14} M_\odot$. The B/n data points are weighted using their Poisson distributed errors. The resulting fits are plotted as dotted curves in Figs. 1.6 and 1.8, and the fitting parameters are given in Table 1.2, 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 to 3 for snipping.

We note that the fitting form of equation (1.14) 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 1.3.3, which is intrinsically asymmetric in M_1 and M_2.

1.5 Tests

1.5.1 Snipping vs Stitching Trees

Fig. 1.9 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% over 2-3 orders of magnitude in both the progenitor mass ratio ξ and the descendant mass M_0. Within this difference, however, Fig. 1.9 and Table 1.2 show that the snipping method systematically 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. 1.2). Moreover, as discussed in Section 1.2.3, 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. 1.9.

A remaining issue is our choice of the stitching criterion: as described in Section 1.2.3, we stitch only FOF fragments that are observed to remerge within the next 3 outputs. This choice is motivated by the fact that about half of the halo fragments at a given output will

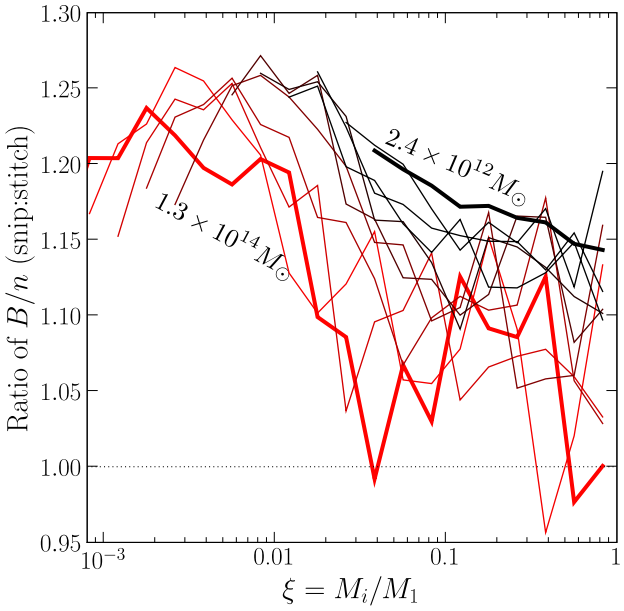

Figure 1.9: The ratio of the snipping and stitching B/n as a function of mass ratio ξ computed using the 0.06:0 catalogue for a variety of mass bins in the range $2.4 \times 10^{12} M_\odot$ (black) $\leq M \leq 1.3 \times 10^{14}$ (red). We find differences at the 25% level at low ξ with the snipping method consistently predicting a higher merger rate at all ξ. We attribute this to the population of remerging orphan haloes.

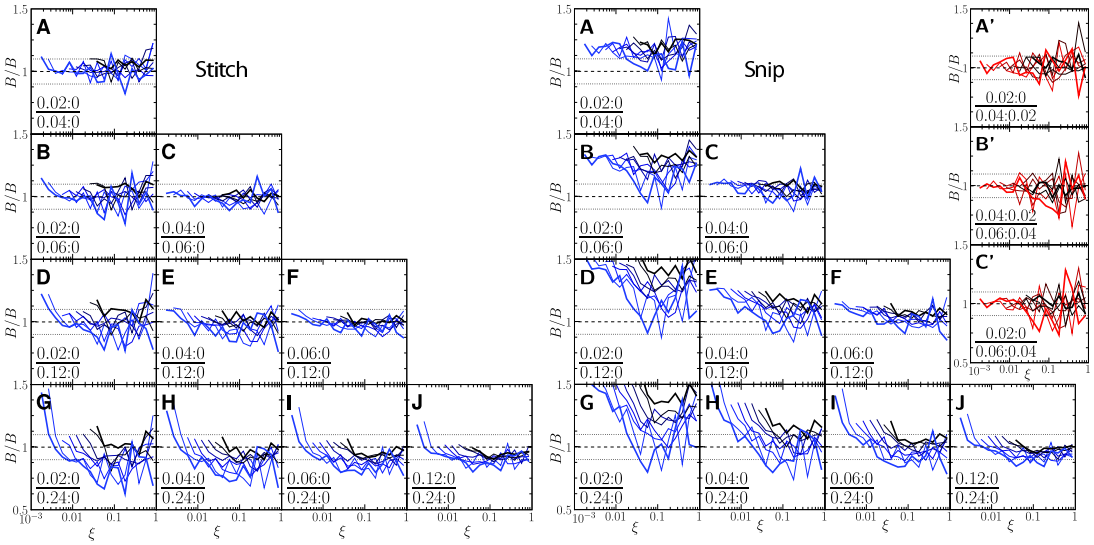

Figure 1.10: Δz Convergence Matrix (stitching left, snipping right). Each subplot is the ratio of B/n for two different catalogues (labelled). The dashed lines denote equality and the dotted lines are the 10% deviation levels. The ratios are presented for a variety of mass bins with the high mass bin highlighted in thick blue (or red) and the low mass bin highlighted in thick black.

have remerged within three outputs (see Fig. 1.3), 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 $\sim 25\%$, similar in magnitude but opposite in sign to the difference between snipping vs 3-stitching shown in Fig. 1.9. The fitting form in equation (1.14) works well for ∞-stitching, where the best fit 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 §1.5.2) and excellent mass conservation properties (§1.5.4) but alters the FOF mass function by a few percent.

Since the snipping algorithm tends to inflate the merger rate and the ∞-stitching algorithm tends to under-estimate it, we believe the 3-stitching used in all the results in Section 1.4 should be a fairly robust scheme.

1.5.2 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. 1.10 shows the ratios of B/n for five pairs of progenitor and descendant redshifts: (z_P, z_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 panel), there is excellent convergence for $\Delta z \lesssim 0.12$ (panels A-F), where the ratios of B/n are centred around 1 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. 1.8. 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 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 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% level) centred around (0.06:0). To emphasise 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 1.2.3 and Fig. 1.2). 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 spacing: (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. 1.11 shows $d\bar{N}_{\rm merge}/dz$ 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 panel) and worse Δz-

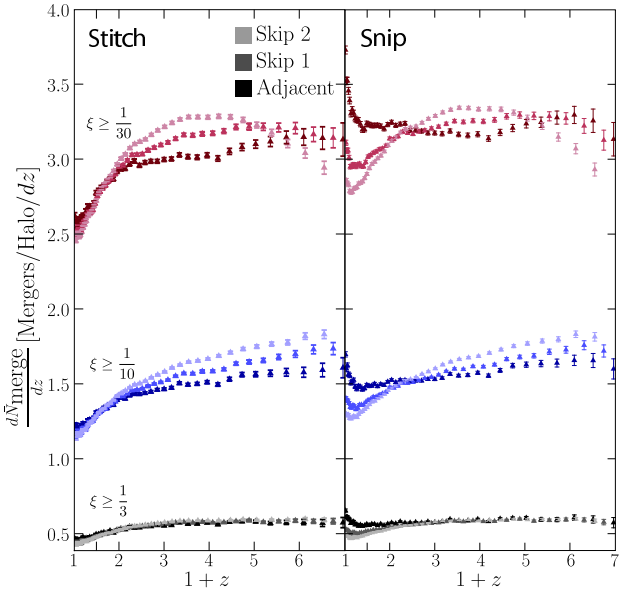

Figure 1.11: Merger rate $d\bar{N}_{\rm merge}/dz$ computed using three types of redshift spacings: *adjacent*, *skip 1* and *skip 2* (see text). For clarity, only the galaxy-mass haloes are shown; the group and cluster haloes behave similarly.

convergence for the snipping method (right panel). The latter follows the behaviour seen in Fig 1.10, with adjacent spacing ($\Delta z = 0.02$ at $z = 0$) over-predicting the merger rate.

At higher redshifts ($z \gtrsim 1.5$), Fig. 1.11 shows that the merger rates in the minor merger regime differ by up to $\sim 15\%$ 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. 1.3 shows, the ratio of fragmentation to merger events is 40% near $z = 0$ but drops to $\lesssim$10% 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 ~ 0.5 at $z \sim 6$ vs $\Delta z = 0.02$ at $z \approx 0$) and the inaccuracy of the multiple counting ordering assumption for large Δz (see Section 1.5.3). At high z, we therefore advocate using the finest output spacings available in the Millennium database, noting the good time convergence for major mergers ($\xi \gtrsim 1/3$) but $\sim 15\%$ variations in the minor merger rates.

1.5.3 Multiple vs Binary Counting

As discussed in Section 1.3.1, 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. 1.12 compares the merger rates (per halo), $B(M_0, \xi)/n$, for these two counting methods (dashed: multiple; solid: binary) as a function

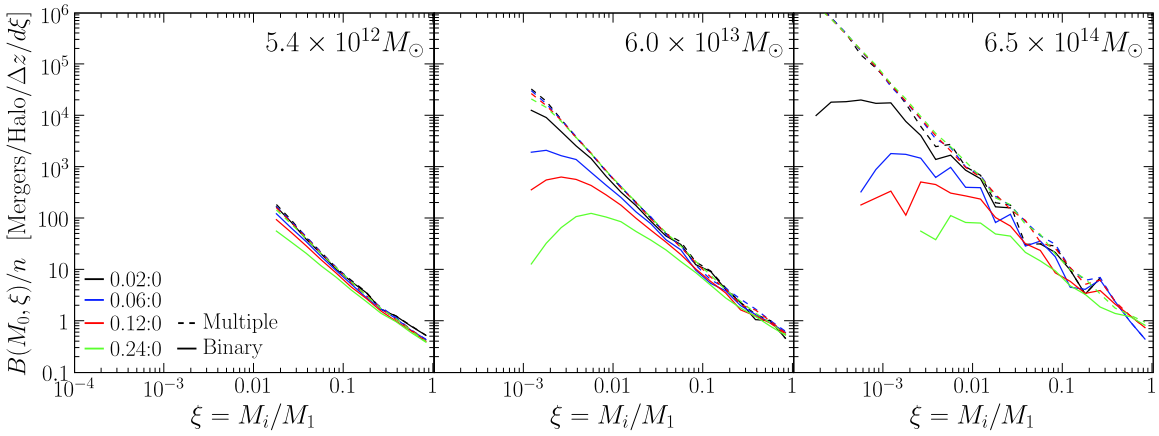

Figure 1.12: Comparisons of merger rate per halo, B/n, computed via multiple counting (dashed lines) and binary counting (solid lines) for 4 merger trees with increasing Δz. Three descendant mass bins are shown (from left to right). We find the multiple counting rate to be in excellent agreement regardless of Δz of the tree, whereas the binary counting B/n curves fall off from the observed power-law behaviour towards lower ξ.

of the progenitor mass ratio ξ for three descendant masses M_0 (increasing from left to right). 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. 1.12 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 1.5.2). 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 under-estimates the minor-merger regime of the multiple counting result.

Fig. 1.12 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 $6 \times 10^{13} M_\odot$ 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 1.3.1. There, we assumed that the less massive progenitors $M_2, M_3, ...$ 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 onto 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 database.) 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 $\lesssim 10\%$. Most of the mergers among multiple progenitors, therefore, do occur between the most massive progenitor and a less massive progenitor, as we have assumed.

1.5.4 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:

$$M_0 = \sum_{i=1}^{N_p} M_i + \Delta M \,, \tag{1.15}$$

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

Fig. 1.13 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_{FOF} and M_{200}. The distribution shows a prominent peak at $\Delta M/M_0 \sim 2.5$ % for both M_{FOF} and M_{200}, indicating that in the majority of the merger events between $z = 0.06$ and 0.0, $\sim 97.5\%$ of the mass of the descendant halo comes from resolvable progenitor haloes.

Even though the two distributions in Fig. 1.13 have similar peaks, the M_{200} mass definition produces longer $\Delta M/M_0$ tails than the M_{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_{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 1.2.1, we have been using the M_{FOF} mass thus far.

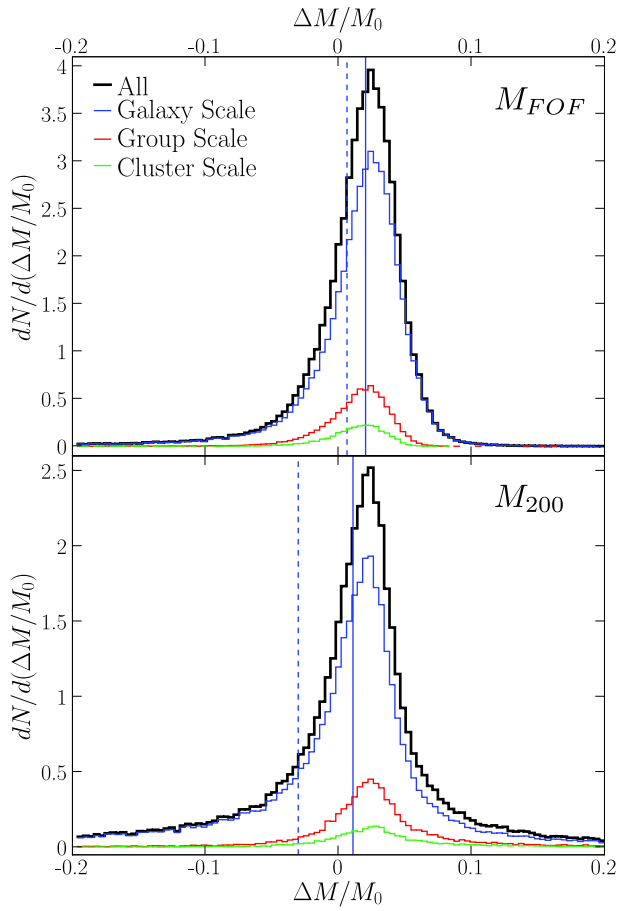

Figure 1.13: Distributions of $\Delta M/M_0$ from the $z = 0.06$:0 Millennium merger tree (using stitching; snipping is nearly identical) computed using the M_{FOF} mass (top panel) vs M_{200} virial mass (bottom panel). The solid vertical line is the median of the distribution for the galaxy-mass bin and the dashed line is the mean. We note a longer negative ΔM tail for the M_{200} tree when compared to the M_{FOF} tree. Note, however, that the peaks of the two distributions are in good agreement ($\Delta M/M_0 \sim 2.5\%$)

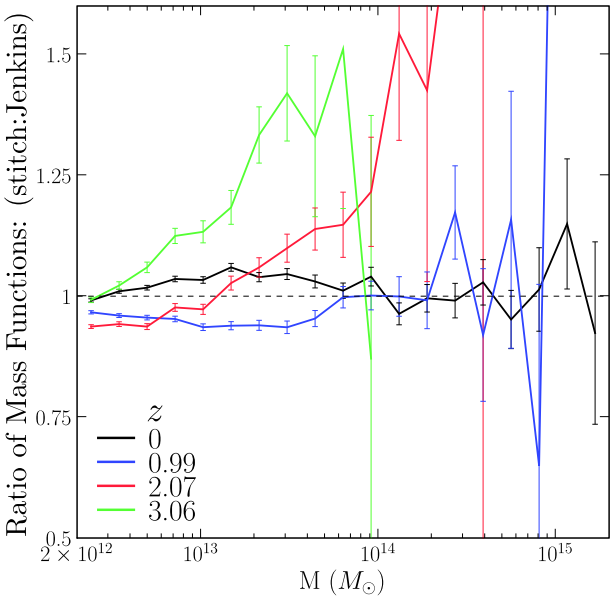

Figure 1.14: Ratios of the Millennium halo mass function (computed from the stitching trees) to the fit of Jenkins et al. [2001] using the M_{FOF} mass. The results for the snipping trees are virtually identical. We note a significant deviation of $\sim 25\%$ at $z \sim 3$ between the Jenkins fit and the Millennium mass function.

Our main results on merger rates in Section 1.4 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.

1.5.5 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. 1.2, 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 $\sim 0.25\%$) so will use the stitching result below.

Fig. 1.14 shows the ratio of the Millennium halo mass function to the fits by Jenkins et al. [2001] for M_{FOF} at four redshifts: $z \approx 0$, 1, 2, and 3. The fit of Jenkins et al is accurate to better than 10% for low redshift ($z \lesssim 1$), but it underestimates the Millennium halo abundance by $\gtrsim 25\%$ 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].

1.6 Theoretical Models for Merger Rates

1.6.1 Extended Press-Schechter Model

In Section 1.3.3 we discussed how our merger rates are related to the conditional probabilities in the EPS model and obtained equation (1.12), where there are two choices for the definition of the progenitor mass M' since the EPS model is not symmetric in the two progenitor masses. In Fig. 1.15 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 (1.12) 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 and Hu [1999].

Fig. 1.15 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-fit β is -2.17 and -2.01 for the snipping and stitching methods, respectively (see Table 1.2). 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. 1.15. 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. 1.15 compares the rates at $z = 0$. At higher redshifts, we find the Millennium merger rate to evolve as $\propto (d\delta_c/dz)^{\eta}$, where $\eta \approx 0.37$ (see equation [1.14] and Table 1.2) and is shallower than the EPS prediction of $\eta = 1$ in equation (1.12). 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. 1.15 will maintain the same shape at higher z, but the amplitude of the ratio will increase. For instance, the ratio shown in Fig. 1.15 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. 1.15 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 et al., 2008b], 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.

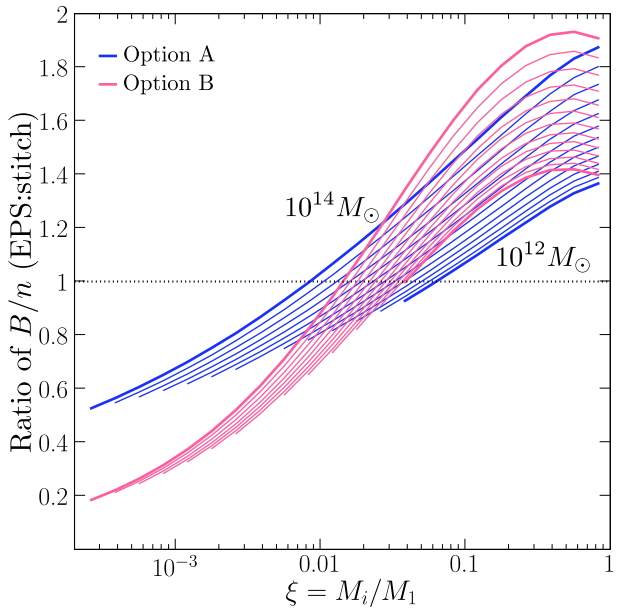

Figure 1.15: Comparison between our Millennium merger rate (from the fit) and the two predictions of the Extended Press-Schechter model. The ratio of B/n from EPS to Millennium is plotted. Blue and red label the two options in assigning progenitor masses in the EPS model (see text); within each colour, the set of curves from bottom to top denotes increasing M_0 bins, from $\sim 10^{12} M_\odot$ 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.

1.6.2 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

$$\begin{aligned}\frac{dn(M)}{dt} = \frac{1}{2}\int_0^M A(M', M{-}M')n(M')n(M{-}M')dM' \\ - \int_0^\infty A(M, M')n(M')n(M)dM' \,,\end{aligned} \tag{1.16}$$

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

$$A(M, M') \leftrightarrow \frac{B(M, M')}{n(M)n(M')} \,. \tag{1.17}$$

We note, however, that the coagulation equation in the form of equation (1.16) 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 (1.16) using equation (1.17) may not yield robust results.

Assuming that $n(M)$ is the Press-Schechter mass function, Benson et al. [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 regularisation 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. 1.16 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.

1.7 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

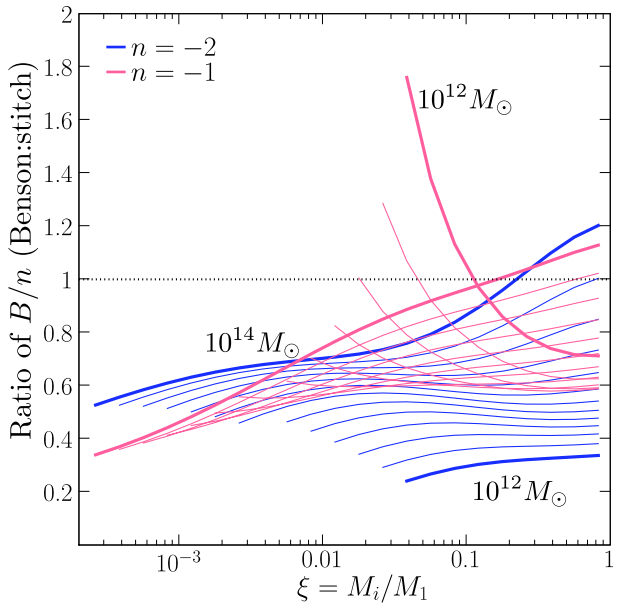

Figure 1.16: Same as Fig. 1.15, only now comparing the merger rates from Benson et al. (2005) to the Millennium B/n for initial power spectrum index $n = -2$ (blue) and $n = -1$ (red).

that we constructed from the halo catalogue of the Millennium simulation. Our main results are presented in Figs. 1.6 to 1.8, 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. 1.6 right panel and Fig. 1.7). 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. 1.8), but when expressed in units of per redshift, the merger rate depends very weakly on z (bottom panel of Fig. 1.8). 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. 1.6). These simple behaviours have allowed us to propose a universal fitting formula in equation (1.14) that is valid for $10^{12} \leq M_0 \lesssim 10^{15} M_\odot$, $\xi \gtrsim 10^{-3}$, and up to $z \sim 6$.

Throughout the paper we have emphasised 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. 1.10 and 1.11). 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. 1.13 and 1.14).

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% (depending on halo mass) and the minor

merger rates to be too low by up to a factor of 2-5 (Fig. 1.15). 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 (1.16). Equation (1.16) 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 $\xi \geq 1/3$ 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 timescale 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 timescale, 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 and White, 2008] finds galaxy growth via major mergers to depend strongly on stellar mass, where mergers are more important in the buildup 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 ($(1+z)^{n_m}$ with $n_m \sim 2 - 2.3$) 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, for example, Burkey et al. 1994, Carlberg et al. 1994, Yee and Ellingson 1995, Woods et al. 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 $n_m < 1$. 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).

1.8 Appendix: 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 [Bower et al., 2006, Harker et al., 2006, Helly et al., 2003].

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 database, 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% 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% 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 database 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. 1.17 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 $\sim 30\%$), 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. 1.18 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 comparison (thin dotted curves). The Durham mass function is systematically lower: the number of $z = 0$ haloes with $M \gtrsim 10^{14} M_\odot$ is $\sim 25\%$ lower, and the difference increases at $z \sim 1$, affecting the halo mass function even at $M \sim 2 \times 10^{12} M_\odot$.

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% of their original mass. This condition may indeed remove spurious FOF linkings in

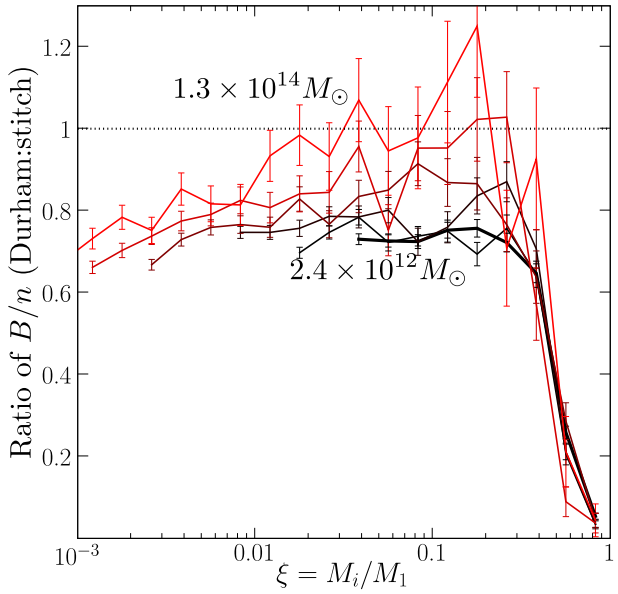

Figure 1.17: The ratio of the Durham merger rate B/n to our stitching rate B/n (section 1.4.3) as a function of progenitor mass ratio ξ for a number of descendant mass bins ranging from $\sim 2 \times 10^{12} M_\odot$ (black) to $\sim 10^{14} M_\odot$ (red). The Durham merger rate tends to be lower than the stitching merger rate, and suffers a sudden drop in the major merger regime ($\xi \gtrsim 0.3$).

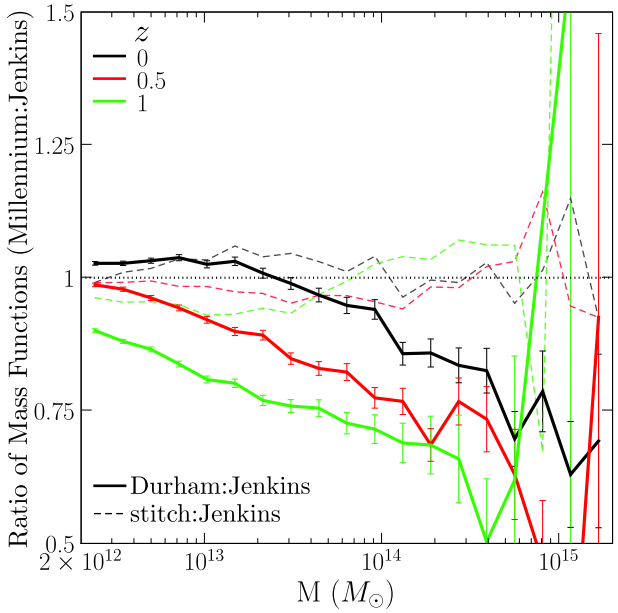

Figure 1.18: The ratio of the Durham halo mass function to the fit of Jenkins et al. (2001) at redshifts 0, 0.5, and 1 (thick solid curves with error bars). The ratio of the halo mass function from our stitching method to the same fit is shown for comparison (thin dashed curves). The Durham algorithm tends to reduce the masses of massive haloes, leading to a deficit that grows to $\sim 50\%$ at $\sim 10^{15} M_{\odot}$ and at higher z.

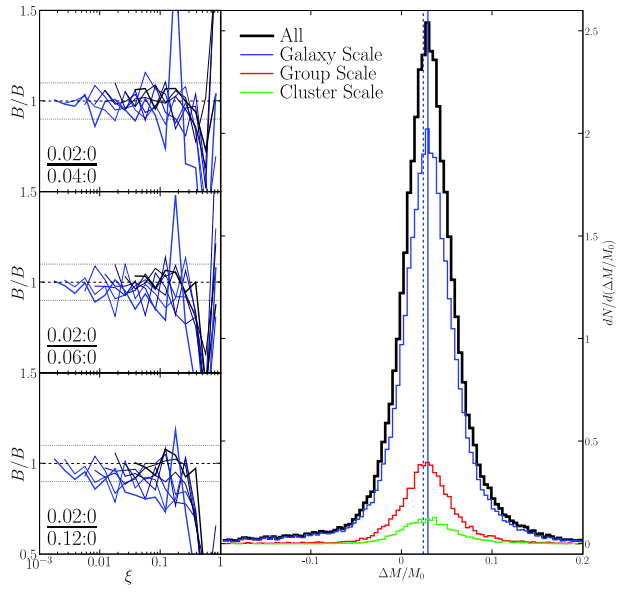

Figure 1.19: Left panels: A subset of the Δz convergence matrix presented in Fig. 1.10, now computed using the Durham tree. Note the poor convergence properties of the major merger end ($\xi \gtrsim 0.3$). This corresponds to the region of the largest difference between the stitching and Durham merger rates (see Fig. 1.17). Right panel: The distribution of $\Delta M/M_0$ for the $z = 0.06{:}0$ Durham catalogue (similar to Fig. 1.13).

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. 1.19 (right panel) shows that the Durham tree has similar mass conservation properties as our stitching tree in Fig. 1.13. The distribution of the mass in the 'diffuse' component, $\Delta M/M_0$, has a very similar peak of $\sim 3\%$, although the negative ΔM events have been suppressed. For Δz convergence (left panels; cf section 1.5.2), 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.

Chapter 2

The Merger Rates and Mass Assembly Histories of Dark Matter Haloes in the Two Millennium Simulations

Fakhouri, Onsi; Ma, Chung-Pei ; Boylan-Kolchin, Michael — June 2010
Monthly Notices of the Royal Astronomical Society

We construct merger trees of dark matter haloes and quantify their merger rates and mass growth rates using the joint dataset from the Millennium and Millennium-II simulations. The finer resolution of the Millennium-II Simulation has allowed us to extend our earlier analysis of halo merger statistics to an unprecedentedly wide range of descendant halo mass ($10^{10} \lesssim M_0 \lesssim 10^{15} M_\odot$), progenitor mass ratio ($10^{-5} \lesssim \xi \leq 1$), and redshift ($0 \leq z \lesssim 15$). We update our earlier fitting form for the mean merger rate per halo as a function of M_0, ξ, and z. The overall behavior of this quantity is unchanged: the rate per unit redshift is nearly independent of z out to $z \sim 15$; the dependence on halo mass is weak ($\propto M_0^{0.13}$); and it is nearly a power law in the progenitor mass ratio ($\propto \xi^{-2}$). We also present a simple and accurate fitting formula for the mean mass growth rate of haloes as a function of mass and redshift. This mean rate is $46 M_\odot\ yr^{-1}$ for $10^{12} M_\odot$ haloes at $z = 0$, and it increases with mass as $\propto M^{1.1}$ and with redshift as $(1+z)^{2.5}$ (for $z \gtrsim 1$). When the fit for the mean mass growth rate is integrated over a halo's history, we find excellent match to the mean mass assembly histories of the simulated haloes. By combining merger rates and mass assembly histories, we present results for the number of mergers over a halo's history and the statistics of the redshift of the last major merger.

2.1 Introduction

Mergers of dark matter haloes are intimately connected to a wide array of phenomena in the now-standard ΛCDM cosmology. In addition to being the dominant channel for mass growth of haloes themselves, mergers are also responsible for the growth of stellar mass in galaxies, both directly via galaxy-galaxy mergers, and indirectly via the accretion of potentially star-forming gas. Furthermore, mergers help shape many important observational properties of galaxies, e.g., star formation rates, color and morphology transformations, dynamical states of stellar disks, and galaxy mass and luminosity functions. Mergers are also responsible for the existence of satellite galaxies such as dwarf spheroidals in the Milky Way and non-cD galaxies in galaxy clusters. Quantifying the rate of halo-halo mergers, and its possible dependence on factors such as halo mass, mass ratio, and time, is therefore of great interest for a theoretical understanding of galaxy formation and its connections to observations.

In a series of papers, we have examined various aspects of the growth of dark matter haloes. In Fakhouri and Ma [2008], we computed the merger rates of dark matter haloes from the Millennium Simulation [Springel et al., 2005] and presented a simple algebraic fitting form for our results. The resolution and size of this simulation allowed us to determine the merger rate over the parameter range of $10^{12} \lesssim M_0 \lesssim 10^{15} M_\odot$ for the mass of the descendant haloes, $10^{-3} \lesssim \xi \leq 1$ for the mass ratio of the progenitor haloes, and $0 \leq z \lesssim 6$ for the redshift. The detailed environmental dependence of the merger rates and halo mass growths was analyzed in two subsequent papers [Fakhouri and Ma, 2009, 2010]. In McBride et al. [2009], we studied the statistics of the halo mass assembly histories and mass growth rates in the Millennium Simulation. Halo mergers have also been studied in a handful of papers by others (e.g., Governato et al. 1999, Gottlöber et al. 2001, Berrier et al. 2006, Maller et al. 2006, Guo and White 2008, Genel et al. 2009, Stewart et al. 2009). The pre-2008 studies were all limited to small simulations that mainly investigated major mergers in a narrow mass range at low redshift (typically $z \lesssim 1$). Some such studies have emphasized potential challenges for hierarchical structure formation; for instance, Stewart et al. [2008] have noted that the frequency of major mergers among Milky-Way sized haloes poses a problem for thin-disk survivability. Much work has also been done in quantifying halo mass accretion and assembly histories using N-body simulations that are smaller than the Millennium runs (e.g., Lacey and Cole 1994, Tormen et al. 1997, Tormen 1998, Wechsler et al. 2002, van den Bosch 2002, Li et al. 2007, Zhao et al. 2009, except Cole et al. 2008).

In this paper, we extend the results presented in Fakhouri and Ma [2008] and McBride et al. [2009] by incorporating the Millennium-II Simulation [Boylan-Kolchin et al., 2009b]. This simulation has the same number of particles as the Millennium Simulation but has 125 times better mass resolution. This new database provides 7.5×10^6 dark matter haloes (each containing more than 1000 simulation particles) between redshift 0 and 15 and their subhalo merger trees for our analysis. Adding to the 11.3×10^6 haloes (between $z = 0$ and 6) available from the Millennium Simulation, this combined dataset allows us to determine the dark matter halo merger rates and mass growth rates from $z = 0$ to up to $z = 15$, for over

five orders of magnitude in the descendant halo mass ($10^{10} \lesssim M_0 \lesssim 10^{15} M_\odot$) and progenitor mass ratio ($10^{-5} \lesssim \xi \leq 1$).

This paper is organized as follows. Section 2.2 describes the dark matter haloes in the Millennium and Millennium-II simulations, and how we construct the merger trees and quantify the merger statistics and mass accretion histories of the haloes. In Section 2.3, we present results for three types of statistics: merger rates at $z = 0$ up to ~ 15 for halo mass $\sim 10^{10}$ to $10^{15} M_\odot$ (§ 2.3.1); the rate at which the haloes are accreting dark matter across the virial radii, and the mass growth history of haloes (§ 2.3.2); and the cumulative merger statistics over a halo's past history, e.g., the mean cumulative number of mergers of a given mass ratio experienced as a function of z and halo mass, and the distribution of the redshift at which the last major merger occurred for haloes at various mass and redshift (§ 2.3.3). The Appendix contains a detailed comparison of the three types of algorithms that we have tested for handling the fragmentation events in a merger tree of FOF haloes [Fakhouri and Ma, 2008, 2010]. They are named "snip," "stitch," and "split," depending on whether the fragmented subhalo was ignored, stitched back to the original FOF halo in subsequent outputs, or split off from the FOF at earlier times. A quantitative assessment of the systematic differences in the merger rates derived from each algorithm is provided in the Appendix.

The cosmology used throughout this paper is identical to that used in the the Millennium simulations: a ΛCDM 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 $\sigma_8 = 0.9$. Masses and lengths are quoted in units of $\mathrm{M}_\odot$ and Mpc without the Hubble parameter h.

2.2 Construction of Halo Merger Trees

2.2.1 The Two Millennium Simulations

The Millennium and Millennium-II simulations are large N-body simulations of cosmological structure formation using the concordance ΛCDM cosmological parameters listed at the end of Section 2.1. The simulations are described in detail in Springel et al. [2005] (Millennium) and Boylan-Kolchin et al. [2009b] (Millennium-II); here we summarize some basic features of the simulations and of the default post-processing procedures that result in subhalo merger trees.

Both simulations follow the evolution of $2160^3 \approx 10^{10}$ particles from redshift 127 to redshift 0 using versions of the `GADGET` tree-PM code [Springel et al., 2001b, Springel, 2005]. The simulations differ in spatial scale and mass resolution: the Millennium Simulation uses a box size of $L = 685\,\mathrm{Mpc}$ and a Plummer-equivalent force softening that is a factor of 10^5 smaller, $\epsilon = 6.85\,\mathrm{kpc}$, with a particle mass of $m_p = 1.18 \times 10^9\,\mathrm{M}_\odot$. The Millennium-II Simulation uses $L = 137\,\mathrm{Mpc}$ and $\epsilon = 1.37\,\mathrm{kpc}$, both of which are a factor of 5 smaller than the values from the Millennium Simulation; the particle mass is therefore 125 times smaller, $m_p = 9.43 \times 10^6\,\mathrm{M}_\odot$. The two simulations have 60 outputs at identical redshifts between

$z \approx 20$ and $z = 0$, spaced approximately equally in $\log z$, as well as additional snapshots (4 for the Millennium, 8 for the Millennium-II) at higher redshifts.

Subhalo merger trees are constructed in an identical fashion for the Millennium and Millennium-II simulations. Dark matter haloes are first identified at each snapshot using a Friends-of-Friends group-finder (FOF; Davis et al. 1985) with a linking length of 0.2 times the mean interparticle separation. All FOF groups with at least 20 particles are stored. The `SUBFIND` algorithm [Springel et al., 2001a] is then applied to each FOF group to identify halo substructure. `SUBFIND` identifies local density maxima and performs an unbinding procedure to determine which particles in the FOF group are bound to each density peak. Substructures with at least 20 particles after unbinding are stored, resulting in a list of subhaloes (SHs) associated with each FOF group in the simulation. Note that some FOF groups do not contain 20 self-bound particles and therefore not every FOF group contains a subhalo, while some FOF groups can contain many self-bound density peaks and therefore have many subhaloes.

These subhaloes are then linked across simulation snapshots to produce subhalo merger trees. This linking is done by establishing a unique descendant for each subhalo in the following manner. First, all particles in the subhalo are rank-ordered by binding energy. A list of candidate descendants – all subhaloes at the subsequent snapshot containing at least one particle from the subhalo in question – is built and a figure of merit is computed for each descendant. The candidate descendant with the highest score is assigned as the actual descendant. The figure of merit for candidate descendants is simply a weighted sum of the rank-ordering of the subhalo's particles; this procedure ensures that the tightly bound center of a subhalo is weighted more heavily than the less-bound outer regions even if the center is subdominant in terms of mass.

In addition to searching for a descendant at the subsequent output, a search is also performed two snapshots later. This additional step accounts for subhaloes that are temporarily unresolved when passing near the center of a more massive system but re-appear later. On occasion, no descendant can be identified at either of the two subsequent snapshots, in which case, the subhalo is not assigned a descendant at all but rather is considered destroyed. With subhaloes and their unique descendants identified, subhalo merger trees are built by linking subhaloes and their descendants: all subhaloes with a common descendant at $z = 0$ are linked to all subhaloes sharing these subhaloes as descendants, and so on. A given subhalo merger tree thus contains all subhaloes that can be linked via their descendants to one specific subhalo at $z = 0$. The trees link 760 million subhaloes for the Millennium Simulation and 590 million subhaloes for the Millennium-II Simulation. For the central subhalo of a $z = 0$ galaxy-mass halo ($M \approx 10^{12}\,\mathrm{M}_\odot$), its subhalo merger tree typically consists of 90 subhaloes in the Millennium Simulation and 2800 subhaloes in the Millennium-II Simulation.

2.2.2 Halo Fragmentation

In this paper, as in our previous work [Fakhouri and Ma, 2008, McBride et al., 2009], our focus is on the merger and assembly histories of FOF haloes. To do this we must

z_d	Sim	$10^{10-11}M_\odot$		$10^{11-12}M_\odot$		$10^{12-13}M_\odot$		$10^{13-14}M_\odot$		$>10^{14}M_\odot$		Total
		$N_p=1$	$N_p\geq 2$	1	≥ 2	1	≥ 2	1	≥ 2	1	≥ 2	
0.06	M	0	0	0	0	321489	90922	14504	45281	3	4817	477016
	MII	214045	25292	12583	17107	13	3279	0	486	0	36	272841
0.51	M	0	0	0	0	306142	98664	11442	39469	0	2757	458474
	MII	224865	29170	12422	18405	7	3199	0	421	0	20	288509
1.08	M	0	0	0	0	236280	137729	4197	32349	0	976	411531
	MII	220703	49811	8473	23221	1	2985	0	316	0	8	305518
2.07	M	0	0	0	0	126926	133965	629	12746	0	73	274339
	MII	202572	80435	4772	24874	2	2128	0	121	0	0	314904

Table 2.1: The number of merger events in the two Millennium simulations at four representative redshifts ($z \approx 0$, 0.5, 1, and 2). At each z, we list the number of descendant FOF haloes that have a single progenitor halo ($N_p = 1$, i.e., no mergers) and multiple progenitors ($N_p \geq 2$), for five descendant mass bins (left to right). The descendant mass here refers to the halo mass at the redshift listed rather than at the present day. Only haloes containing more particles than our minimum cutoff (1000 for descendants; 40 for progenitors) are counted. The higher-resolution Millennium-II Simulation dominates the contribution to the merger statistics of $M_0 \lesssim 10^{12}M_\odot$ haloes, while the larger-volume Millennium Simulation dominates the contribution to cluster-mass haloes.

first construct merger trees of the FOF haloes from the underlying subhalo trees described in Section 2.2.1. Such construction is nontrivial due to halo fragmentations: subhaloes of a progenitor FOF halo may have descendants that reside in more than one FOF halo. Sometimes this is due to a physical unbinding event in which a subhalo formerly bound to an FOF is ejected out of the FOF system. Sometimes the fragmentation is spurious – a subhalo may oscillate in and out of the FOF group before finally settling in. Sometimes the FOF algorithm incorrectly groups subhaloes that are unbound but only happen to pass by one another and should not be associated as a single FOF group.

We presented detailed comparisons in Fakhouri and Ma [2008, 2010] of three types of algorithms – snip, stitch, and split – for handling these fragmentation events. In the Appendix we summarize these algorithms and quantify the systematic differences in the merger rates derived from each algorithm.

For the main results presented in § 2.3 below, we use the split-3 algorithm, in which the subhalo fragments that pop out of an FOF halo are either snipped or split depending on a simple criterion. The fragmented subhalo is snipped if it is observed to remain in the FOF halo for all 3 snapshots immediately preceding the fragmentation event; in this case the ancestral link between fragment and FOF is severed. If the fragmented subhalo is not in the FOF halo for all 3 preceding snapshots, it is interpreted as distinct and is split off from the FOF. The split-3 algorithm generally gives very similar results to the stitch-3 algorithm used in Fakhouri and Ma [2008], e.g., the two methods produce merger rates that agree to within 10% for the redshifts and mass ranges that we have statistics for. The only exception is in the major merger regime for low-mass haloes at low redshift ($z \lesssim 1$), where split-3 is lower than stitch-3 by up to 30% (see Fig. 2.10 in Appendix). Overall, split-3 appears slightly more robust at handling spurious FOF linking events in this regime (also see Fakhouri and

Ma 2010, Genel et al. 2009). As discussed in the Appendix, however, the exact definition of what constitutes a merger may be situation-dependent, meaning that no single method is perfect in all cases.

2.2.3 Extracting Merger Rates and Mass Accretion Histories

From the merger trees of FOF haloes obtained by applying a given fragmentation algorithm, we extract a merger catalog. Each catalog provides us with a list of descendant FOF haloes at redshift $z_d \geq 0$ with mass M_0, and for each descendant halo, its set of N_p FOF progenitors at $z_p = z_d + \Delta z$, where N_p can range from 1 (i.e. a single progenitor) to a large number, depending on the halo mass and the value of Δz. We label the rank-ordered progenitor mass with M_i, $i \in (1, 2, \ldots, N_p)$, and $M_1 \geq M_2 \geq \ldots M_{N_p}$. To ensure that only numerically resolved haloes are included in our study, we impose a minimum of 1000 particles for the descendant haloes and 40 particles for the progenitor haloes. For the Millennium Simulation, this criterion corresponds to a minimum halo mass of $1.2 \times 10^{12}\,\mathrm{M}_\odot$ for the descendant and $4.7 \times 10^{10}\,\mathrm{M}_\odot$ for the progenitor. For Millennium-II, the minimum masses are $9.4 \times 10^{9}\,\mathrm{M}_\odot$ and $3.8 \times 10^{8}\,\mathrm{M}_\odot$ for the descendant and progenitor haloes. We emphasize that the mass of a descendant halo refers to its mass at a given redshift z_d and not its ultimate mass at $z = 0$ (unless $z_d = 0$).

We compute the merger rates at redshift z as a function of descendant mass M_0 and progenitor mass ratio $\xi = M_i/M_1$ (for $i > 1$). We define $B(M_0, \xi, z)$ to be the number of mergers per Mpc^3, dM_0, $d\xi$, and Δz with mass $M_0 \pm dM_0/2$ and mass ratio $\xi \pm d\xi/2$. As discussed in Fakhouri and Ma [2008], we find the mean merger rate *per halo*, $B(M_0, \xi, z)/n(M_0, z) \equiv dN_m/d\xi/dz$, where $n(M_0, z)$ is the number density of haloes, to have a particularly simple dependence on the merger parameters. This rate, when expressed in per redshift units, is a dimensionless quantity that gives the mean number of mergers per halo per unit z per unit ξ. To avoid artificial boundary effects at $z = 0$, we use the two outputs at $z = 0.12$ and 0.06 to compute the $z \sim 0$ merger rate.

To compute the mass accretion histories and accretion rates of haloes, we start with a given descendant FOF halo at some redshift and identify the mass of its most massive progenitor at an earlier snapshot. This process is iterated backwards in time to construct the main branch of the descendant's merger tree. The mass trajectory along the main branch of a descendant gives us its mass accretion history $M(z)$ (see. e.g., Lacey and Cole 1993), from which we can compute the mass accretion rate $\dot{M}$ as a function of z. Note that the progenitor halo on the main branch of a descendant halo at a given snapshot need not be the most massive progenitor of that descendant at that snapshot.

2.2.4 Definitions of Halo Mass

In our prior analysis of the Millennium halo merger rate [Fakhouri and Ma, 2008], we assigned the halo mass using the standard FOF mass M_{FOF}. This mass is simply proportional to the number of particles assigned to each FOF halo by the FOF group finder. An alternative

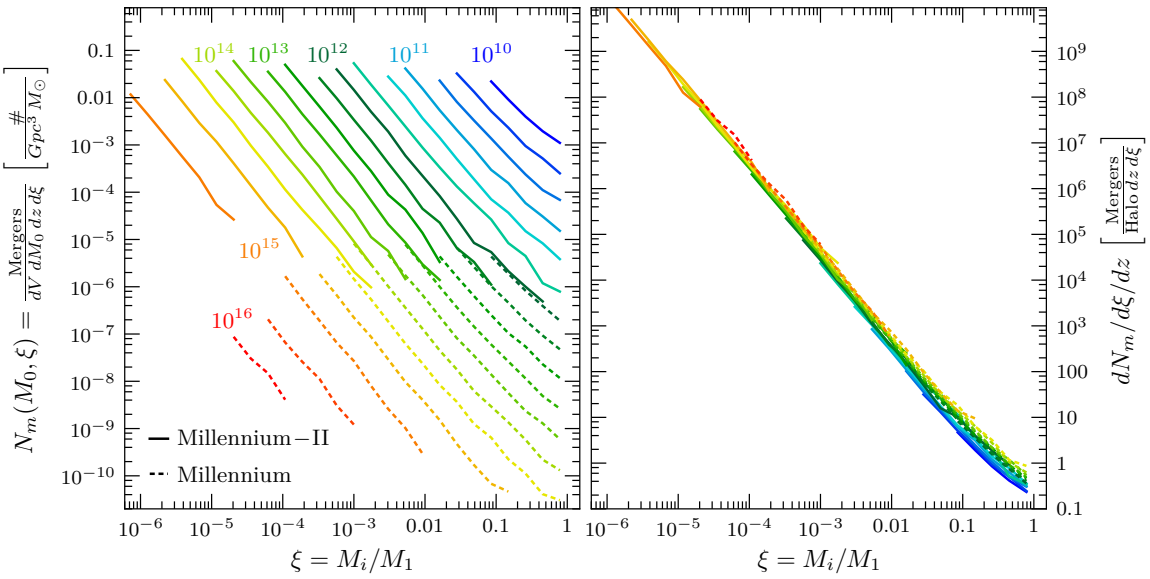

Figure 2.1: Left panel: The mean merger rate of $z = 0$ FOF haloes, $B(M_0, \xi)$, as a function of the mass ratio of the progenitors (ξ) and the descendant halo mass (M_0) over 6 orders of magnitude: 10^{10} to $10^{16} M_\odot$ from right (blue) to left (red). The Millennium-II results are shown in solid, while the results from the Millennium are in dashed curves. Right panel: The mean merger rate *per halo*, $B(M_0, \xi)/n(M_0) \equiv dN_m/d\xi/dz$. Normalizing $B(M_0, \xi)$ by the halo number density $n(M_0)$ collapses the curves in the left panel to nearly a single curve, indicating that $dN_m/d\xi/dz$ is nearly independent of M_0 and has a simple universal form.

definition that we will use throughout this paper is the sum of the masses of an FOF's subhaloes, M_{SH}. This definition has been shown recently by Genel et al. [2009] to be more robust than M_{FOF} since the SUBFIND algorithm assigns only gravitationally bound particles to each subhalo.

Overall, we find the halo mass functions computed using these two mass definitions to differ at the 5% level at all halo masses. This difference can be caused by a slight excess of mass in FOF haloes due to unbound or spuriously linked particles, as well as by a slight deficit in M_{SH} when SUBFIND does not account for all the mass physically associated with a subhalo. When restricted to the subset of haloes that are undergoing very minor mergers, however, Genel et al. [2009] noted that the FOF mass of the smaller progenitor increases as it approaches the more massive progenitor. For minor mergers involving mass ratios as low as $\xi \sim 0.001$, the ratio M_{FOF}/M_{SH} for the *smaller* progenitor can rise from 1.03 to 1.5 prior to mergers. We will therefore use M_{SH} for halo masses in this study. We note that this discrepancy occurs only for the small subset of low-mass haloes that are in the process of merging onto a much larger halo; its effect on the total halo mass function is therefore limited to $\sim 5\%$.

2.3 Results

To provide a sense for the halo statistics and merger events available from the two Millennium simulations, we list in Table 2.1 the number of descendant haloes (above 1000 particles) and their progenitors (above 40 particles) at four representative redshifts for five broad mass ranges from 10^{10} to $> 10^{14} M_{\odot}$. The results presented below are based on these events and those at other redshifts.

2.3.1 Merger Rates

Present-Day Merger Rates

The left panel of Fig. 2.1 presents $B(M_0, \xi, z = 0)$, the $z = 0$ mean number of mergers per unit volume, descedant mass M_0, mass ratioξ, and redshift as a function of progenitor mass ratio ξ from the two Millennium simulations (solid for Millennium II; dashed for Millennium). The colored curves correspond to different mass bins ranging from $10^{10}\,\mathrm{M}_{\odot}$ (blue) to $10^{16}\,\mathrm{M}_{\odot}$ (red). The rates are determined from the $z = 0.06$ and 0.12 merger tree catalogue since, as described in Section 2.2.3, we would like to avoid the $z = 0.0$ snapshot due to the boundary effects that interfere with the post-processing algorithms used to handle the halo fragmentation events. The split-3 algorithm is used here; other algorithms yield qualitatively similar agreement between the two simulations (see Appendix for details).

The right panel of Fig. 2.1 shows the mean merger rate per halo, $B/n = dN_m/d\xi/dz$, where each of the curves in the left panel has been divided by the number density of haloes in that mass bin. The collapse of the curves to nearly a single curve shows that the per halo merger rate $dN_m/d\xi/dz$ is nearly independent of halo mass. This collapse is similar to that

seen in Fig. 6 of Fakhouri and Ma [2008] for the Millennium Simulation. A comparison of the two figures helps illustrate the large dynamic range achieved when the two Millennium simulations are combined: the halo mass range has been increased by two orders of magnitude in Fig. 2.1, and for each mass bin, the progenitor mass ratio is extended downward by also a factor of ~ 100, reaching $\xi \sim 10^{-6}$ for $M_0 = 10^{15}\,\mathrm{M}_\odot$.

The overlap in the merger parameter space between the two simulations is seen to be fairly small in Fig. 2.1. The two simulations are therefore quite complementary: Millennium II allows us to probe descendant and progenitor masses that are a factor of 125 smaller than Millennium, whereas the larger box of the Millennium Simulation provides robust statistics for the rare events that are poorly sampled in Millennium II, e.g., major mergers of massive haloes (i.e. the lower right corner of left panel of Fig. 2.1). Over the small region of overlap, Fig. 2.1 shows good agreements between the merger rates determined from the two simulations: both the power-law dependence on ξ and the weak dependence on M_0 carry over from Millennium to Millennium II. Boylan-Kolchin et al. [2009b] show that many other quantities, such as halo mass functions, formation times, and subhalo abundances, have a much wider range of overlap and that the two simulations are in excellent agreement for these quantities as well.

The weak dependence of the merger rate on M_0 is shown explicitly in Fig. 2.2. Each curve here shows the mean rate per halo, $dN_m/d\xi/dz$, integrated over different ranges of $\xi \geq \xi_{\rm min}$, where $\xi_{\rm min} = 0.3, 0.10.01$, and 10^{-3} (from bottom up). Major mergers with mass ratio within 1:3 (bottom curve) are clearly much more rare than minor mergers (top curves), but all the curves have very similar power-law dependence on M_0. Over about 4.5 orders of magnitude in M_0, the rate increases by only a factor of ~ 3, suggesting that the merger rate scales roughly as $\sim M_0^{0.1}$. A more accurate fit is provided in Sec 2.3.1 below.

$z > 0$ Merger Rates

The Millennium Simulation provided sufficient halo statistics for us to determine the halo merger rates up to $z \sim 6$ in our previous study. The higher mass resolution of the Millennium-II Simulation now allows us to probe redshifts up to ~ 15. The combined results from the two simulations are shown in Fig. 2.3, which plots the mean merger rate per unit redshift (left panel), dN_m/dz, and per unit time (right panel), dN_m/dt, as a function of redshift. These merger rates have been integrated over different ranges of $\xi \geq \xi_{\rm min}$, ranging from major mergers with $\xi_{\rm min} = 0.3$ (solid curves at bottom), to extreme minor mergers with $\xi_{\rm min} = 10^{-5}$ (dotted curve at top). Within each line type, the colors indicate different descendant mass bins ranging from 10^{10} (blue) to $> 10^{14}\mathrm{M}_\odot$ (red). Only the higher mass bins are plotted as $\xi_{\rm min}$ is lowered. This is because minor mergers of low-mass haloes fall below the mass resolution limit.

Fig. 2.3 indicates that the general trends reported in Fig. 8 of Fakhouri and Ma [2008] continue to hold in the Millennium-II Simulation. The dimensionless rate dN_m/dz is remarkably independent of redshift up to $z \sim 15$, whereas the rate per Gyr, dN_m/dt, rises with increasing z because a unit redshift corresponds to a shorter time interval at higher z. This

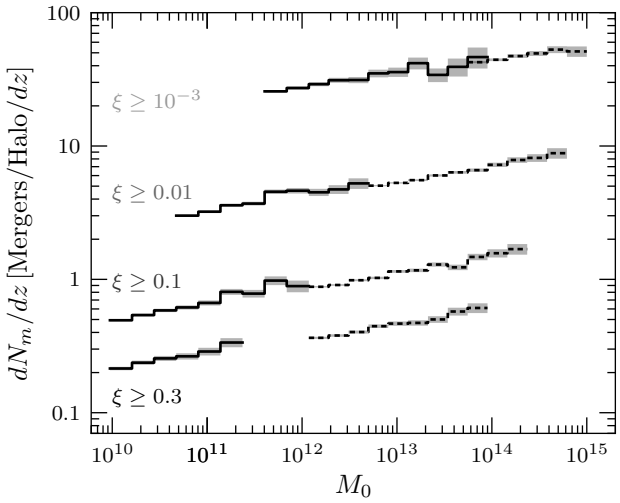

Figure 2.2: The $z \approx 0$ mean merger rate per halo (per unit z), dN_m/dz, as a function of descendant mass, M_0, for four ranges of progenitor mass ratio ξ. The Millennium-II results are shown in solid, while the original Millennium results are the dashed curves. The upper curves include increasingly more minor mergers. The mass dependence is weak over five orders of magnitude in mass and is well approximated by a power law $\propto M_0^{0.133}$.

redshift dependence is similar to that obtained by Guo and White [2008] for the dimensionless growth rates due to mergers of both haloes and galaxies (based on semi-analytic models) from the Millennium Simulation.

Merger Rate Fitting Forms

Since the merger statistics in Figs. 2.1-2.3 are very consistent between the two Millennium simulations, we use an analytical form similar to equation (12) of Fakhouri and Ma [2008] to fit the dimensionless mean merger rate $dN_m/d\xi/dz$ (in units of mergers per halo per unit redshift per unit ξ) from the combined Millennium dataset. An appealingly simple feature of this fitting form is that it is separable in the variables M_0, ξ, and z:

$$\frac{dN_m}{d\xi dz}(M,\xi,z) = A\left(\frac{M}{10^{12}M_\odot}\right)^{\alpha} \xi^{\beta} \exp\left[\left(\frac{\xi}{\tilde{\xi}}\right)^{\gamma}\right](1+z)^{\eta}\,. \tag{2.1}$$

We find the best-fit parameters to be $(\alpha,\beta,\gamma,\eta) = (0.133, -1.995, 0.263, 0.0993)$ and $(A,\tilde{\xi}) = (0.0104, 9.72\times10^{-3})$. The near z-independence in the left panel of Fig. 2.3 is more striking than in our 2008 study due to the larger coverage in redshift here. In view of this lack of z-dependence, we choose to use the simpler factor of $(1+z)^{\eta}$ here rather than the growth rate of the density field used in Fakhouri and Ma [2008]. In comparison to our 2008 study, the power-law slope of the mass dependence is slightly steeper here ($\alpha = 0.133$ vs. 0.089), whereas the power-law slope of the ξ dependence is slightly shallower here ($\beta = -1.995$ vs.

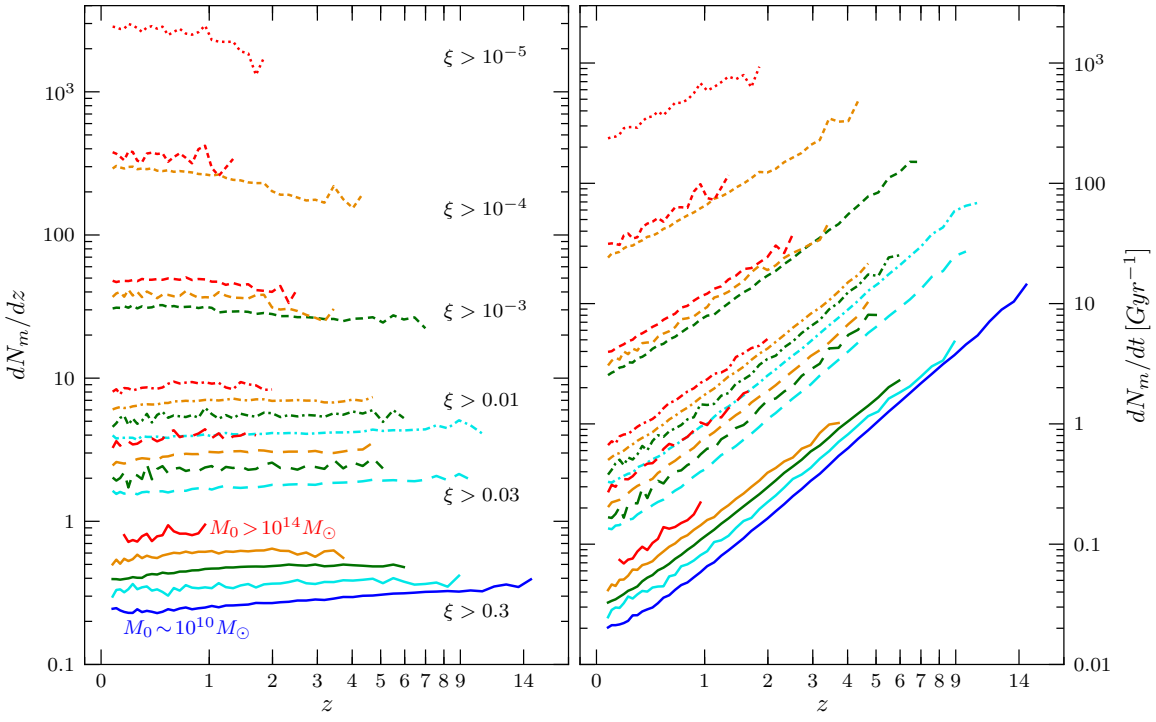

Figure 2.3: Time evolution of the mean halo merger rates *per halo* in units of per redshift, dN_m/dz (left panel), and in units of per Gyr, dN_m/dt (right panel) from the two Millennium simulations. The descendant mass M_0 and progenitor mass ratio ξ over five orders of magnitude are plotted. The weak dependence of the rates on M_0 is shown by the different colors: $\approx 10^{10}$ (blue), 10^{11} (cyan), 10^{12} (green), 10^{13} (orange). to $> 10^{14}\mathrm{M}_\odot$ (red). The line types denote different types of mergers, ranging from major mergers (solid) to extreme minor mergers (dotted) The rate dN_m/dz on the left is remarkably constant out to $z \sim 15$; the rapid rise of dN_m/dt with increasing z on the right is therefore largely due to the cosmological factor dt/dz, which spans a shorter time per unit z with increasing z.

-2.17). These differences are primarily due to the differing definitions of halo mass used in the two studies (FOF vs. sum of subhalos; see Sec. 2.4) and the refinements in our stitch-3 algorithm (see the Appendix).

We note that the left panel of Fig. 2.3 does show mild variations in the redshift dependence among the different ξ bins: the rate increases slightly with increasing z for major mergers, while it declines somewhat for the very minor mergers ($\xi_{\rm min} \sim 10^{-4}$ to 10^{-5}). Since this variation is so minor and the minor merger regime is more prone to numerical resolution issues, we have opted for simplicity rather than a more complicated fitting form.

2.3.2 Mass Growth Rates and Assembly Histories

In the last section, we presented results for the instantaneous rates of halo mergers as a function of redshift, descendant mass, and progenitor mass ratio. Here, we examine a related set of statistics that quantify the mass growth of haloes. These two quantities are clearly related since mergers are a primary process for haloes to gain mass, but mergers are not the only process. As discussed at length in Fakhouri and Ma [2010], "diffuse" accretion of unresolved haloes or dark matter particles also makes an important contribution to halo growth. In the mass assembly history of a halo, mergers with other haloes typically result in more discrete but less frequent changes in the halo mass, while diffuse accretion leads to a more continuous change.

Mass Accretion Rates

To compute the total mass growth rate of a halo of a given mass M_0 at time t, we follow the main branch of its merger tree (see § 2.2.3) and set $\dot{M} = (M_0 - M_1)/\Delta t$, where M_0 is the descendant mass at time t and M_1 is the mass of its most massive progenitor at time $t - \Delta t$. The mean value of $\dot{M}$ as a function of z for the complete set of resolved haloes in the two Millennium simulations is plotted in Fig. 2.4 (solid curves). Nine ranges of M_0 spanning five orders of magnitude ($10^{10}\,{\rm M}_\odot$ to $10^{15}\,{\rm M}_\odot$ from bottom up) are shown. Fig. 2.4 can be compared directly to Fig. 5 of McBride et al. [2009] for the Millennium Simulation alone. The rising $\langle \dot{M} \rangle$ with increasing redshift in our earlier study is seen to continue to $z \sim 14$, and the nearly linear scaling of $\langle \dot{M} \rangle$ with halo mass is extended down to $\sim 10^{10} M_\odot$.

We find the mass accretion rates shown in Fig. 2.4 to be very well fit by the forms given by equations (8) and (9) of McBride et al. [2009]. The coefficients quoted there only need minor adjustments after the Millennium-II results are added. We suggest the following updated

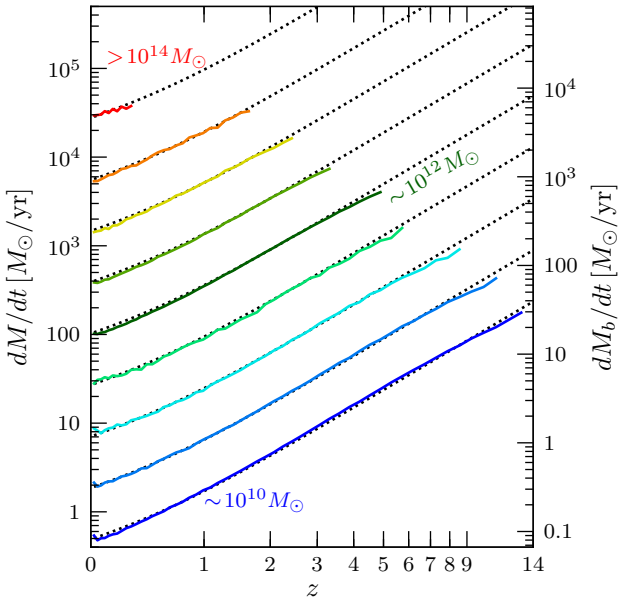

Figure 2.4: Mean mass accretion rate of dark matter onto haloes as a function of redshift from the two Millennium simulations (solid curves). Halo masses ranging from $10^{10} M_\odot$ to $> 10^{14} M_\odot$ are plotted. The dashed curves show the accurate fit provided by equation (2.2). The right-hand side of the vertical axis labels the mean accretion rate of baryons, M_b, assuming a cosmological baryon-to-dark matter ratio of 1/6.

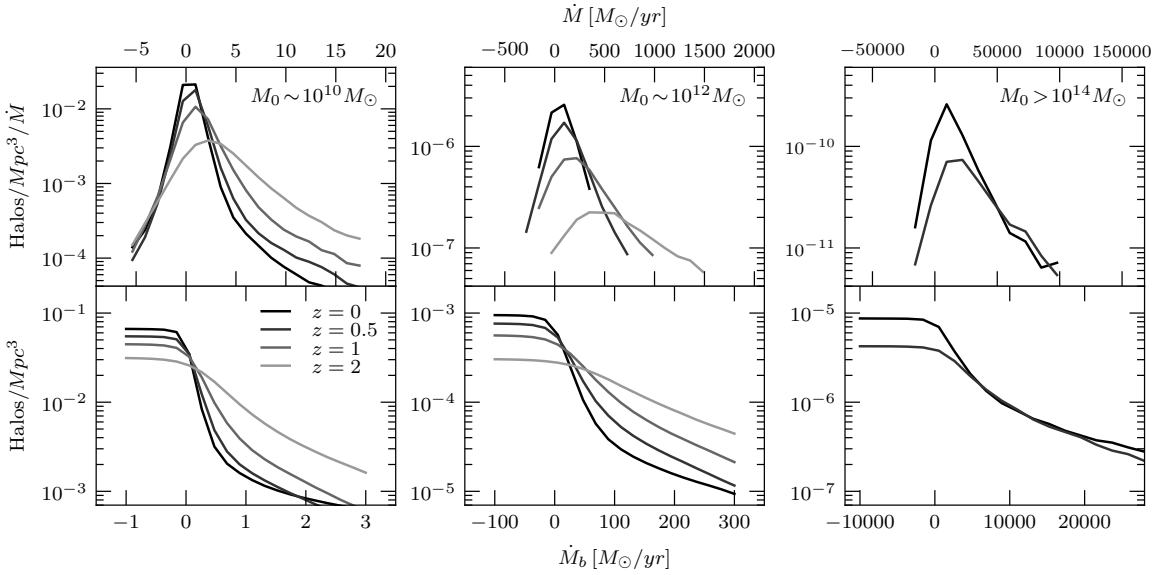

Figure 2.5: Differential (top) and cumulative (bottom) distributions of the baryonic accretion rates, $\dot{M}_b$, for halo masses 10^{10} (left), 10^{12} (middle), and $> 10^{14} M_\odot$ (right). Within each panel, the accretion rates at $z = 0, 0.5, 1$, and 2 are shown (except the right panel, where such massive haloes are present only at $z = 0$ and 0.5). The distributions are seen to broaden significantly with increasing z. The vertical axis in each of the bottom panels labels the number of haloes per comoving Mpc^3 that are accreting at a rate of $\dot{M}_b$ or above.

fits to the mean and median mass growth rates of haloes of mass M at redshift z:

$$\begin{aligned}
\langle \dot{M} \rangle_{\rm mean} &= 46.1\ {\rm M}_\odot {\rm yr}^{-1} \left(\frac{M}{10^{12}\,{\rm M}_\odot} \right)^{1.1} \\
&\quad \times (1+1.11z)\sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda} \\
\langle \dot{M} \rangle_{\rm median} &= 25.3\ {\rm M}_\odot {\rm yr}^{-1} \left(\frac{M}{10^{12}\,{\rm M}_\odot} \right)^{1.1} \\
&\quad \times (1+1.65z)\sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda}\,. \qquad (2.2)
\end{aligned}$$

At a given mass and redshift, the mean rate is overall higher than the median rate since the distribution of $\dot{M}$ has a long positive tail (see Fig. 2.5). The dashed curves in Fig. 2.4 illustrate the remarkable accuracy of this formula in matching the simulation results over the broad ranges of halo mass and redshift shown.

The right-hand-side label along the vertical axis of Fig. 2.4 shows the corresponding mean accretion rate of baryons, $\dot{M}_b$, where we have assumed a cosmological baryon-to-dark matter ratio of $\Omega_b/\Omega_m = 1/6$. These values are meant to provide a rough approximation for the mean rate at which baryons are being accreted near the virial radius of a dark matter halo. Most of these baryons are presumably in the form of warm or hot ionized hydrogen gas that is being channeled into the haloes along cosmic filaments, and various gas cooling and feedback processes will likely affect the baryon accretion rate. Many studies on galaxy formation are aimed at quantifying these physical processes under which these baryons are cooled to form neutral gas, molecular gas, and stars, and the feedback processes that heat up the baryons and lead to large-scale outflows.

In Fig. 2.5 we show the differential (top) and cumulative (bottom) *distributions* of the baryonic accretion rate for three halo masses (left to right panels) and four redshifts. A cosmic ratio of $\Omega_b/\Omega_m = 1/6$ is again assumed to convert the dark matter rate into a baryonic rate. The distributions are strongly peaked at the mean values presented in Fig. 2.4 but exhibit long tails towards high positive values due to major merger events and towards negative values due to tidal stripping and halo fragmentation. Not only is the mean accretion rate higher at higher z, the distribution of $\dot{M}_b$ is also broader at higher z. For example, the comoving density of Milky Way-mass haloes that are accreting baryons at a rate of at least $100 M_\odot$ per year is approximately ten times greater at $z = 2$ ($\sim 2 \times 10^{-4}$ Mpc^{-3}) than at $z = 0$ ($\sim 3 \times 10^{-5}$ Mpc^{-3}).

Mass Assembly Histories

The top panel of Fig. 2.6 shows the mean mass assembly history $M(z)$ for nine bins of M_0 (at $z = 0$) from 10^{10} to $> 10^{14}$ (from top to bottom). The solid curves show the results obtained from the main branch (i.e. the most massive progenitor) along the merger tree for all the $z = 0$ haloes in the two Millennium simulations. The dotted curves show the $M(z)$ obtained from integrating the fitting formula for the mean $\dot{M}$ in equation (2) from the present-day to some redshift z. The agreement is generally very good, in particular at

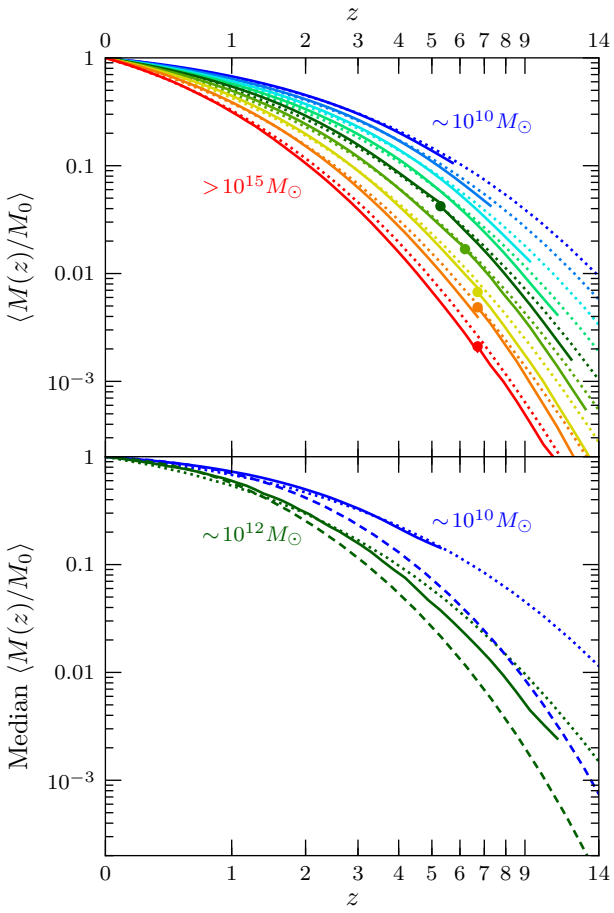

Figure 2.6: Top panel: *Mean* mass assembly history $M(z)$ of all $z = 0$ resolved dark matter haloes in the two Millennium simulations (solid curves). Nine ranges of halo mass from $10^{10} M_\odot$ (top blue) to $10^{15} M_\odot$ (bottom red) are plotted. The dotted curves show the predictions given by integrating the mean $\dot{M}$ of our fitting formula (eq. 2.2). The lower four mass bins contain only haloes from the Millennium-II Simulation. For the upper five mass bins in which the haloes are drawn from both simulations, we use a solid circle to label the redshift above which only Millennium-II haloes contribute since the Millennium Simulation can no longer resolve haloes at such high z. The relatively smooth connection at the circle illustrates the consistency between the two simulations. Lower panel: *Median* mass assembly history $M(z)$. For clarity, only two mass bins are plotted. Solid lines are from the simulations, dotted lines from the integration of the *mean* $\dot{M}$ from eq. (2.2), and dashed lines show the fits from Zhao et al. (2009).

$z \lesssim 8$. A perfect agreement is not expected because the two quantities, $\dot{M}$ and $M(z)$, are not determined from the same set of haloes in the simulations: the $\dot{M}$ statistics are obtained from all haloes of a given mass M at a given z, whereas the $M(z)$ curves show only the mean mass of the most massive progenitors at redshift z for the $z = 0$ haloes of mass M, which are a small subset of the haloes of the same mass that are present at z in the simulation boxes. In view of this difference, the agreement between the solid and dotted curves in the top panel of Fig. 2.6 is in fact quite remarkable. Over the large range of mass and redshift shown in Fig. 2.6, we have checked that the direct fits for the mean $M(z)$ proposed in recent literature (e.g. Boylan-Kolchin et al. 2009a, McBride et al. 2009) provide a good match at low z, but integrating the fit for $\langle \dot{M} \rangle$ in equation (2.2) provides a closer match at high z,

The solid curves in the bottom panel of Fig. 2.6 show the median, rather than the mean, mass assembly history obtained from the simulations for two mass bins centered at $M_0 = 10^{10} M_\odot$ and $10^{12} M_\odot$. We note that integrating our fit to the *median* $\dot{M}$ does not yield the median $M(z)$ because unlike the *mean* $\dot{M}$, the median and derivative operations do not commute. The median and mean $M(z)$ are sufficiently similar, however, that we find integrating our *mean* $\dot{M}$ to yield relatively good agreement with the median $M(z)$ (dotted curves). For comparison, the fit of Zhao et al. [2009] to the median $M(z)$ is shown as dashed curves. Their fit appears to be systematically lower than the Millennium results at $z > 1$.

In principle, we can integrate the (mass-weighted) halo merger rate in equation (1) and obtain the portion of the dark matter accretion rate $\langle \dot{M} \rangle$ in equation (2) that is due to mergers. As emphasized in Fakhouri and Ma [2010], however, accretion of "diffuse" material (consisting of unresolved haloes and tidally stripped mass) also makes a non-negligible contribution to $\langle \dot{M} \rangle$; equation (2) therefore can not be obtained solely from equation (1).

2.3.3 Merger Statistics over a Halo's History

In the last two sections we quantified the halo merger rates, the mass growth rates, and the assembly histories of haloes. These quantities can be combined to predict a number of additional useful merger statistics over a halo's history.

Cumulative number of mergers

One such statistic is $N_m(\xi_{\rm min}, M_0, z_0, z)$, the total number of mergers that a halo of mass M_0 at redshift z_0 has encountered between z_0 and an earlier z. The mergers can be characterized by major or minor mergers by imposing a limit of $\xi_{\rm min}$ on the progenitor mass ratio (evaluated at the redshift of the merger). These numbers are essential for making theoretical predictions of galaxy properties that are impacted by mergers, e.g., the dynamics and stability of stellar disks, the star formation rate, and the color and morphology transformation due to mergers.

Fig. 2.7 shows $N_m(\xi_{\rm min}, M_0, z_0, z)$ for the complete set of resolved haloes at $z_0 = 0$ (left), 1 (middle), and 2 (right) from the two Millennium simulations. In each panel, five ranges of

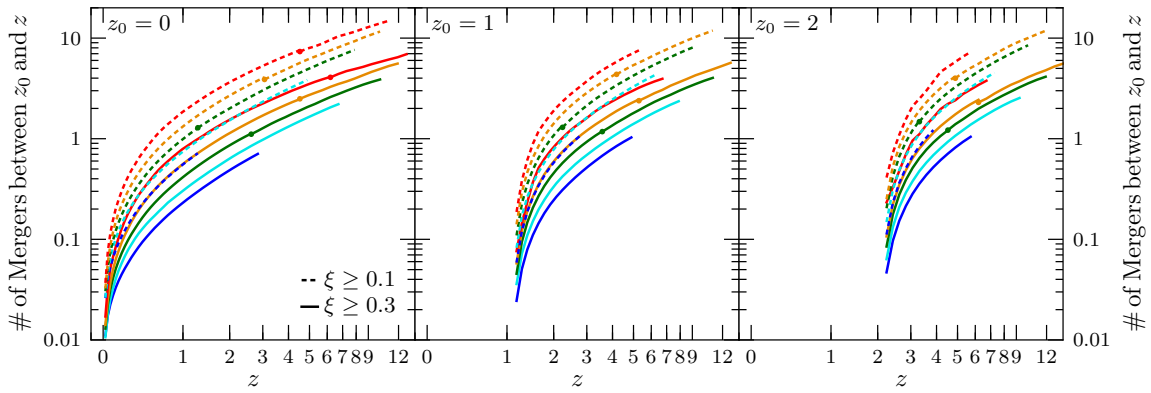

Figure 2.7: Mean number of mergers between redshifts z_0 and z experienced by a halo at $z_0 = 0$ (left), 1 (middle), and 2 (right) from the joint dataset of the two Millennium simulations. In each panel, the solid and dashed curves represent mergers with a progenitor mass ratio (defined at the time of merger) of $\xi \geq 0.3$ and $\xi \geq 0.1$. For each mass ratio cutoff, five ranges of halo mass are shown (from bottom up): 10^{10} (blue), 10^{11} (cyan), 10^{12} (green), 10^{13} (orange), and $> 10^{14} M_{\odot}$ (red). The lower-mass haloes are from the Millennium-II Simulation, whereas the cluster mass haloes are mainly from the Millennium Simulation. For mass bins in which the haloes are drawn from both simulations, we use a solid circle to label the redshift above which only Millennium-II haloes contribute since the Millennium can no longer resolve haloes at such high z. The fact that the curves connect quite smoothly are another indication of the consistency between the two simulations.

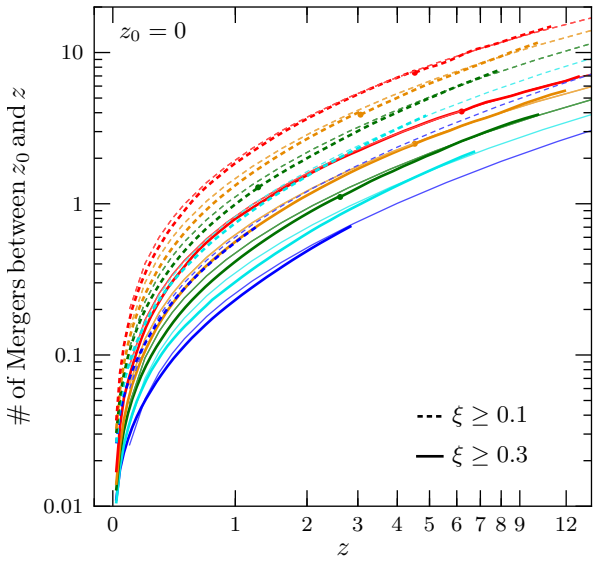

Figure 2.8: Same as the left panel of Fig. 2.7, with the addition of the predictions (thin curves) computed from eq. (2.3). The agreement with the simulation results (thick curves) is excellent, suggesting that eq. (2.3) can be used to make analytic predictions for merger statistics over a halo's history.

M_0 are plotted for redshift up to 12. Major mergers with $\xi \geq 0.3$ are shown by solid curves, while the more minor mergers with $\xi \geq 0.1$ are shown in dashed curves.

Fig. 2.7 shows that the mean trend of the number of mergers experienced over a halo's lifetime is a sensitive function of the halo mass and merger mass ratio. haloes of Milky-Way mass at the present day (green curves) have on average experienced one major merger event ($\xi \geq 0.3$) per halo since $z \approx 2.3$, and one merger with $\xi \geq 0.1$ per halo since $z \approx 1$. When extended to $z \approx 7$, these haloes have on average encountered ~ 3 mergers with $\xi \geq 0.3$, and ~ 7 mergers with $\xi \geq 0.1$. The formation redshifts as well as the last merger epoch for more massive haloes are both lower, a well-known trend in CDM-based cosmology (see, e.g., Lacey and Cole 1993, 1994). Cluster-sized haloes with $M_0 \sim 10^{14} M_\odot$, for instance, have on average experienced one major merger event since $z \approx 1.2$, and one merger with $\xi \geq 0.1$ since $z \approx 0.6$.

It is possible to compute the cumulative number of mergers, $N_m(\xi_{\rm min}, M_0, z_0, z)$, shown in Fig. 2.7 from the fitting formula for the merger rate $dN_m/d\xi/dz$ in equation (2.1) and the mass accretion history $M(z)$ obtained by integrating $\langle \dot{M} \rangle$ in equation (2.2). Specifically, these quantities are related by

$$N_m(\xi_{\rm min}, M_0, z_0, z) = \int_{z_0}^{z} dz \int_{\xi_{\rm min}}^{1} d\xi \frac{dN_m}{d\xi dz} \left(M(z), \xi, z\right) . \qquad (2.3)$$

Since we are interested in the number of mergers over a halo's past history, we must take

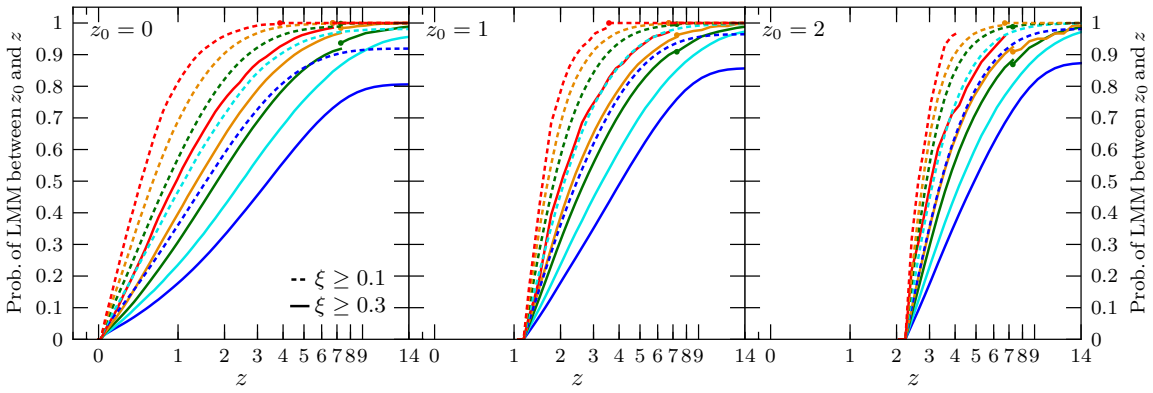

Figure 2.9: Cumulative distribution of the redshift at which the last (i.e. most recent) major merger occurred in a halo's past history for haloes at $z_0 = 0$, 1, and 2 (left to right) in the two Millennium simulations. The vertical axis gives the probability that the last major merger occurred between z_0 and redshift z. The curves are defined the same way as in Fig. 2.7.

into account the fact that a halo's mass generally decreases with increasing z, and that the merger rate depends on the halo mass (albeit weakly). The merger rate $dN_m/d\xi/dz$ at redshift z in the integrand above therefore should be evaluated using the mean mass $M(z)$ that a halo of mass M_0 at z_0 had at the earlier z. The results are shown in Fig. 2.8, which is identical to the left panel of Fig. 2.7 except that we have added the theoretical curves (thin curves) for comparison. The agreement with the simulation results (thick curves) is excellent, suggesting that equation (2.3) can be used to make analytic predictions for merger statistics over a halo's history.

Redshift of last major merger

The redshift at which each curve in Fig. 2.7 crosses one merger event along the vertical axis is a useful quantity since it gives the mean redshift at which a halo has experienced its last major merger (LMM). The LMM redshift of a halo is closely related to its formation redshift and may be linked to the time at which the associated galaxy last experienced prominent star formation activity and morphological changes. To analyze this quantity further, we show in Fig. 2.9 the *distribution* of the LMM redshift for haloes at $z_0 = 0$, 1, and 2 (from left to right). Within each panel, five halo masses and two ranges of ξ are plotted. The vertical axis gives the probability that a halo at a given z_0 has had a last major merger between z_0 and z.

Useful merger statistics can be read off from Fig. 2.9. For instance, 50% of present-day haloes have had a major merger ($\xi \geq 0.3$) since $z \approx 1$, 1.8, and 3.4 for halo mass $10^{14}, 10^{12}$, and $10^{10} M_\odot$, respectively. When more minor mergers with $\xi \geq 0.1$ are considered, the median redshift of the last merger is lowered to 0.4, 0.8, and 1.6 for the three masses. The assembly history of Milky Way-size haloes is of particular interest; see Boylan-Kolchin et al.

[2009a] and references therein for a detailed statistical study of this topic. For haloes of $M_0 \sim 10^{12} M_\odot$ today, the left panel of Fig. 2.9 shows that $\sim 31\%$, 53%, and 69% of these haloes have experienced a major merger since $z = 1$, 2, and 3, respectively. For haloes of $\sim 10^{12} M_\odot$ at $z_0 = 1$ (middle panel), about 50% of them have had a major merger since $z \approx 2.7$, and for haloes of the same mass at $z_0 = 2$ (right panel), about 50% of them have had a major merger since $z \approx 3.7$.

2.4 Summary and Conclusions

We have combined the halo catalogs from the two Millennium simulations to form an unprecedentedly large dataset for studying the merger statistics and assembly histories of dark matter haloes in the ΛCDM cosmology. The two simulations provide, respectively, 11.3×10^6 haloes (between redshift 0 and 6) and 7.5×10^6 haloes (between redshift 0 and 15) above 1000 particles for our study. These haloes and their merger trees have allowed us to determine the dark matter halo merger rates and mass growth rates from $z = 0$ to up to $z = 15$, for over five orders of magnitude in the descendant halo mass ($10^{10} \lesssim M_0 \lesssim 10^{15} M_\odot$) and progenitor mass ratio ($10^{-5} \lesssim \xi \leq 1$). For the small range of overlapping parameter space between the two simulations, we have found the agreement to be excellent.

For the merger rates, the basic features reported in our earlier study based on the Millennium Simulation alone [Fakhouri and Ma, 2008] are largely preserved in the Millennium-II Simulation. The mean merger rate *per halo*, $dN_m/d\xi/dz$, is nearly independent of the descendant mass (Fig. 2.1 and 2.2) and scales as $\propto M_0^{0.133}$ at all redshifts. The merger rate in units of per redshift is nearly independent of redshift out to $z \sim 15$ (left panel of Fig. 2.3); the rate in units of per Gyr is therefore largely determined by the cosmological factor of dt/dz and increases roughly as $(1+z)^{2.5}$ at $z \gtrsim 1$ (right panel of Fig. 2.3). Equation (2.1) provides an update on our simple analytical fitting form for the merger rate as a function of M_0, ξ, and z.

For the mass growth rates of individual haloes, we have found the mean and median statistics (Fig. 2.4) to be well approximated by the simple fitting form of our earlier study [McBride et al., 2009]. The updated coefficients based on the joint dataset from the two Millennium simulations are given by equation (2.2). The present-day mean and median rates at which a $10^{12} M_\odot$ dark matter halo is accreting mass (at the virial radii) are 46.1 and 25.3 $M_\odot$ yr^{-1}, respectively. This rate increases nearly linearly with the halo mass, and increases with redshift approximately as $(1+z)^{1.5}$ at low z and $(1+z)^{2.5}$ at $z \gtrsim 1$.

We have also presented statistical quantities that track the merger histories of dark matter haloes cumulatively. Fig. 2.7 presents the number of major mergers experienced by haloes of various mass between redshift z_0 and z for $z_0 = 0, 1$, and 2. Fig. 2.9 presents the probability that a dark matter halo at redshift z_0 will have last experienced a major merger at some earlier redshift z. Much interesting and useful information regarding the contribution to halo growth made by major mergers can be read off these figures with ease.

With the addition of results from the Millennium-II Simulation to our previous analysis of

the Millennium Simulation, the merger rate of dark matter haloes is now well-quantified for haloes with masses between 10^{10} and $10^{15}\,\mathrm{M}_{\odot}$ for redshifts $z \lesssim 15$, modulo the uncertainties inherent in halo definitions and in algorithms for handling fragmentation (see Appendix), for the cosmology used in the Millennium simulations. Several avenues remain open for future work, however.

One obvious extension of the results in this paper is to consider the mergers of *subhaloes* themselves, as subhalo mergers can be more directly linked to galaxy mergers than can FOF halo mergers [Angulo et al., 2009, Wetzel et al., 2009]. Furthermore, the structure of the merger trees produced for the Millennium simulations lends itself naturally to computing subhalo merger properties. While computing subhalo merger rates and connecting them to galaxy mergers presents additional challenges – in particular, the issues of assigning stellar masses to subhaloes, numerical resolution effects, and subhalo identification within larger FOF haloes – a thorough theoretical understanding of such rates is essential for disentangling the relative contributions of merging and star formation to the growth of galaxies.

2.5 Appendix: Comparison of Different FOF Merger Trees

We refer the reader to Section 5 and Figure 8 of Fakhouri and Ma [2009] for a detailed discussion of the three basic operations – "snip," "stitch," and "snip" – that we have implemented and tested for handling the issue of halo fragmentations during the construction of a merger tree for FOF haloes (see also Section 2.2 of this paper). Briefly, "snip" removes halo fragmentation events by severing the ancestral link between the fragment subhalo and its progenitor, "stitch" places the fragment subhalo back into the FOF halo from which it emerged, whereas "split" removes the fragment subhalo's progenitor from its FOF halo, thereby generating a new FOF at the progenitor redshift.

Within the stitch and split algorithms, the operations can be applied either on a subset of fragments or on all fragments in a given FOF tree. We therefore subdivide each algorithm into two: stitch-∞ vs stitch-3, and split-∞ vs split-3. The stitch-∞ and split-∞ algorithms perform the given operation on all FOF fragments. This is done recursively from the redshift of fragmentation, going forward in redshift for stitch-∞, and backwards in redshift for split-∞, until there are no more fragments present in the simulation merger trees. As a result of this recursive process, stitch-∞ identifies the first (highest-z) snapshot in which two subhaloes join the same FOF to be their merger time, whereas split-∞ selects the last (lowest-z) snapshot.

These algorithms introduce some complications, however. One particular problem faced by split-∞ is the fact that there exists a firm cutoff at $z = 0$, beyond which we do not have merger or fragmentation information. As a result, although a fragment may actually finally merge beyond $z = 0$, split-∞ will incorrectly assign its final merger to an earlier redshift. This results in a pile-up of merger events at $z = 0$ and, as we will show, artificially raises the low-z merger rate with respect to the high-z rate. Since there is no analogous hard

limit at high z, stitch-∞ does not suffer from this same behavior, and fragment mergers are re-distributed across all high redshifts evenly. On the other hand, any chance encounter between subhaloes that results in the FOF algorithm spuriously linking them together is interpreted as a real merger event by stitch-∞. The subhaloes, which may never interact again, are nonetheless forced to join the same FOF group down to $z = 0$.

The stitch-3 and split-3 algorithms are designed to limit the propagation effects of stitch-∞ and split-∞. Stitch-3 performs the stitching operation on any FOF fragment that is observed to remerge with its progenitor FOF's main branch within 3 snapshots of the fragmentation event. Any fragments that do not satisfy this criterion are snipped, resulting in an orphan halo that may or may not later remerge. The split-3 algorithm performs the split operation on any FOF fragment that is *not* a member of the main branch FOF at some point in the 3 snapshots *before* the fragmentation event. Again, fragments that do not satisfy this criterion are snipped.

Neither stitch-3 nor split-3 adequately removes all remerger events. Depending on the context this may be either a weakness or a strength: although the notion of halo remergers may be considered as multiple counting from a theoretical perspective, observers will likely count as signatures all events that trigger mergers, regardless of whether they are the first or last entry.

Moreover, both split-3 and stitch-3 have superior time convergence properties to the snip algorithm, in which the remerger problem is entirely unmitigated. Thus, stitch-3 and split-3 stand as intermediates between the snip and stitch-∞/split-∞ algorithms.

An immediate concern is whether the halo mass function is heavily modified by the destruction/creation of FOFs due to the stitch/split operations. We have verified that these operations do not modify the mass function severely. For stitch-3 and split-3, the deviations are within 3% of the unprocessed mass function at all redshifts, while deviations of up to 10% exist for the stitch-∞ (split-∞) algorithm at low (high) redshifts.

Fig. 2.10 compares the five post-processing algorithms directly by presenting ratios of the per-halo merger rate, $dN_m/d\xi/dz$, as a function of progenitor mass ratio ξ at five redshifts ($z = 0.06$, 0.5, 1, 2, and 4 from top bottom). Each column presents the ratio of the merger rate of a particular post-processing algorithm (left to right: snip, stitch-3, stitch-∞, split-∞) to the merger rate extracted from the split-3 trees presented throughout this paper. Different mass bins are presented by different colored curves ranging from $10^{10}\,\mathrm{M}_\odot$ (blue) to $10^{15}\,\mathrm{M}_\odot$ (red). We note that though the region of overlap between Millennium (dashed) and Millennium II (solid) is small, there appears to be smooth continuation between these two sets of curves for all post-processing algorithms.

The merger rates computed by all algorithms converge towards high z, though there is some residual disagreement with split-∞ and snip at the $\sim 20\%$ level. There are, however, distinct systematic differences among the algorithms when $z < 4$. Since we presented stitch-3 as our algorithm of choice for handling halo fragmentation in Fakhouri and Ma [2008], we focus on the comparison of stitch-3 and split-3 in this section. The origins of the differences between the other algorithms and split-3 can be inferred from the discussion of the algorithms earlier in this section.

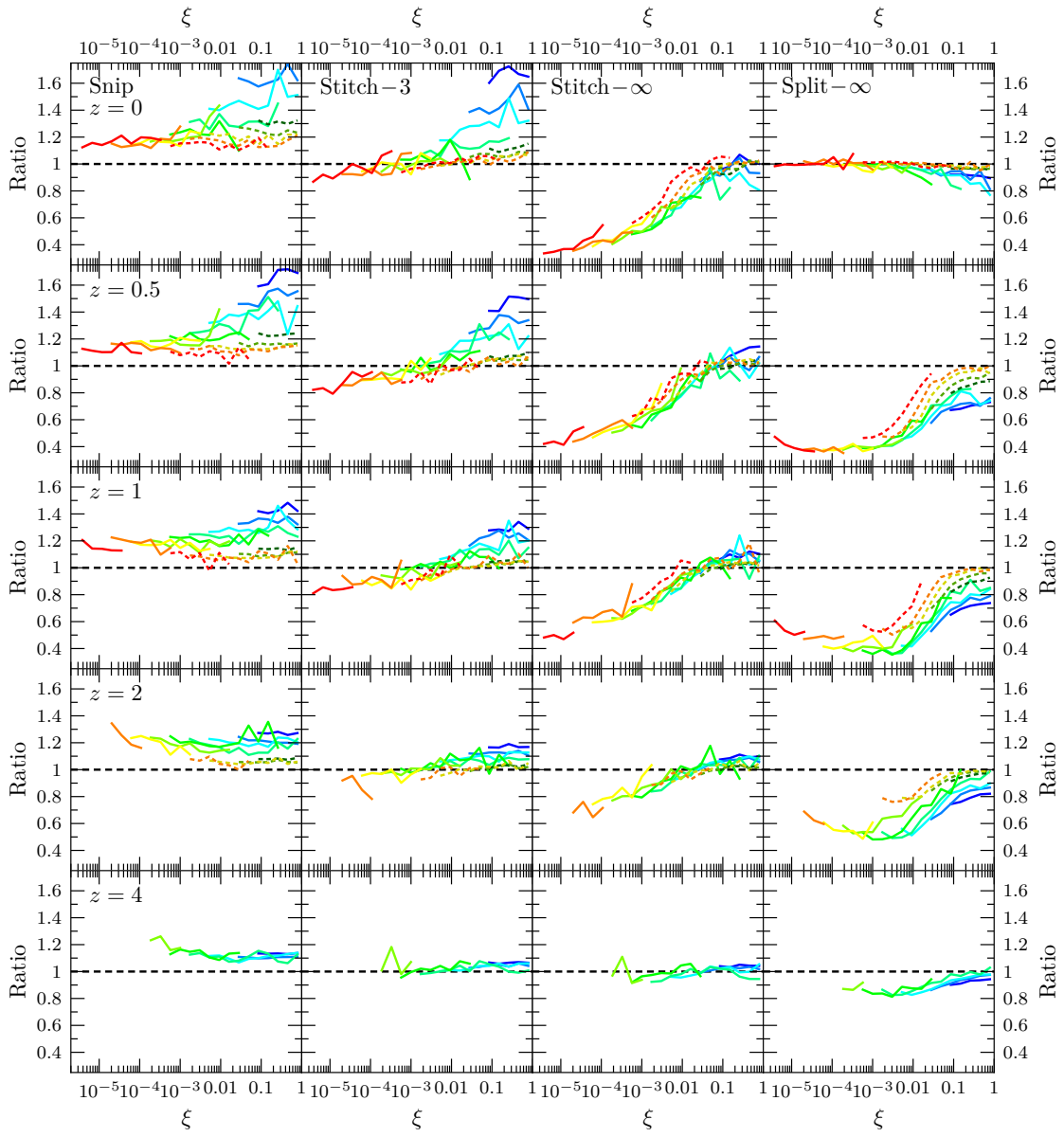

Figure 2.10: Comparison of five algorithms used to handle halo fragmentation events in the FOF merger trees: snip, stitch-3, stitch-∞, split-∞, and split-3. Results presented throughout this paper are based on the split-3 tree. Plotted as a function of the progenitor mass ratio ξ (left to right) are the ratios of the merger rates, $dN_m/d\xi/dz$, between each of the first four algorithms relative to the split-3 results. Five redshifts are shown: $z = 0, 0.5, 1, 2$ and 4 (top to bottom). Within each panel, up to nine mass bins are shown: $10^{10} M_\odot$ (blue) to $> 10^{14} M_\odot$ (red). The Millennium Simulation results are presented with dashed curves and Millennium II with solid curves. The systematic differences amongst the five algorithms are discussed in the text.

The second column of Fig. 2.10 shows that stitch-3 and split-3 are in excellent agreement at all ξ for high mass haloes ($M_0 > 10^{12}$). Low mass haloes, however, can show significant deviations in the merger rate. This is true especially in the major merger regime, where the merger rate predicted by stitch-3 is over 50% higher than split-3. This distinction was not detectable using Millennium alone, as the mass resolution limited our analysis to $M_0 \geq 10^{12} M_\odot$.

To understand this deviation we have studied a subset of halo mergers in detail by analyzing halo tracks, velocities, and merger histories. In particular, we have constructed a number of criteria to determine whether a given merger is actually a spurious encounter: if the relative velocity of the two haloes greatly exceeds the more massive halo's maximum circular velocity, if the angle between the velocity vectors of the two haloes exceeds 70° at the time of merger, or if the FOF algorithm only associates the two haloes for two snapshots out of the eight snapshots centered on the merger snapshot, then the merger is deemed spurious. A qualitative look at three-dimensional halo trajectories finds that this criteria does a good job of identifying chance halo encounters and premature mergers.

For halo mergers with $1.1 \times 10^{10} < M_0 < 1.3 \times 10^{10}$ and $\xi > 0.1$ at $z = 0$, we find that stitch-3 identifies $1,304$ mergers, while split-3 only identifies 738 mergers. Of these, 505 mergers are in common, leaving split-3 with 233 mergers that are not in stitch-3 and stitch-3 with 799 mergers that are not in split-3. Of the 505 mergers in common, only 4 (0.8%) are deemed spurious by our criterion. Similarly, of the 233 mergers unique to split-3, only 12 (5.2%) are deemed spurious. Of the 799 mergers unique to stitch-3, however, 589 (73.7%) are deemed spurious. These spurious mergers are primarily comprised of chance encounters in which the two otherwise unassociated haloes merge for a snapshot or two and then disconnect. While split-3 correctly splits these events, stitch-3 does not and consequently inflates the merger rate. When these spurious mergers are removed, the remaining 210 mergers unique to the stitch-3 algorithm bring the stitch-3 and split-3 rates into close agreement.

We note that depending on the context, one may choose one algorithm over another. Stitch-∞ provides the *first* encounter merger rate, but is known to link chance-encounter haloes that should not be linked. Split-∞ provides the *last* encounter merger rate, but cannot be trusted for $z < 1$ and may incorrectly underpredict the merger rate due to spurious fragmentation. Split-3 stands in between both algorithms: it does not propagate up and down the tree and does not heavily modify the distribution of FOFs, but it does double count some halo remerger events. This may be odious to the theorist, but may yield the most appropriate merger rate for comparison to observation.

Part III

The Mass Accretion Histories of Dark Matter Haloes

Chapter 3

Mass Accretion Rates and Histories of Dark Matter Haloes

McBride, James; Fakhouri, Onsi ; Ma, Chung-Pei — October 2009
Monthly Notices of the Royal Astronomical Society, Volume 398, Issue 4, pp. 1858-1868

We use the extensive catalog of dark matter haloes from the Millennium simulation to investigate the statistics of the mass accretion histories (MAHs) and accretion rates of $\sim 500,000$ *haloes from redshift* $z = 0$ *to 6. We find only about 25% of the haloes to have MAHs that are well described by a 1-parameter exponential form. For the rest of the haloes, between 20% (Milky-Way mass) to 50% (cluster mass) experience late-time growth that is steeper than an exponential, whereas the remaining haloes show plateau-ed late-time growth that is shallower than an exponential. The haloes with slower late-time growth tend to reside in denser environments, suggesting that either tidal stripping or the "hotter" dynamics are suppressing the accretion rate of dark matter onto these haloes. These deviations from exponential growth are well fit by introducing a second parameter:* $M(z) \propto (1+z)^{\beta} e^{-\gamma z}$*. The full distribution of* β *and* γ *as a function of halo mass is provided. From the analytic form of* $M(z)$*, we obtain a simple formula for the mean accretion rate of dark matter,* $\dot{M}$*, as a function of redshift and mass. At* $z = 0$*, this rate is* $42 M_{\odot}\,\mathrm{yr}^{-1}$ *for* $10^{12} M_{\odot}$ *haloes, which corresponds to a mean baryon accretion rate of* $\dot{M}_b = 7 M_{\odot}\,\mathrm{yr}^{-1}$*. This mean rate increases approximately as* $(1+z)^{1.5}$ *at low* z *and* $(1+z)^{2.5}$ *at high* z*, reaching* $\dot{M}_b = 27, 69$*, and 140* $M_{\odot}\,\mathrm{yr}^{-1}$ *at* $z = 1$*, 2, and 3. The specific rate depends on halo mass weakly:* $\dot{M}/M \propto M^{0.127}$*. Results for the broad distributions about the mean rates are also discussed.*

3.1 Introduction

The mass growth history is a basic property of dark matter haloes. Haloes in numerical simulations are seen to be assembled through a number of processes: mergers with comparable mass haloes ("major mergers"), mergers with smaller satellite haloes ("minor mergers"),

and accretion of non-halo material that is composed of either haloes below the numerical resolution or diffuse particles. Following the mass history of the most massive progenitor halo as a function of redshift z is a useful way to quantify a halo's mass assembly history. These mass accretion histories (MAHs, or $M(z)$) are important for statistical studies of the distributions of halo formation redshifts, and the correlations between formation time and other halo properties such as environment, concentration, substructure fraction, spin, and relative contributions to mass growth from major vs minor mergers. Moreover, the time derivative of the MAH gives the mass growth rate of dark matter haloes, which is directly related to the accretion rate of baryons from the cosmic web onto dark matter haloes.

A number of earlier papers have investigated various aspects of the halo MAHs. For instance, Wechsler et al. [2002] analyzed ~ 900 haloes (above $10^{12}h^{-1}M_{\odot}$ at $z = 0$) in a ΛCDM simulation in a $60\,h^{-1}$ Mpc box with 256^3 particles. The values from a 1-parameter fitting function for the MAHs were presented for 8 haloes. Clear correlations between the formation redshift z_f and concentration c of haloes were seen, with late-forming haloes being less concentrated. The scatter in c was attributed to the scatter in z_f. An alternative 2-parameter fitting function was demonstrated by van den Bosch [2002] to be superior to a 1-parameter fit to haloes in a simulation with the same particle number in a $141h^{-1}$ Mpc box.

The relationship between halo structure and accretion was further addressed in Zhao et al. [2003b] and Zhao et al. [2003a], where the redshift dependence of c was observed to be more complicated than a simple proportionality. Tasitsiomi et al. [2004] examined 14 haloes, ranging in mass from group to cluster scale (.58 to $2.5 \times 10^{14}h^{-1}M_{\odot}$) and also found that a 2-parameter fit for $M(z)$ worked better. Cohn and White [2005] studied the mass accretion histories of ~ 1500 cluster-sized haloes and characterized several properties of galaxy cluster formation.

Maulbetsch et al. [2007] studied the environmental dependence of the formation of ~ 4700 galaxy-sized haloes (above $10^{11}h^{-1}M_{\odot}$) in a $50\,h^{-1}$ Mpc simulation box. In higher-density environments, they found the haloes to form earlier with a higher fraction of their final mass gained via major mergers. Li et al. [2008] studied 8 different definitions of halo formation time using the haloes from the Millennium simulation [Springel et al., 2005]. The motivation was to search for halo formation definitions that better characterize the downsizing trend in star formation histories, as opposed to the hierarchical growth of haloes in the ΛCDM cosmology. Zhao et al. [2009] (Z08) investigated the mean MAH in different cosmological models – scale-free, ΛCDM, standard CDM, and open CDM – and searched for scaled mass and redshift variables that would lead to a universal fitting form for the median MAH for all models.

The results in these earlier papers were presented either for $M(z)$ of a handful of individual haloes, or for the global mean growth of a selection of haloes. Our aim here is to quantify systematically the diversity of growth histories and rates using the $\sim 500,000$ $z = 0$ haloes with $M > 10^{12}M_{\odot}$ (i.e. above 1000 simulation particles) and their progenitors in the Millennium simulation. Over this large range of haloes, we find that an exponential fit does not adequately capture the behavior of halo growth. Many haloes experience large changes

in the rate at which they accrete mass. Some haloes grow more slowly at late times, and occasionally even lose mass, while other haloes undergo late bursts of growth. All of these MAHs are poorly fit by an exponential, and suggest the need for a fitting form with more flexibility. We find it helpful to classify the MAHs into four types based on their late-time accretion rate. The large ensemble of haloes allows us to quantify the mean values as well as the dispersions of the mass accretion rates and halo formation redshifts as a function of mass and redshift.

This paper is organized as follows. Sec. 2 provides some background information about the haloes in the Millennium simulation and describes how we construct halo merger trees. This post-processing of the Millennium public data is necessary for identifying the thickest branch (i.e. the most massive progenitor) along each final halo's past history. The masses of these progenitors will then allow us to quantify the MAH, $M(z)$. In Sec. 3, we first assess the accuracy of the 1-parameter exponential form for $M(z)$. We then propose a more accurate two-parameter function for $M(z)$ and classify the diverse assembly histories into four broad types according to their late-time growth behavior. We further quantify the statistics of the two fitting parameters, providing (in the Appendix) algebraic fits for their joint distributions that can be used to generate Monte Carlo realizations of an ensemble of halo growth tracks. The applicability of $M(z)$, which is derived for $z = 0$ haloes, for the mass accretion history of higher-redshift haloes is discussed in Sec. 3.3. Sec. 4 is focused on the statistics of the mass accretion rates. A simple analytic expression is obtained for the mean accretion rate, $\left\langle \dot{M} \right\rangle$, of dark matter as a function of halo mass and redshift. The dispersions about the mean rates are significant, as evidenced by the differential and cumulative distributions of $\dot{M}$ presented here. Sec. 5 discusses the mean and the distribution of the halo formation redshift as a function of halo mass. In Sec. 6 we report the correlations of MAHs with halo environment, the last major merger redshift, and the fraction of haloes' final masses assembled via different types of mergers.

3.2 Halo Merger Trees in the Millennium Simulation

The Millennium simulation [Springel et al., 2005] provides a database for the evolution of roughly 2×10^7 $z = 0$ dark matter haloes from redshifts as high as $z = 127$ in a $500h^{-1}$ Mpc box using 2160^3 particles of mass $1.2 \times 10^9 M_\odot$ (all masses quoted in this paper are in units of $M_\odot$ and not $h^{-1}M_\odot$). It assumes a ΛCDM model with $\Omega_m = 0.25$, $\Omega_b = 0.045$, $\Omega_\Lambda = 0.75$, $h = 0.73$, and a spectral index of $n = 1$ for the density perturbation power spectrum with a normalization of $\sigma_8 = 0.9$.

Dark matter haloes are identified with a friends-of-friends (FOF) group finder [Davis et al., 1985] with a linking length of $b = 0.2$. Throughout this paper we use the number of particles linked by the FOF finder to define the halo's mass. Once identified, each FOF halo is then broken into gravitationally bound substructures (subhaloes) by the SUBFIND algorithm (see Springel et al. 2001a). These subhaloes are connected across the 64 available redshift outputs: a subhalo is the descendant of a subhalo at the preceding output if it

hosts the largest number of the progenitor's bound particles. The resulting subhalo merger tree can be used to construct merger trees of FOF haloes, although we have discussed at length in Fakhouri and Ma [2008, 2009] the complications due to halo fragmentation and have presented comparisons of several post-processing algorithms that handle fragmentation events.

Our results in this paper are based on the stitch-3 post-processing algorithm described in Fakhouri and Ma [2008]. In this algorithm, fragmented haloes that remerge within 3 outputs after fragmentation are stitched into a single FOF descendant; those that do not remerge within 3 outputs are snipped and become orphan haloes. After applying the stitch-3 algorithm, we extract the mass accretion history, $M(z)$, of each halo at $z = 0$ (or at any higher redshift) by following the halo's main branch of progenitors. We have compared the resulting $M(z)$ and formation redshifts to those obtained from the alternative algorithms (e.g., "snip," "split," and subhalo vs FOF mass) discussed in Fakhouri and Ma [2008, 2009]. We find the systematic variations to all be within 5-10% of the stitch-3 values of these quantities.

3.3 Fitting Mass Accretion Histories

3.3.1 Previous MAH Forms

To quantify the limitations of the exponential fit in capturing halo growth, consider the formation redshift z_f, here defined as the redshift at which $M(z)$ is equal to $M_0/2$.

For the 1-parameter exponential form (e.g. Wechsler et al. 2002)

$$M(z) = M_0 e^{-\alpha z}, \tag{3.1}$$

the parameter α is simply related to z_f by

$$z_f = \frac{\ln(2)}{\alpha}. \tag{3.2}$$

We have compared z_f as determined by the exponential fit to each halo's $M(z)$ from the simulation with the z_f determined directly from the $M(z)$ tracks such that $M(z_f) = M(0)/2$ (using interpolation between output redshifts). We find the exponential fit to err systematically in its determination of z_f, significantly overestimating the formation redshift for haloes that form recently and underestimating it for haloes that form early. The mean value of z_f from the exponential fit, for instance, is 0.3 higher than the actual value for young haloes and is 0.8 lower for old haloes across all masses.

A more complicated functional representation of MAHs was put forth by van den Bosch [2002]:

$$\log\left(\frac{M(z)}{M_0}\right) = -0.301\left[\frac{\log(1+z)}{\log(1+z_f)}\right]^{\nu}, \tag{3.3}$$

where ν and z_f are fitting parameters. The use of an additional parameter provided significant improvement in the quality of the fits for many MAHs, especially those that formed late. This 2-parameter form, however, is not flexible enough to handle haloes that have lost mass, as it cannot take on values that would give $M(z)/M_0$ greater than 1. Moreover, over the sample of haloes tested in van den Bosch [2002], comparison between the goodness-of-fit of this two-parameter form and the exponential fit showed that the exponential fit actually performs better for early forming haloes.

3.3.2 A Revised MAH Form

To address the need for a fit that is both effective and simple, we find a 2-parameter function of the form

$$M(z) = M_0(1+z)^{\beta}e^{-\gamma z}, \tag{3.4}$$

to be versatile enough to accurately capture the main features of most MAHs in the Millennium Simulation. This form has also been studied in Tasitsiomi et al. [2004] for cluster-mass haloes, but it has not been tested over a large number of haloes of different mass. The form reduces to an exponential when β is 0, and γ in this case is simply the inverse of the formation redshift: $\gamma = \ln(2)/z_f$. A large fraction of the haloes, however, are better fit when the additional factor of $(1+z)^{\beta}$ in equation (3.4) is included. In general, β can be either positive or negative, but $\gamma \geq 0$. We find the combination $\beta - \gamma$ to be a useful parameter for characterizing these MAHs as $\beta - \gamma$ gives the mass growth rate at small redshifts:

$$\frac{d\ln(M(z))}{dz} = \frac{\beta}{1+z} - \gamma \approx \beta - \gamma + \mathcal{O}(z)\,. \tag{3.5}$$

This late-time trend can be used to characterize the MAH as described below.

To obtain the best-fit values for β and γ in equation (3.4), we have performed a χ^2-like minimization of the quantity

$$\Delta^2 = \frac{1}{N}\sum_{N}\frac{[M(z_i)/M_0 - (1+z_i)^{\beta}e^{-\gamma z_i}]^2}{M(z_i)/M_0}\,, \tag{3.6}$$

where the sum is over the N-available simulation redshift outputs at $z_i (i = 1, ..., N)$ for each halo. The choice of the factor $M(z_i)$ in the denominator is akin to assuming Poissonian errors for halo masses. We found this choice to be a suitable middle ground between minimizing simply the sum of squares and minimizing the fractional deviation (i.e. with a factor of $M^2(z)$ in the denominator). The former tended to fit the finely sampled low-z points well at the expense of the sparsely spaced high-z points, whereas the latter tended to do the opposite. Equation (3.6), on other hand, provides reasonable fits for the entire history of the halo growth. Fig. 3.1 shows the cumulative distribution for the rms deviation of the fits from the Nbody data (normalized by M_0) for all $\sim 500,000$ $z = 0$ haloes. The deviation is less than 6% for over 75% of the haloes, and only a few percent of haloes have deviations larger than 10%. Of this most poorly fit subset of haloes, nearly half underwent mass loss

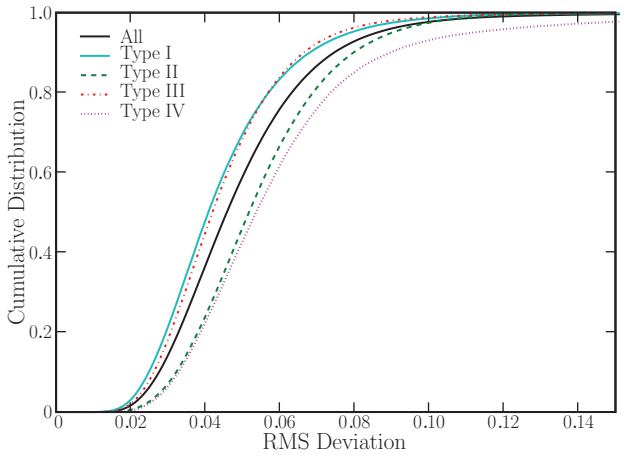

Figure 3.1: Cumulative distribution for the RMS deviations between the fitting formula eq. (3.4) and the Millennium Nbody output for the mass growths of 478,781 $z = 0$ haloes (solid black) and the four sub-types (colored curves) listed in Table 1. The figure shows that the fits perform well overall: about 75% of the haloes have RMS deviations less than $\sim 6\%$.

at late times. As expected, the fits become progressively worse at higher redshifts; for over 75% of haloes, the maximum fractional deviation between the fits and Nbody results occurs above $z = 4$.

We suggest that the parameters β and γ allow for rough classifications of MAHs into a few basic groups, summarized in Table 3.1. The classification scheme is quite straightforward. Fits with small values for β indicate a weak contribution from the $(1 + z)$ term, and deviate minimally from an exponential curve. These haloes with $|\beta| < 0.35$ are labeled Type I.

The rest of the classifications are dependent upon the value of $\beta - \gamma$. The motivation for this is the fact that the difference represents the value of the derivative at $z = 0$, as noted in

Type	Criteria	Characteristics	$\frac{\chi_1^2}{\chi_2^2}$
I	$\lvert\beta\rvert < 0.35$	Good exponential	1.09
II	$\beta - \gamma <$ -0.45	Steep late growth	1.61
III	-0.45 $< \beta - \gamma < 0$	Shallow late growth	2.15
IV	$\beta - \gamma > 0$	Late plateau/decline	3.31

Table 3.1: MAHs are categorized based upon the best fit parameters β and γ of equation (3.4). Categorization is done in order by type; thus MAHs that satisfy the criteria for both Type I and Type II belong to Type I. The right-most column is the mean of χ_1^2/χ_2^2, the ratio of the χ^2 computed for the 1-parameter exponential form to the χ^2 computed for the 2-parameter form in equation (3.4). Values > 1 imply that the 2-parameter form provides a more accurate fit than the 1-parameter form.

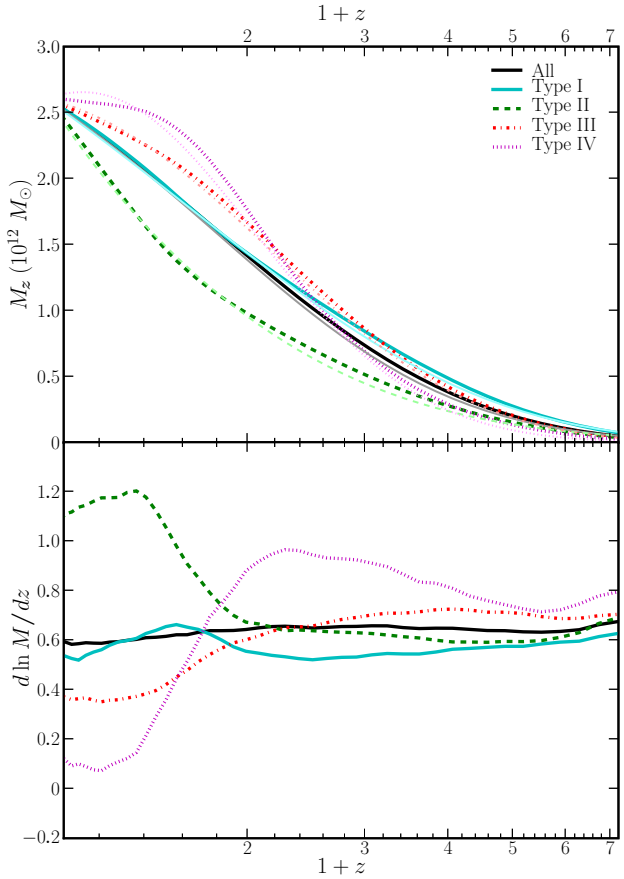

Figure 3.2: Average mass accretion histories (top) and their derivatives (bottom) for the total population (black solid) and the four types of halo growths listed in Table 3.1. The average is taken over the $M(z)$ for galaxy-sized haloes with masses between $2.1 \times 10^{12} M_\odot$ and $3.3 \times 10^{12} M_\odot$ in the Millennium simulation, although haloes at different mass show similar behavior. In the top panel, the set of curves with the lighter shading shows the average $M(z)$ computed from our fits of equation (3.4) to each halo's MAH. The bottom panel illustrates that the late-type growth rates differ greatly, ranging from $d \ln M/dz \sim 1.2$ for Type II to 0.1 for Type IV.

equation (3.5). Hence Type II haloes, defined to be those haloes with $\beta - \gamma < -0.45$, feature steep growth at late times, typically steeper than can be captured by an exponential fit.

Type III haloes have fit parameters that fall in the range $-0.45 < \beta - \gamma < 0$ and exhibit flat late time growth. Like Type II, these tend to deviate from the fit that would be found using the exponential form, but Type III haloes do so in the opposite direction to Type II haloes. A typical Type III halo has undergone limited growth during recent times, sometimes after a spurt of growth at earlier times.

Type IV, with $\beta - \gamma > 0$, represents the most extreme deviation from an exponential. The majority of Type IV haloes have shed mass, some of them by significant amounts. Some Type IV haloes have merely seen their growth slow down like the Type III haloes, but over a more significant period of time. As such, Type IV haloes are extreme cases of Type III haloes, perhaps representing the future growth for some Type III haloes.

The boundaries delineating these classifications are rough guidelines at best. For example, consider the definition for Type I of $|\beta| < 0.35$. For the largest values of β in this group, which should be considered the worst of the "good exponentials" that constitute Type I, the fractional difference between the formation redshift as determined by the simple exponential and the modified exponential is a little under 8%. The agreement is not perfect, but the two fits are similar enough for these haloes that the use of the power law parameter adds little. Of course, there is no reason why we should not instead demand that the formation redshifts differ on average by no more than 5%, or perhaps 10%. In the end, the combination of the formation redshift metric and a couple of others for comparing the fits suggested that demanding $|\beta| < 0.35$ was inclusive enough to capture the majority of haloes for which an exponential is an adequate fit, without unduly diminishing the integrity of the group.

Fig. 3.2 compares the shapes of the average MAH for haloes of galaxy-size mass from the Millennium simulation for the overall distribution and for each type. The bottom panel shows the derivative $d \ln M/dz$ to highlight the different late-time accretion rates among the four types. Haloes of other mass show similar behavior. Clearly, the late time growth rate is an important factor in distinguishing haloes from one another. The average MAH for Type I haloes is quite similar to that of the overall distribution, which indicates that the average MAH is approximately exponential. However, the behavior of about 75% of individual haloes deviates from an exponential noticeably. This fact is quantified in the right-most column of Table 3.1, where the ratio of χ^2 for the exponential fit to the 2-parameter fit is seen to increase with the MAH types.

Since the mean MAH is approximately exponential, the accretion rate $d \ln M/dz$ averaged over the whole population is also nearly independent of redshift (black solid curves in Fig. 3.2) when expressed in units of per redshift, with $d \ln M/dz$ being between 0.6 and 0.7 for $z = 0$ up to 5. This weak dependence on redshift is similar to that of the halo merger rates (per unit z) reported in Fakhouri and Ma [2008]. The different types of haloes, however, show significant dispersions in the late-time accretion rates, with $d \ln M/dz$ being as high as 1.2 for Type II and as low as 0.1 for Type IV at $z \approx 0$.

For each of the mean profiles shown in Fig. 3.2, we have fit the analytic form in equation (3.4). The best-fit values of (β, γ) are (0.10, 0.69) for all haloes, and (−0.04, 0.54),

Mass Range ($10^{12}M_\odot$)	Halo Number	Type I	II	III	IV
1.2 to 2.1	191421	29%	27%	32%	12%
2.1 to 4.5	143356	27%	29%	32%	12%
4.5 to 14	95744	24%	34%	31%	11%
14 to 110	43089	20%	42%	26%	11%
> 110	4787	18%	57%	17%	8%
	478781	27%	31%	31%	11%

Table 3.2: Within each mass range, the percentage of haloes that belong to each type are provided. For each type, there is a noticeable trend with mass, though the strength of the trend varies.

$(-0.9, 0.35)$, $(0.62, 0.88)$, and $(1.42, 1.39)$ for each of the four types, respectively.

The statistics of the 478,781 $z = 0$ haloes (above 1000 particles, or a mass of $1.2\times10^{12}M_\odot$) belonging to each MAH type across different mass bins are given in Table 3.2 and Fig. 3.3. Exponential MAH (Type I) is seen to apply to only 20 to 30% of the haloes. There is also interesting dependence of the type on halo mass. Most notably, Type II haloes feature a strong dependence on mass, where the fraction rises from 27% at $\sim 10^{12}M_\odot$ to 60% at $\gtrsim 10^{14}M_\odot$. Cluster-size haloes therefore not only form late, which is a natural consequence of the ΛCDM cosmology, but the majority of their mass accretion rates is also faster than an exponential at low redshifts.

The results presented thus far are for the mean MAH and mean values of (β, γ). We find, however, significant dispersions about the mean behavior that are also important to characterize. For completeness, we show the distributions of our best-fit (β, γ) for all halo MAHs in the Appendix and Fig. 3.12. We also present there an accurate fitting form that we have obtained for the two-dimensional probability distribution of β and γ as a function of halo mass. This formula can be used to generate a Monte Carlo ensemble of realistic halo growth histories. The details of the formula, its usage, and comparison to the Millennium data are described in the Appendix. We emphasize that the results presented for the rest of this paper are obtained from the Millennium haloes directly rather than from this Monte Carlo realization.

3.3.3 MAHs for Haloes at Higher Redshifts

The MAHs presented thus far are obtained from the main branches of the descendant haloes at $z = 0$. Thus, for a higher redshift $z_1 > 0$, the distribution of $M(z_1)$ contains only information about the main branch progenitors, which is a subset of all the haloes at z_1 since many haloes do not belong to main branches.

Since the formation of higher-redshift galaxies and their host haloes is of much interest, it is useful to quantify the behavior of MAHs for haloes at z_1, where $z_1 > 0$. In particular,

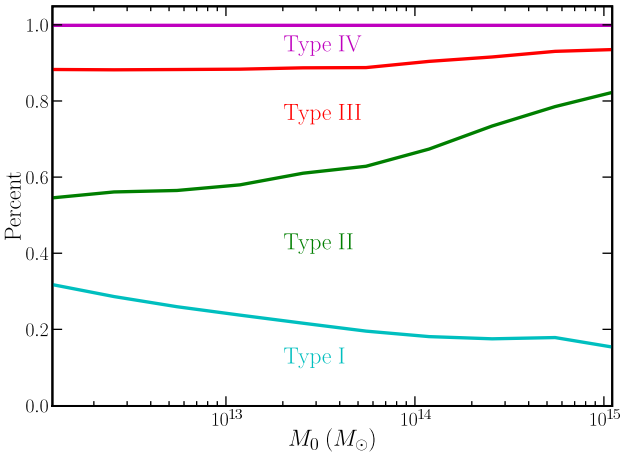

Figure 3.3: Cumulative fraction of haloes belonging to a given MAH type as a function of halo mass (e.g. the magenta type-IV curve includes the contributions from the three other types). The exponential form (Type I) is a good fit for only 20% to 30% of the haloes at all masses. The type II fraction shows a strong mass dependence, reaching $\sim 60\%$ for cluster-mass haloes.

we ask whether the mean MAH for haloes of mass M_1 at z_1 for $z > z_1$ can be related to the MAHs of haloes at $z = 0$ that we have studied thus far.

We find that the mean MAH of haloes of mass M_1 at z_1 is nearly identical to the mean MAH of haloes at $z = 0$ that satisfy $M(z_1) = M_1$. That is, the mean MAH for $z > z_1$ of the main branch subset with mean mass $M(z_1) = M_1$ at z_1 is very similar to the mean MAH of the complete population of haloes with mean mass M_1 at z_1. As a specific example, the mean MAH of the $10^{13}M_\odot$ $z = 0$ haloes in the simulations had the value $M(z = 1) = 4.5 \times 10^{12}M_\odot$ at $z = 1$. We find that the mean MAH of these $10^{13}M_\odot$ haloes at $z > 1$ is nearly identical (within 2%) to the $z > 1$ evolution of the mean MAH of *all* $M_1 = 4.5 \times 10^{12}M_\odot$ haloes at $z = 1$. This property for the mean MAH is in fact a natural consequence of the Markovian nature of the Extended Press-Schechter theory (see, e.g., Sec 2.3 of White 1994).

This self-similar property implies that in order to study the MAH properties of haloes with mass M_1 at redshift z_1, one simply needs to determine which set of haloes at $z = 0$ have $M(z_1) = M_1$. In particular, one needs to compute the average mass M_0 of the haloes at $z = 0$ that map onto $M(z_1) = M_1$ at z_1. This mapping is shown in Fig. 3.4 with M_1 along the x-axis and M_0 along the y-axis for $z_1 = 0, 0.5, 1$, and 2. Note that $M_0 = M_1$ at $z_1 = 0$ by construction, and as z_1 increases, the mass M_0 that maps onto some fixed M_1 by redshift z_1 also increases.

We note that the mapping in Fig. 3.4 implies that haloes of some mass M_1 at some redshift $z_1 > 0$ do *not* have the same shape of MAH as haloes of mass $M_0 = M_1$ at $z = 0$. That is, the MAH of a $10^{13}M_\odot$ halo at $z = 0$ and the MAH of a $10^{13}M_\odot$ halo at $z_1 > 0$ are not simply related by a shift from z to $z - z_1$ in equation (3.4). This is because haloes at higher z_1 have a *relative* formation redshift $z_f - z_1$ that is smaller than haloes of the same mass at

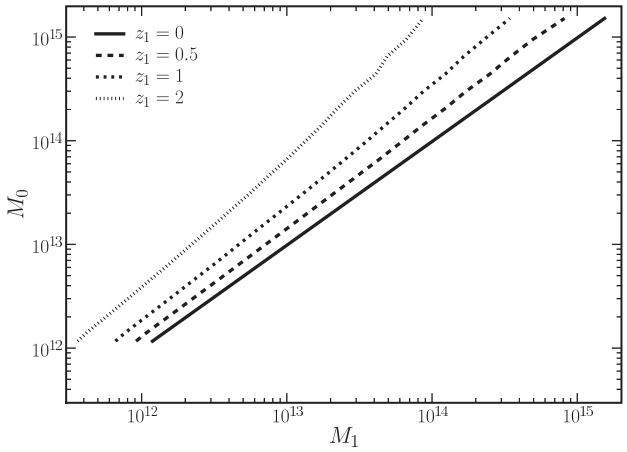

Figure 3.4: For haloes of mass M_1 at redshift z_1, the y-axis plots their corresponding mean mass M_0 today. Four values of z_1 are shown: 0 (solid), 0.5 (dashed), 1 (dot dashed), and 2 (dotted). This mapping allows one to use equation (3.4) for the MAH of higher-redshift haloes (see text).

$z = 0$. This result is not surprising since haloes of the same mass at different redshifts in the ΛCDM model represent different part of the mass spectrum and are not generally expected to have identical properties.

We have tested the self-similar property of the fitting form of Z08 (using their online code) by comparing their mean MAH for $z = 0$ $10^{13} M_\odot$ haloes and the MAH for their $M(z = 1)$ haloes at $z > 1$. Their latter MAH is higher than the former by about 15%, while ours differ by less than 2%.

3.4 Mass Accretion Rates: Mean and Dispersion

Having quantified $M(z)$ in Sec. 3, we now examine its time derivative – the mass accretion rate – in more detail. In particular, we would like to obtain a general formula for the mean accretion rates of dark matter for a wide range of halo mass and redshift. To achieve this, we note that our analytical form in equation (3.4) for individual halo MAHs gives:

$$\frac{\dot{M}}{M} = 0.10h\,\mathrm{Gyr}^{-1}\,[\gamma(1+z) - \beta]\,\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda} \tag{3.7}$$

where Ω_m and Ω_Λ are the present-day density parameters in matter and the cosmological constant, and we have assumed $\Omega_m + \Omega_\Lambda = 1$ (used in the Millennium simulation) and matter-dominated era in computing dz/dt. As shown in Sec. 3, the parameters β and γ in equation (3.7) generally depend on the halo mass. We find, however, that the mass dependence follows a simple power law independent of the redshift, and the simple analytic form in equation (3.7) provides an excellent approximation for the *mean* mass accretion rate

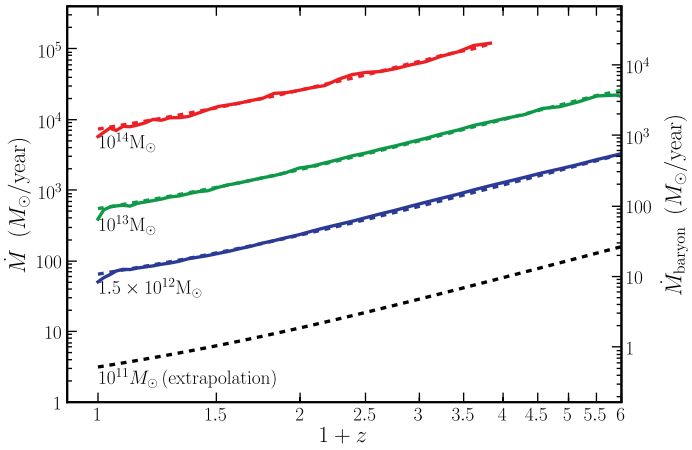

Figure 3.5: Mean mass accretion rate of dark matter as a function of redshift at halo mass $10^{11}, 1.5\times10^{12}, 10^{13}$ and $10^{14} M_\odot$. The solid curves are computed from the Millennium haloes (except $10^{11} M_\odot$, which falls below our resolution limit of 1000 particles per halo) at a given mass ($\pm$20% range); the dashed curves show the accurate fit provided by eq. (3.8). The right side of the vertical axis labels the mean accretion rate of baryons, $\dot{M}_b$, assuming a cosmic baryon-to-dark matter ratio of $\sim 1/6$. The slight dip in $\dot{M}$ at $z = 0$ is due to the artificial edge effect inherent in the stitch-3 algorithm used to process the FOF merger trees.

as a function of redshift and halo mass:

$$\left\langle \dot{M} \right\rangle = 42\, M_\odot \mathrm{yr}^{-1} \left(\frac{M}{10^{12} M_\odot}\right)^{1.127} \times (1+1.17z)\sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda}\,. \quad (3.8)$$

For completeness, the best fit for the *median* growth rate computed in the Millennium simulation is

$$\left\langle \dot{M} \right\rangle_{\rm median} = 24.1\, M_\odot \mathrm{yr}^{-1} \left(\frac{M}{10^{12} M_\odot}\right)^{1.094} \times (1+1.75z)\sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda}\,. \quad (3.9)$$

We note that the overall amplitude of the mean is higher than the median due to the long, positive, $\dot{M}$ tail (see Fig. 3.6).

Fig. 3.5 compares the mean accretion rates of dark matter in $M_\odot$ per year computed from the Millennium simulation (solid curves) and this formula (dashed curves) for haloes of mass $10^{12} M_\odot$ to $10^{15} M_\odot$ over the redshift range of 0 and 5. The overall trend of the accretion rate is such that $\dot{M}/M$ has a weak dependence on M ($\propto M^{0.127}$), and its dependence on redshift is approximately $(1+z)^{1.5}$ at low z and $(1+z)^{2.5}$ at $z > 1$. This z-dependence is motivated by our 2-parameter form for $M(z)$ and is more accurate than the simple power law used in Genel et al. [2008], Neistein et al. [2006], and Neistein and Dekel [2008b]; our $z \sim 0$ value, on

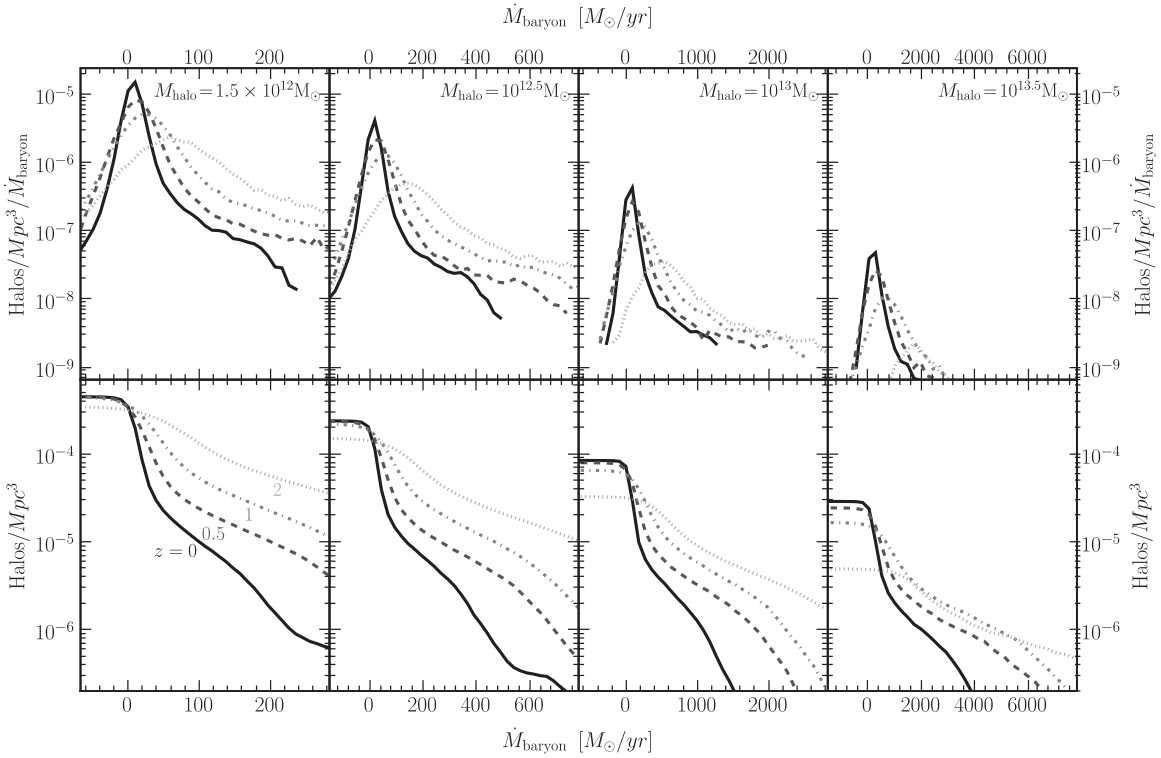

Figure 3.6: Differential (top) and cumulative (bottom) distributions of the accretion rates of cosmic baryons, $\dot{M}_b$, for four halo masses (left to right). Within each panel, the accretion rates at four redshifts $z = 0$ (solid), 0.5 (dashed), 1 (dashed dotted), and 2 (dotted) are shown, where the distributions are seen to broaden significantly with increasing z. The vertical axis labels the number of haloes per comoving Mpc^3 at or above a given $\dot{M}_b$.

the other hand, is consistent with theirs to within 20%. We have also computed $\dot{M}$ from the fitting form for the median MAH in the recent preprint by Z08. We found their $\dot{M}$ to have a slightly steeper z-dependence than our equation 3.9 where their median value is within 20% of our median $\dot{M}$ at $z \sim 0$ but exceeds ours by a factor of ~ 2 at $z \sim 4$.

Along the right side of the vertical axis of Fig. 3.5, we label the corresponding mean accretion rates of baryons, $\dot{M}_b$, assuming a cosmic baryon-to-dark matter ratio of $\Omega_b/\Omega_m \approx 1/6$. The results shown should be a reasonable approximation for the mean rate of baryon mass that is being accreted at the virial radius of a dark matter halo of a given mass. Fig. 3.5 and equation (3.8) indicate that this rate is $\dot{M}_b \approx 7\,M_\odot \mathrm{yr}^{-1}$ for $10^{12} M_\odot$ haloes today, and it increases to 27, 69, and 140 $M_\odot \mathrm{yr}^{-1}$ for $10^{12} M_\odot$ haloes at $z = 1$, 2, and 3, respectively. Since the infalling baryons are a reservoir for the gas that fuels star formation, it is interesting to compare $\dot{M}_b$ with the mean star formation rates of different types of galaxies, e.g., $\dot{M}_* \sim 4 M_\odot \mathrm{yr}^{-1}$ for the Milky Way (e.g., Diehl et al. 2006), suggesting that about half of the infalling $\dot{M}_b \approx 7\,M_\odot \mathrm{yr}^{-1}$ for Galactic-size haloes needs to be converted into stars. The relations among these different accretion rates and the implications will be investigated in a subsequent work.

Having determined the mean rates, we show their distributions and dispersions in Fig. 3.6. Four redshifts, $z = 0$, 0.5, 1, and 2, and four ranges of halo masses (left to right panel) are shown. Both the differential (top panels) and cumulative (bottom panels) distributions of $\dot{M}_b$ are plotted for comparison. Within each panel, the distribution of $\dot{M}_b$ at a given halo mass is seen to broaden significantly with increasing redshift. For instance, the (comoving) number density of $1.5 \times 10^{12} M_\odot$ haloes with $\dot{M}_b > 250 M_\odot$ yr^{-1} increases dramatically from 5×10^{-7} Mpc^{-3} at $z = 0$ to 5×10^{-5} Mpc^{-3} at $z = 2$. At a given redshift, the distribution of $\dot{M}_b$ also broadens with increasing halo mass, although the distribution (and dispersion) of the ratio $\dot{M}_b/M_b$ is largely independent of mass. The latter is similar to the weak mass dependence of the mean $\dot{M}/M$ given by equation (3.8).

3.5 Formation Redshifts: Mean and Dispersion

It is well established that on average, more massive haloes form later than less massive haloes in the ΛCDM cosmology. The Millennium database provides sufficient statistics for us to quantify the distributions of the formation redshift z_f and its mean *and* scatter over a wide range of halo masses ($\sim 10^{12}$ to $\sim 10^{15} M_\odot$). The formation redshift, along with the late-time growth rate $\beta - \gamma$, can be thought of as two physically motivated quantities parameterizing the halo MAH.

The distributions of z_f for each type of MAHs for three halo mass bins are plotted in Fig. 3.7. The overall trend of decreasing z_f with increasing halo mass is evident across the three panels. Within each panel, a correlation between z_f and the MAH type is clearly seen. Nearly all haloes in the smallest few z_f bins are Type II. This means that despite the fact that Type II dominates the highest mass bins, the haloes that constitute Type II are not merely the especially massive haloes which formed late, but also include less massive haloes

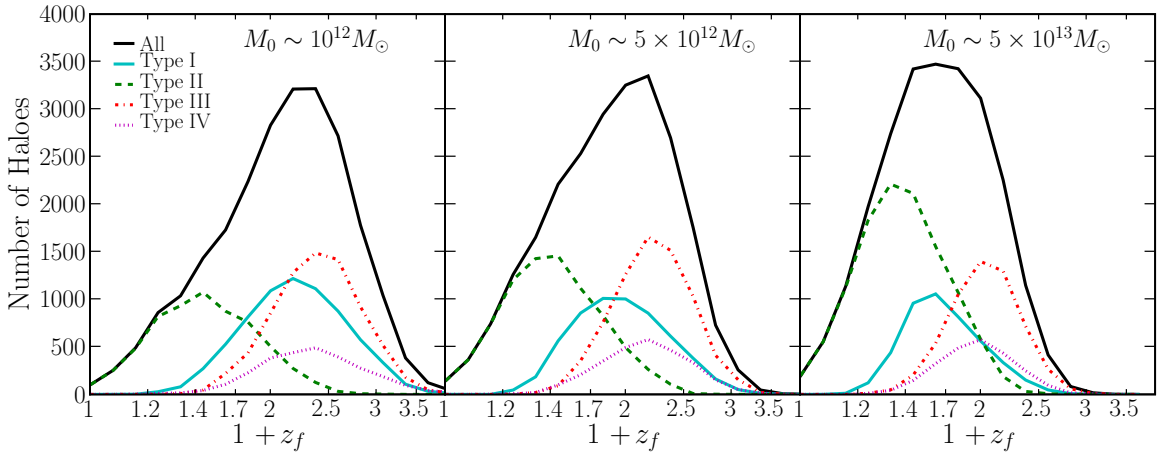

Figure 3.7: Distributions of the formation redshift z_f for three mass bins (left to right). Within each mass bin, the z_f distribution for all haloes is plotted (solid black), as well as the distribution for each halo type. While the relative amplitudes of the distributions do change from one mass bin to another, the overall shapes remain similar across all masses.

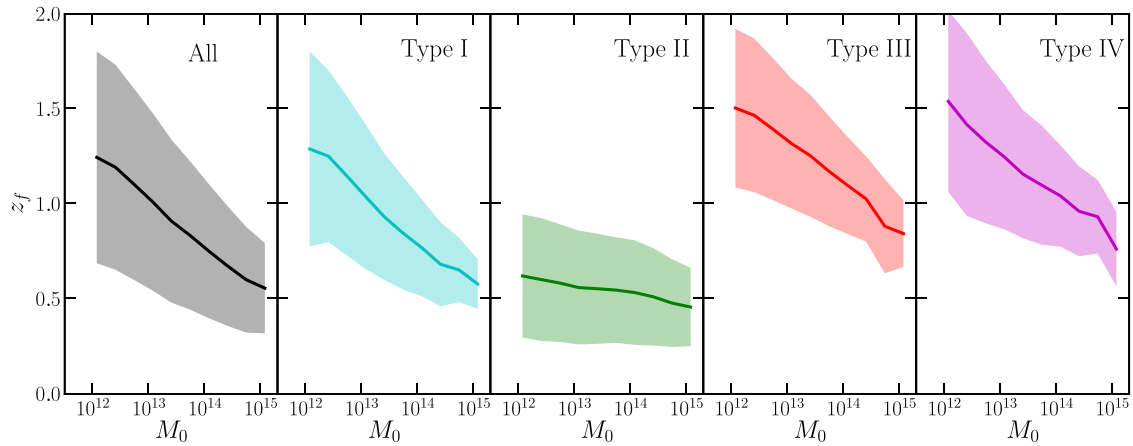

Figure 3.8: Mass dependence of the mean (solid curves) and one standard deviation scatter (shaded regions) of the formation redshifts of the Millennium haloes. More massive haloes on average form more recently, but the scatter is large. One exception is Type II haloes that have a mean z_f of ≈ 0.5 independent of mass. Fits to the mean and scatter of z_f as a function of mass are given in Table 3.3.

	z_f vs. M_0		$\gamma - \beta$ vs. M_0	
	$\langle z_f \rangle$	σ_{z_f}	$\langle \gamma - \beta \rangle$	$\sigma_{(\gamma-\beta)}$
Overall	$-0.24x + 1.26$	$-0.11x + 0.58$	$-0.25x - 0.29$	$0.14x + 0.64$
Type I	$-0.25x + 1.30$	$-0.12x + 0.51$	$-0.22x - 0.45$	$0.03x + 0.21$
Type II	$-0.05x + 0.62$	$-0.04x + 0.34$	$-0.12x - 1.15$	$0.14x + 0.60$
Type III	$-0.23x + 1.56$	$-0.08x + 0.43$	$-0.03x - 0.20$	$0.01x + 0.11$
Type IV	$-0.23x + 1.56$	$-0.08x + 0.43$	$0.03x + 0.15$	0.36

Table 3.3: Linear fits for the mass dependence of the mean formation redshift z_f, mean $\gamma - \beta$, and their respective 1σ scatter about the mean, where $x \equiv \log_{10}(M_0/10^{12} M_\odot)$.

which formed late. On average, a Type II halo has a formation redshift 0.5 smaller than a typical halo. To a lesser degree, Type III and Type IV are also distinct from the overall distribution. Both tend to form early, Type III more so than Type IV, and together the two types account for most of the haloes that formed early.

Fig. 3.8 shows the mean formation redshift z_f as a function of halo mass for all 478,781 $z = 0$ Millennium haloes (leftmost panel) as well as for each type of MAH. As the scatter about the line is, to a good approximation, Gaussian, the 1σ range about the line is also provided in the plots in each panel (light shaded areas). From these shaded areas, it is clear that there is considerable scatter for the overall distribution. The relationship between M_0 and z_f is different from the overall distribution for all types except Type I, which suggests that the types discriminate by formation redshift to some extent. Also note that the separation of haloes into types also produces more limited scatter about the mean.

To approximate the mass dependence of the mean and scatter of z_f, we use the linear form

$$\langle z_f \rangle = a \log_{10} \frac{M_0}{10^{12} M_\odot} + b \,, \quad \sigma_{z_f} = c \log_{10} \frac{M_0}{10^{12} M_\odot} + d \tag{3.10}$$

and find it to fit the simulation data accurately. Table 3.3 lists the best-fit coefficients for all the halo MAHs (above 1000 particles at $z = 0$) and for each of the four types of MAHs shown in Fig. 3.8. Table 3.3 also includes the same fit performed for the fit parameters $(\beta - \gamma)$. The mean formation redshifts differ significantly among the types, with $\langle z_f \rangle \approx 0.6, 1.3$, and 1.5 for Type II, I, and III (plus IV), respectively, for galaxy-size haloes. The dependence of $\langle z_f \rangle$ on mass is noticeably weak for Type II; the other types show similar mass dependence, where $d \langle z_f \rangle / d \log M$ ranges from -0.23 to -0.25.

With the relationships between formation redshift and mass for each type, we can look at how these dependences relate to the basic halo characteristics given in Table 3.1. Recall that Type II haloes were marked by steep growth at late times, which is captured by the very negative value of $\beta - \gamma$. Type III haloes, on the other hand, have small values for $\beta - \gamma$, and thus grow slowly at late times. The relationships shown in Fig. 3.8 are then no surprise. Type II haloes are also associated with late formation times, while Type III haloes tend to have formed quite early.

3.6 Correlations with Halo Environments and Major Merger Frequencies

Thus far we have discussed how the halo MAHs and mass accretion rates vary with halo mass and redshift. We have also shown that the mean z_f depends on halo mass strongly, but the scatter in z_f does not depend strongly on the MAH type nor halo mass. In this section, we investigate if the mean and scatter in z_f are correlated with quantities other than halo mass. In particular, we ask if the shapes of MAHs (1) differ systematically between underdense vs overdense regions, and (2) are correlated with the time and frequency of major mergers and mass brought in by these events during a halo's lifetime.

3.6.1 Environment

An extensive discussion and tests of halo environments can be found in Fakhouri and Ma [2009]. Four definitions of a halo's local environment based on the local mass density centered at the halo were compared. Three of them were computed using the dark matter particles in a sphere of radius R centered at a halo, either with or without the central region carved out; the fourth definition was computed using the masses of only the haloes rather than all the dark matter. Here we use $\delta_{R-\mathrm{FOF}}$, computed by subtracting out the FOF mass M of the central halo within a sphere of radius R:

$$\delta_{R-\mathrm{FOF}} \equiv \delta_R - \frac{M}{V_R \bar{\rho}_m}\,, \tag{3.11}$$

where V_R is the volume of a sphere of radius R, and δ_R is the mean mass overdensity within R. This measure makes no assumption about the central halo's shape. By taking out the mass of the central halo itself, this density was shown to be a more robust measure of the environment *outside* of a halo's virial radius. Otherwise, the halo mass itself dominates the density centered at massive haloes (see Fig. 1 of Fakhouri and Ma 2009), and it becomes difficult to distinguish whether any correlations are due to the mass or the larger-scale environment in which the halo resides.

Fig. 3.9 shows the distribution of the environmental densities (evaluated at $z = 0$) for haloes within each MAH type and the total population. The distribution for Type IV haloes is quite distinct from all other distributions and is offset towards higher densities. Since Type IV haloes experience very little mass growth or even mass loss at late times, denser environments appear to impede mass accretion onto haloes.

3.6.2 Mass Growth due to Major Mergers

Major mergers are more rare than minor mergers, but they can contribute to a significant fraction of a halo's final mass, and have a strong impact on halo structures and galaxy properties such as the star formation rate.

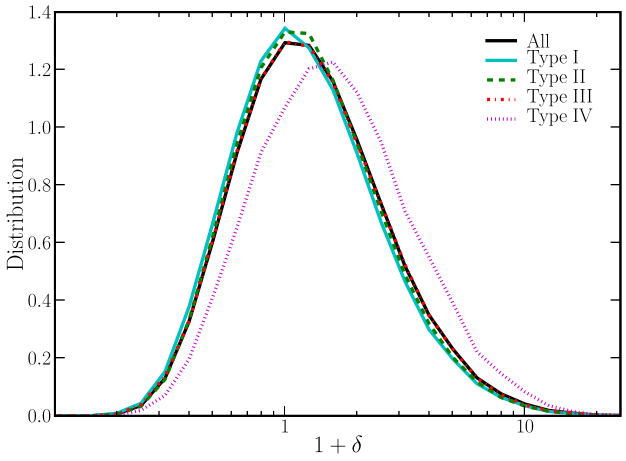

Figure 3.9: Normalized distributions in δ are plotted for all haloes, as well as each type of halo. The most prominent feature is the clear separation between Type IV and all other types. The center of the Type IV is well to the right of the other types, meaning that Type IV haloes are predominantly found in denser environments than any other type.

A useful quantity for assessing the role of major mergers on halo MAHs is z_{lmm}, the redshift of the last major merger in a halo's history. Fig. 3.10 shows the fraction of haloes whose last major merger occurred at or before z. Type II haloes are seen to experience a major merger in the much more recent past than the other types: about 65% of them had encountered a major merger within redshifts 0 and 0.3, and only 25% of them had their last major merger before $z = 1$. In sharp contrast, only about 5% of Type III and IV haloes had a major merger later than $z = 0.3$, and over 75% of them had their last major merger earlier than $z = 1$.

Another useful parameter for quantifying the role of major mergers in its MAH is $F(\xi > \xi_{\rm min})$, which is the fraction of mass at z_0 that came from mergers above some progenitor mass ratio $\xi_{\rm min}$. We choose to define the mass ratio in relation to the mass of the progenitor at the time of merger, as opposed to being defined in relation to the halo's present mass. The exact value of F is strongly dependent upon the choice of $\xi_{\rm min}$. The overall features, however, do not change significantly, so different values for $\xi_{\rm min}$ only change the values for $F(\xi > \xi_{\rm min})$, but leave the overall characteristics in place.

Fig. 3.11 shows the differential (top) and cumulative (bottom) distribution of the major merger mass fraction, $F(\xi > 0.33)$ for each type of MAHs. Type II haloes (dashed curves), which feature steep growth at late times, are seen to have the highest F among the four types. The distribution peaks at $F \approx 0.4$, indicating that $\sim 40\%$ of their final mass was acquired through major mergers. Type I haloes (solid cyan curves), by contrast, feature a dearth of major mergers, which is unsurprising given the fact that large mergers are poorly handled by the simple exponential. Likewise, Type IV haloes (dotted curves), which grow quickly early on only to lose mass at late times, are dominated by major mergers with a number of Type IV haloes even having $F(\xi > \xi_{\rm min}) > 1$. This occurs when a halo has less

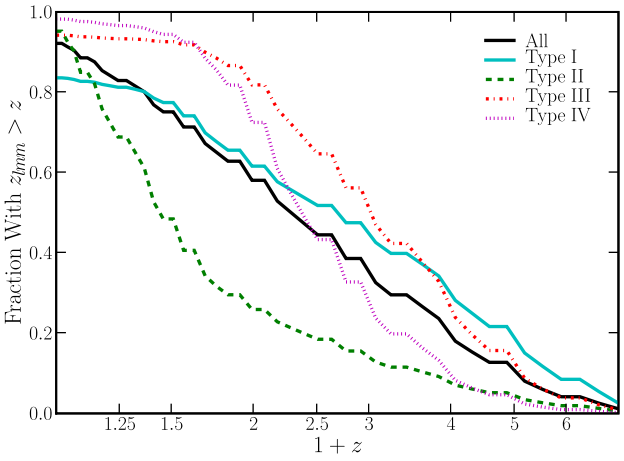

Figure 3.10: Distribution of the redshift of the last merger major for the four types of haloes (colored curves) and all $z = 0$ haloes (solid black). The majority of the $z = 0$ haloes with Type II MAH (dashed green curve) have experienced a major merger (of mass ratio > 0.33) very recently, whereas the last major merger occurred earlier than $z = 1$ for more than 75% of Type III and IV haloes.

mass presently than it has gained overall via major merger. Type III haloes (dotted dashed curves) tend to be relatively major merger free, which is again to be expected due to the decelerating growth of Type III haloes at later times during which much of the overall mass is accreted.

3.7 Conclusions

We have examined the mass growth histories of $\sim 500,000$ $z = 0$ haloes and their progenitors in the Millennium simulation. The two-parameter function in equation (3.4) provides a reasonable fit for the MAHs of these haloes, as shown in Fig. 3.2. The mean mass accretion rate of dark matter (and baryons) as a function of halo mass and redshift is well approximated by equation (3.8), as shown in Fig. 3.5. The distributions of $\dot{M}$ are broad, and the number density of high-$\dot{M}$ haloes increases sharply with increasing z at a given halo mass (see Fig. 3.6). The mean halo formation redshift as a function of mass is given by equation (3.10) and Fig. 3.8.

To facilitate the analysis of the halo MAH, we have classified $M(z)$ into four types based on their shapes. We have shown that only 20 to 30% of the Millennium haloes follow an exponential form ("Type I") in their mass accretion history $M(z)$. Only one parameter is needed to specify their MAH, e.g., the formation redshift z_f. The formation redshift depends strongly on halo mass, as expected for hierarchical cosmological models such as the ΛCDM. The median z_f ranges from 1.3 for $10^{12} M_\odot$ haloes to 0.6 for $10^{15} M_\odot$ haloes, and is dispersed over a range roughly equal to the median value for all masses.

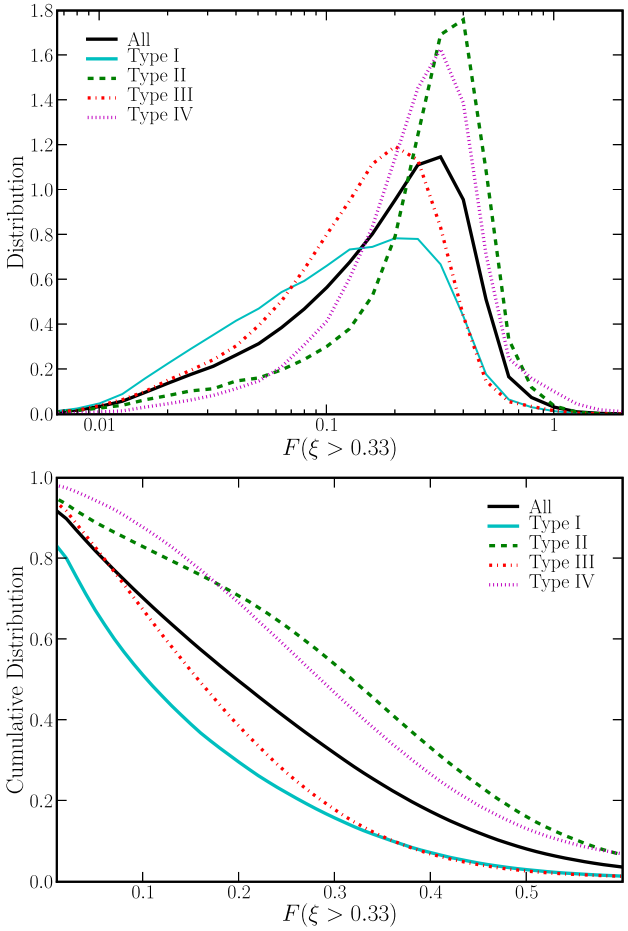

Figure 3.11: Differential (top) and cumulative (bottom) distributions of the fraction of halo mass gained from major mergers of mass ratio $\xi > 0.33$. Much like the z_f distributions, each Type is quite distinct from the overall distribution. Type I and III haloes feature few major mergers, while Type II and Type IV haloes tend to be dominated by major mergers. Type IV mergers also feature a noticeable rise in the highest bin, due to having gained more mass via major mergers across their history than they currently have.

About 20% of galaxy-size and 60% of cluster-size haloes have late-time growth that is steeper than an exponential ("Type II"). These haloes are formed more recently, with a median z_f of about 0.5 for all masses. The redshift at which they experience the last major merger is also significantly later than the exponential haloes: about 50% of them have had the last major merger between $z = 0$ and 0.3, as opposed to 10% of the rest of the haloes, including exponential haloes. Correspondingly, a higher fraction of Type II haloes' final mass is acquired through major mergers, e.g. 60% of these haloes obtained more than 30% of their final mass from major mergers, whereas a little over 30% of all haloes obtained more than 30% of their final mass from major mergers, and fewer than 20% of exponential haloes did.

The rest of the haloes have stunted late-time growth relative to an exponential form. The median z_f ranges from 1.5 at low mass to 0.8 at high mass. These haloes can be further separated into two groups (Type III and IV), where the two are primarily distinguished by the roles that major mergers have played in their growth; that is, Type III haloes tend to experience few major mergers, whereas Type IV haloes grew predominantly from major mergers at early redshifts. The MAHs of the two can be distinguished by the sharpness in the downturn of late time growth. Type IV haloes also live in somewhat denser environments, where the stronger tidal fields and more frequent interactions may have contributed to rapid accretion at early times followed by a slow down of their late time mass growth.

Despite this diverse behavior of halo MAHs, we have found the individual $M(z)$ to be well fit when a second parameter is introduced (eq. 3.4). To quantify the statistics of $M(z)$, we have provided fits to the joint probability distribution of the two MAH parameters β and γ in the Appendix. These can be used to generate realizations of halo mass growth histories in semi-analytic models of galaxy formation that incorporate realistic scatters about the mean trends.

3.8 Appendix: Joint Distribution of β and γ

We have seen that halo MAHs are well-fit by equation (3.4) with two parameters β and γ. Applying this fit to haloes in the Millennium simulation yields a joint distribution of β and γ. In this appendix we provide a fitting form to this distribution as a function of β, γ, and halo mass that is intended to allow the reader to generate rapidly a mock catalog of MAH tracks. We find that a straightforward rejection method can generate 300,000 mock MAH trajectories in under a minute on a standard laptop. The mean properties of the resulting trajectories match the mean properties of the Millennium trajectories at the 10% level. The fitting forms presented below are chosen for the practical purpose of matching the underlying distribution as closely as possible.

We find that 95.34% of the haloes occupy a smooth region in the (β, γ) plane shown in the left panel of Fig. 3.12. The remaining 4.66% of the haloes live along a distinct line with $\gamma = 0$ and $\beta \leq 0$, where the distribution of β is shown in the right panel of Fig. 3.12. That is, their MAHs are better approximated by a power law in $1 + z$ rather than an exponential.

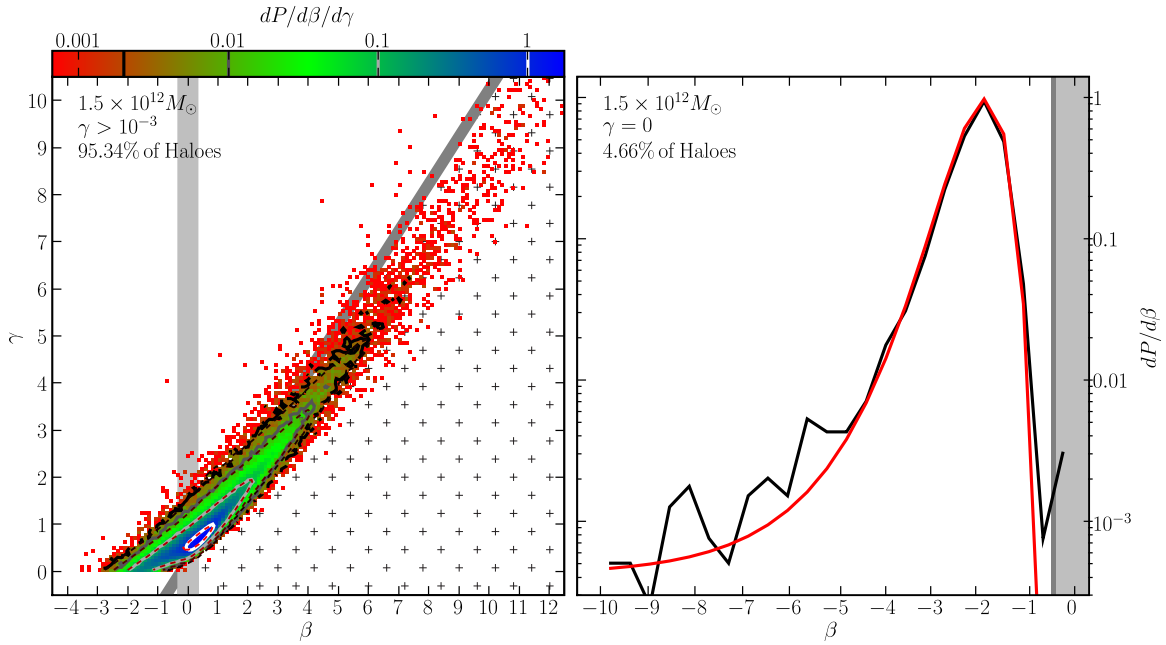

Figure 3.12: Probability distributions of β and γ for haloes of mass $1.5 \times 10^{12} M_\odot$. The left panel is the PDF of β and γ for haloes with $\gamma > 10^{-3}$ (95.34% of all haloes). The shaded 2D histogram presents the distribution of β and γ obtained from Millennium (see color scale for units) with corresponding contours drawn at the 0.005, 0.01,0.1 and 1 levels (black to white). The appropriately normalized fitting form D_{2D} is overlaid as red dashed contour lines. The right panel is the PDF of β for haloes with $\gamma = 0$ (4.66%). The appropriately normalized fitting form D_{1D} is overlaid in red. For both panels, the background shading corresponds to regions of phase space denote type I (grey), type II (red), type III (green) and type IV (blue) MAHs.

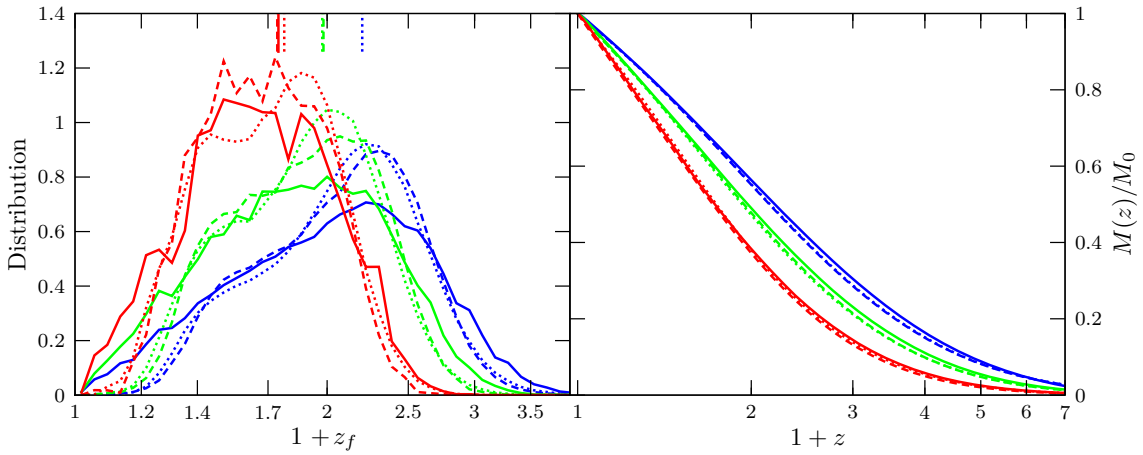

Figure 3.13: Distribution of the formation redshifts (left) and the mean MAHs (right) from the simulation directly (solid), the (β, γ) fits to the simulation MAHs (dashed), and a Monte Carlo ensemble of 300,000 halos per mass bin generated using equations (A1)-(A3) (dotted). In each panel, three mass bins are shown:$10^{12}M_\odot$ (blue), $10^{13}M_\odot$ (green), and $10^{14}M_\odot$ (red). Note that the dashed and dotted curves are almost indistinguishable in the right panel.

Interestingly, this 95.34% vs 4.66% division is independent of halo mass, even though the shape of the distributions generally depends on mass. For accuracy, we choose to separate the distribution of β and γ into these two components and fit to them separately.

For the 4.66% of haloes with $\gamma = 0$, their β distribution is well approximated by

$$\frac{dP}{d\beta} \propto e^{-X^2} \tag{3.12}$$

where

$$X = 7.443\, e^{0.6335\beta + 0.2626 M^{0.1992}} - 2.852 M^{-0.05412} \tag{3.13}$$

and $M \equiv M_{\rm halo}/10^{12}M_\odot$. This fit is valid in the range $-10 < \beta \leq 0$.

For the 95.34% of haloes with $\gamma > 10^{-3}$, the joint distribution in β and γ is well approximated by

$$\frac{dP}{d\beta d\gamma} \propto e^{-(XM^{-0.05569})^2 - (YM^{-0.05697})^2} \tag{3.14}$$

where

$$
\begin{aligned}
X &= (-1.722 - 0.1568\beta + 0.007592\beta^2)(1 - T_2) + \\
&\quad (1.242 + 0.3138\beta - 0.01399\beta^2)T_2 \\
Y &= 13.39[1-1.224\tanh(1.043Y')](1-0.08018\beta) \\
Y' &= \gamma - (28.85 + 0.4537\beta)(1 - T_1) - \\
&\quad (28.38 + 0.7624\beta)T_1 + 29.21M^{-0.001933} \\
T_1 &= 0.5[1 + \tanh(1.174\beta)] \\
T_2 &= 0.5[1 + \tanh(0.7671\beta - 0.1269)]\,.
\end{aligned}
$$

This is valid in the range $-8 < \beta < 12$, $0 < Y' < 3$. Since the rejection method does not require a normalized PDF for input, we leave these probability distributions unnormalized.

Fig. 3.13 illustrates that Monte Carlo realizations generated from the probability distributions above (dotted curves) reproduce accurately the formation redshift distributions (left panel) and the mean MAHs (right panel) obtained from the (β, γ) fits to the Millennium MAHs (dashed curves), and both match closely the results computed directly from the Millennium simulation (solid curves).

Part IV

The Environmental Dependence of Dark Matter Halo Growth

Chapter 4

Environmental Dependence of Dark Matter Halo Growth I: Halo Merger Rates

Fakhouri, Onsi; Ma, Chung-Pei — April 2009
Monthly Notices of the Royal Astronomical Society, Volume 394, Issue 4, pp. 1825-1840

In an earlier paper we quantified the mean merger rate of dark matter haloes as a function of redshift z, descendant halo mass M_0, and progenitor halo mass ratio ξ using the Millennium simulation of the ΛCDM cosmology. Here we broaden that study and investigate the dependence of the merger rate of haloes on their surrounding environment. A number of local mass overdensity variables, both including and excluding the halo mass itself, are tested as measures of a halo's environment. The simple functional dependence on z, M_0, and ξ of the merger rate found in our earlier work is largely preserved in different environments, but we find that the overall amplitude of the merger rate has a strong positive correlation with the environmental densities. For galaxy-mass haloes, we find mergers to occur ~ 2.5 times more frequently in the densest regions than in voids at both $z = 0$ and higher redshifts. Higher-mass haloes show similar trends. We present a fitting form for this environmental dependence that is a function of both mass and local density and is valid out to $z = 2$. The amplitude of the progenitor (or conditional) mass function shows a similar correlation with local overdensity, suggesting that the extended Press-Schechter model for halo growth needs to be modified to incorporate environmental effects.

4.1 Introduction

In studies of cosmological structure formation, the mass of a dark matter halo is a key variable upon which many properties of galaxies and their host haloes depend. For instance, dark matter haloes of lower mass are expected to form earlier on average than more massive haloes in hierarchical cosmological models such as ΛCDM. In semi-analytical modelling of

galaxy formation (see Baugh 2006 for a review), properties such as the formation redshift, halo occupation number, galaxy colour and morphology, and stellar vs AGN feedback processes are all assumed to depend on the mass of the halo (sometimes better characterised by the halo circular velocity).

In addition to the halo mass, however, recent work based on numerical simulations has shown that a halo's local environment also affects various aspects of halo formation. For instance, at a *fixed* mass, older haloes are found to cluster more strongly than more recently formed haloes [Gottlöber et al., 2001, Sheth and Tormen, 2004, Gao et al., 2005, Harker et al., 2006, Wechsler et al., 2006, Jing et al., 2007, Wang et al., 2007, Gao and White, 2007, Maulbetsch et al., 2007]. Other halo properties such as concentration, spin, shape, and substructure mass fraction have also been found to vary with halo environment (e.g., Avila-Reese et al. 2005, Wechsler et al. 2006, Jing et al. 2007, Gao and White 2007, Bett et al. 2007).

In contrast, no such environmental dependence is predicted in the extended Press-Schechter (EPS) and excursion set models [Press and Schechter, 1974, Bond et al., 1991, Lacey and Cole, 1993] that are widely used for making theoretical predictions of galaxy statistics and for Monte Carlo constructions of merger trees. The lack of environmental correlation arises from the Markovian nature of the random walks in the excursion set model. This limitation is not built into the model per se, but is an assumption stemming from the use of a tophat Fourier-space window function. When a Gaussian window function is used, for instance, Zentner [2007] finds an environmental dependence in the halo formation redshift, but the dependence is *opposite* to that seen in the numerical simulations cited above. Other attempts at incorporating environmental effects into the excursion set model thus far have not been able to reproduce the correlations in simulations (e.g., Sandvik et al. 2007, Desjacques 2008).

In this paper, we focus on the environmental dependence of the merger rate of dark matter haloes, a topic that has not been studied in detail. The merger rate is an important quantity for understanding and interpreting observational data on galaxy formation, growth, and feedback processes. While the mergers of galaxies and the mergers of dark matter haloes are not identical processes, the two processes are closely related, and quantifying the latter is the first key step in understanding the former. There have been few theoretical studies of merger rates (e.g., Gottlöber et al. 2001, Fakhouri and Ma 2008, Stewart et al. 2008) probably because mergers are two- (or more-) body processes, and a large ensemble of descendent haloes *and* their progenitor haloes must be identified from merger trees before the rate can be reliably calculated. In comparison, studies of halo properties such as the mass function, density and velocity profiles, concentration, triaxiality, spin, and substructure distribution require only the particle information from a single simulation output.

This paper is an extension of our earlier study Fakhouri and Ma [2008] (henceforth FM08). There we quantified the global mean merger rates of haloes in the Millennium simulation [Springel et al., 2005] over a wide range of descendant halo mass ($10^{12} \lesssim M_0 \lesssim 10^{15} M_\odot$), progenitor mass ratio ($10^{-3} \lesssim \xi \leq 1$), and redshift ($0 \leq z \lesssim 6$). We found that when expressed in units of the mean number of mergers *per halo* per unit redshift, the merger rate has a very simple dependence on M_0, ξ, and z: the rate depends very weakly on halo mass

($\propto M_0^{0.08}$) and redshift, and scales as a power law in the progenitor mass ratio ($\propto \xi^{-2.01}$) for minor mergers ($\xi \lesssim 0.1$), with a mild upturn for major mergers. These simple trends allowed us to propose a universal fitting form for the mean merger rate that is accurate to 10-20%.

Here we go beyond the global merger rate and use the rich halo statistics in the Millennium database to quantify the merger rate as a function of halo environment, in addition to descendant mass, progenitor mass ratio, and redshift. We also investigate the environmental dependence of the progenitor (or conditional) mass function. This quantity is closely related to the merger rate and is also the most important ingredient in the EPS and excursion set models for constructing Monte Carlo merger trees.

Several recent environmental studies have used halo clustering, quantified by the halo bias, as a measure of environment (e.g. Gottlöber et al. 2002, Sheth and Tormen 2004, Gao et al. 2005, Harker et al. 2006, Jing et al. 2007, Wechsler et al. 2006, Gao and White 2007). While halo bias is a powerful statistical quantity, we choose a simpler and more intuitive local environment measure and use the local mass density centred at each halo. The earlier studies that have used local overdensities as measures of halo environment have used a variety of definitions, e.g., the mean density within a sphere of some radius (ranging from 4 to $10\,h^{-1}$ Mpc) or within a spherical shell (e.g. between 2 and $5\,h^{-1}$ Mpc) [Lemson and Kauffmann, 1999, Harker et al., 2006, Wang et al., 2007, Maulbetsch et al., 2007, Hahn et al., 2008]. In this paper we compare different definitions of the local overdensity, both including and excluding the mass of the central halo itself.

In this paper we also provide an in-depth investigation of the effects of halo fragmentation on the merger rate and its environmental dependence. In FM08, we discussed how fragmentation is a generic feature of all merger trees and compared the *stitching* method with the conventional *snipping* method for handling these events. We will show here that fragmentation occurs more frequently in dense regions than in voids; understanding the effects of fragmentation on merger rates is therefore essential for obtaining robust results in dense environments. There are three general types of approaches to handling fragmentations: do nothing (*snipping*), *stitching* together fragmented haloes, or *splitting* up the common progenitor of the fragmented haloes. We will compare five algorithms for handling fragmentations based on these three approaches and show that except for one algorithm, all the algorithms give similar merger rates to within 20%.

This paper is organised as follows. In § 4.2 we briefly review how haloes and merger trees are constructed from the particle data in the Millennium simulation. Statistics detailing the distribution of halo mass at different redshifts are summarised in Table 4.1. In § 4.3 we compare four local density measures and their distributions in relation to halo mass. Three of the measures use the dark matter mass in a sphere centred at a given halo, either including or excluding the mass of the central halo. The fourth measure is motivated by observables such as luminosity-weighted galaxy counts and uses only the masses of the haloes within a sphere. § 4.4 contains the main results of this paper, where we quantify how the merger rate is amplified in denser regions and suppressed in voids for redshifts $z = 0$ to 2 over three decades of halo mass ($10^{12} - 5 \times 10^{15} M_\odot$). A simple power-law fitting function for this environmental

Redshift	Quantity	Mass Percentile 0-40%	40-70%	70-90%	90-99%	99-100%
	Number of haloes	192038	144028	96019	43208	4800
$z = 0$	Mass bins ($10^{12} M_\odot$)	$1.2 - 2.1$	$2.1 - 4.5$	$4.5 - 14$	$14 - 110$	> 110
	ν bins	0.75-0.81	0.81-0.92	0.92-1.11	1.11-1.63	1.63-4.30
	Number of haloes	188258	141194	94129	42358	4706
$z = 0.51$	Mass bins ($10^{12} M_\odot$)	$1.2 - 2.0$	$2.0 - 4.1$	$4.1 - 12$	$12 - 74$	> 74
	ν bins	0.95-1.03	1.03-1.15	1.15-1.37	1.37-1.93	1.93-4.66
	Number of haloes	172568	129426	86284	38827	4314
$z = 1.08$	Mass bins ($10^{12} M_\odot$)	$1.2 - 1.9$	$1.9 - 3.7$	$3.7 - 9.5$	$9.5 - 48$	> 48
	ν bins	1.22-1.32	1.32-1.46	1.46-1.70	1.70-2.27	2.27-4.73
	Number of haloes	116830	87622	58415	26286	2920
$z = 2.07$	Mass bins ($10^{12} M_\odot$)	$1.2 - 1.8$	$1.8 - 3.0$	$3.0 - 6.5$	$6.5 - 24$	> 24
	ν bins	1.74-1.86	1.86-2.01	2.01-2.27	2.27-2.85	2.85-5.20

Table 4.1: Halo mass bins and number statistics at redshifts $z = 0, 0.51, 1.08$ and 2.07 ($\Delta z = 0.06, 0.06, 0.09, 0.17$) from the Millennium simulation used in this paper. The bins are computed assuming fixed mass-percentile bins (header row). Listed are the number of haloes in each bin, and the corresponding mass and ν boundaries for each percentile bin. The highest 1% mass bins extend out to $5.2 \times 10^{15}, 3 \times 10^{15}, 1.3 \times 10^{15}$, and $4.4 \times 10^{14} M_\odot$ for $z = 0, 0.51, 1.08$ and 2.07 respectively.

dependence is proposed, which can be used in combination with the fit for the global rate presented in FM08. We also show that the progenitor (or conditional) mass function has a similar environmental trend as the merger rate. Even though this is expected given that the two quantities are closely related, this result demonstrates directly that the excursion set model is incomplete. In § 5, we present statistics of halo fragmentations, compare five algorithms for handling these events, and illustrate the robustness of the results reported in § 4. The Appendix provides a discussion of the self-similarity of the merger rate and its environmental dependence in the context of the choice of mass and environment variables used in the fitting formula.

4.2 Haloes in the Millennium Simulation

The Millennium simulation [Springel et al., 2005] follows the evolution of roughly 2×10^7 dark matter haloes from redshift $z = 127$ to $z = 0$ in a $500 h^{-1}$ Mpc box using 2160^3 particles of mass $1.2 \times 10^9 M_\odot$ (all masses quoted in this paper include the factor of h^{-1}). It assumes a ΛCDM model with $\Omega_m = 0.25$, $\Omega_b = 0.045$, $\Omega_\Lambda = 0.75$, $h = 0.73$ and an initial power-law distribution of density perturbations with index $n = 1$ and normalisation $\sigma_8 = 0.9$.

A friends-of-friends (FOF) group finder [Davis et al., 1985] with a linking length of $b = 0.2$ is used to identify haloes in the simulation. Each FOF halo (henceforth *halo*) thus identified is further broken into constituent subhaloes (each with at least 20 particles or $2.35 \times 10^{10} M_\odot$)

by the SUBFIND algorithm which identifies gravitationally bound substructures within the host FOF halo (for more on SUBFIND, see Springel et al. 2001a).

The subhaloes are connected across the 64 available redshift outputs to form a subhalo merger tree. Mergers are complicated processes and the particles in a given subhalo will not necessarily end up in a *single* subhalo in the subsequent output. As such, a subhalo is chosen to be the descendent of a progenitor subhalo at an earlier output if it hosts the largest number of bound particles in the progenitor subhalo. The resulting merger tree of the subhaloes can be used to construct the merger tree of the FOF haloes, although we have discussed at length in FM08 that this construction is non-trivial due to the fragmentation of FOF haloes. Our main results reported in Sec. 4 use the *stitching* tree of FM08. Since fragmentation occurs more frequently in denser environments, we provide a detailed comparison in Sec. 5 between stitching and four alternative algorithms to test the robustness of our results.

The Millennium database provides a number of mass measurements for each identified FOF halo. We use the total mass of the particles connected to an FOF by the group finder. Tinker et al. [2008] argue that spherical overdensity measures of mass are more closely linked to cluster observables than FOF measures and, therefore, are to be preferred. We have found, however, that the FOF mass definition is more robust in the context of merging haloes than definitions that make assumptions about halo geometry and virialization (for the simple reason that merging haloes are typically not virialized at the simulation outputs immediately preceding and following a merger event; see also White 2001).

We study the dependence of halo growth on halo environment in a variety of halo mass bins at different redshifts. Since the most massive haloes at $z = 0$ are more massive than the most massive haloes at higher redshifts, we use mass bins with boundaries that vary with redshift such that each mass bin contains a fixed percentage of haloes. Table 4.1 lists the five percentile bins used in our study, and the corresponding number of haloes and halo masses at $z = 0, 0.51, 1.08$, and 2.07. We note that even in the highest 1% mass bin, there are 4000 to 5000 cluster-mass haloes at $z \lesssim 1$ available for this study.

Table 4.1 also lists the range of ν for each mass bin, where $\nu = \delta_c(z)/\sigma(M)$ is often used as a mass variable for comparing haloes over different redshifts. Here $\sigma(M)$ is the variance of the linear density perturbations and $\delta_c(z)$ is the critical overdensity at redshift z, where $\delta_c(z) = 1.686/D(z)$ and $D(z)$ is the linear growth function of density perturbations in ΛCDM. A comparison of the FOF mass versus ν as the mass variable is provided in the Appendix.

Our notation is as follows. When computing the merger rate, we refer to the haloes at the lower redshift as the descendants and label their masses by M_0. The progenitors of a given descendant halo at a (slightly) higher redshift are labelled $M_1, M_2, M_3...$, where $M_1 \geq M_2 \geq M_3...$ by our convention. The mass ratio of the progenitors is defined as $\xi = M_{i\geq 2}/M_1$. In this paper we find that there are sufficient halo statistics from the Millennium simulation for studying the environmental dependence of the descendant haloes over a range of redshifts, and shall present results at $z_0 = 0, 0.51, 1.08$, and 2.07 and their progenitors at $z_1 = z_0 + \Delta z$, where $\Delta z = 0.06, 0.06, 0.09$, and 0.17, respectively.

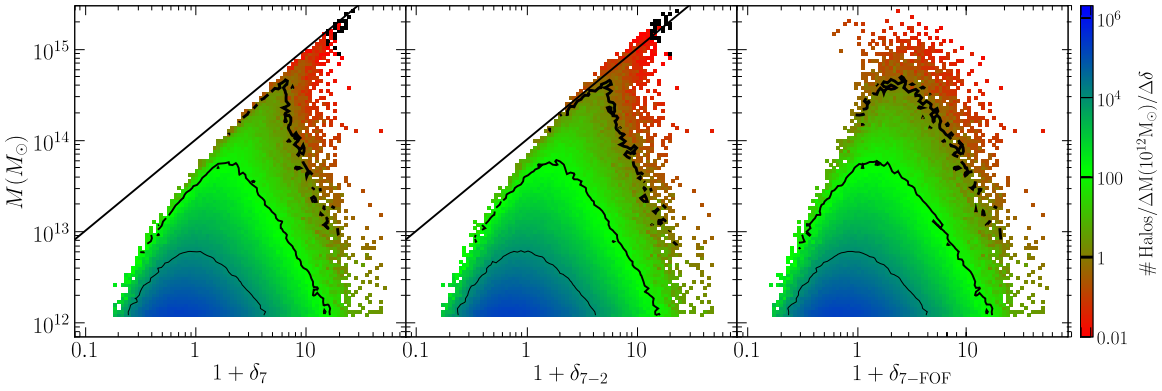

Figure 4.1: Scatter plots of halo mass vs three measures of halo environment for all FOF haloes above $1.2 \times 10^{12} M_\odot$ (1000 particles or more) in the $z = 0$ Millennium simulation output. The colour scale indicates the number of haloes present in each (δ, M) grid cell normalised by the bin size, and the contours are drawn at the 1, 100, and 10^4 bin levels (decreasing line width). The left panel uses $1 + \delta_7$, the density in a sphere of radius $7h^{-1}$ Mpc centred at each halo. The black line is $1 + \delta_7 = M/V_7/\bar{\rho}_m$ (see text). The middle panel shows $1 + \delta_{7-2}$, the density in a shell between $2h^{-1}$ and $7h^{-1}$ Mpc. The right panel uses $1 + \delta_{7-\mathrm{FOF}}$ by subtracting the halo mass from δ_7. At the high mass end, the halo itself is the main contribution to δ_7 and δ_{7-2}, leading to the tight correlation between δ and M in the upper right region in the left and middle panels. The right panels shows that this correlation is largely removed when $\delta_{7-\mathrm{FOF}}$ is used, which subtracts out the FOF mass of the central halo. The variable $\delta_{7-\mathrm{FOF}}$ is therefore a more independent measure of the immediate environment *outside* of the haloes.

4.3 Measuring Halo Environment

In this paper we quantify a halo's local environment using the local mass density centred at the halo. In this section we examine four definitions of density. Three of them are computed using the dark matter particles in a sphere of radius R centred at a halo, either with or without the central region carved out (see Sec 3.1-3.3). The fourth definition is computed using the masses of only the haloes rather than all the dark matter (Sec. 3.4). This last environmental measure based on mass-weighted halo counts has the advantage that it can be linked to observables such as luminosity-weighted galaxy counts.

4.3.1 Definitions of Environment

Only a few studies of halo environment have used local overdensities as measures of environment [Lemson and Kauffmann, 1999, Harker et al., 2006, Wang et al., 2007, Hahn et al., 2008]. By contrast, many studies have used the halo bias as a proxy for halo environment, which is obtained by taking the ratio of the halo-halo (or halo-mass) two-point correlation function to the underlying dark matter two-point correlation function (e.g., Gottlöber et al.

2002, Sheth and Tormen 2004, Gao et al. 2005, Harker et al. 2006, Jing et al. 2007, Wechsler et al. 2006, Gao and White 2007). Typically, these studies explore the dependence of bias on a variety of tracers of the halo growth history such as formation redshift, concentration, and number of major mergers. This technique has yielded clear signs of environmental dependence, particularly when combined with the marked correlation function statistical test [Gottlöber et al., 2002, Sheth and Tormen, 2004, Harker et al., 2006, Wechsler et al., 2006].

The connection between a halo's local density and the halo bias, however, is not entirely straightforward. The two quantities are certainly correlated, e.g., the two-point correlation function of objects in denser regions is typically higher than that in less dense regions [Abbas and Sheth, 2005]. However, the local density is a simple quantity that can be computed for each halo, whereas the bias is a statistical measure of clustering strength computed by averaging over a large number of pairs of haloes and particles over a range of pair separations.

The rich statistics of the Millennium simulation over large dynamic ranges in both mass and redshift make it possible to use the more intuitive local density as a measure of environment.

To compute the local overdensity in a halo's neighbourhood, we centre either a sphere or shell on the halo at spatial coordinates $\mathbf{x}$ and define the halo's environment by

$$\delta_R(\mathbf{x}) \equiv \frac{\rho_R(\mathbf{x}) - \bar{\rho}_m}{\bar{\rho}_m} \tag{4.1}$$

for a sphere of radius R, or

$$\delta_{R_o - R_i} \equiv \frac{\delta_{R_o} R_o^3 - \delta_{R_i} R_i^3}{R_o^3 - R_i^3} \tag{4.2}$$

for a shell of inner and outer radii R_i and R_o. Here $\bar{\rho}_m$ is the mean matter density in the simulation box, and $\rho_R(\mathbf{x})$ is the mean density of a sphere of radius R centred at $\mathbf{x}$. We also propose an environmental measure, $\delta_{R-\mathrm{FOF}}$, computed by subtracting out the FOF mass M of the central halo within a sphere of radius R:

$$\delta_{R-\mathrm{FOF}} \equiv \delta_R - \frac{M}{V_R \bar{\rho}_m}\,, \tag{4.3}$$

where V_R is the volume of a sphere of radius R. Note that unlike the shell measure, this measure makes no assumption about the central halo's shape.

To compute $\rho_R(\mathbf{x})$, one would need all the particle positions from the Millennium simulation, which are not available on the online public database. The database, however, does provide the density on a 256^3 cubic grid (with a grid spacing of $1.95\,h^{-1}$ Mpc) computed from the dark matter particles in the simulation using the Cloud-in-Cell (CIC) interpolation scheme. We use this data to sum up the contributions within the sphere centred at $\mathbf{x}$ to evaluate $\rho_R(\mathbf{x})$. We note that the grid in the database is indexed using a Peano-Hilbert space filling curve, which we have mapped to spatial coordinates in order to compute ρ_R.

It is important to choose an appropriate radius R (or $R_o - R_i$) in equations (4.1)-(4.3) when computing halo environment. Lemson and Kauffmann [1999] used both δ_{10} and δ_{5-2}

(in units of h^{-1} Mpc) but failed to detect any environmental dependence in the formation redshift for haloes with masses between 2×10^{12} and $10^{14}h^{-1}M_{\odot}$. Harker et al. [2006], following Lemson and Kauffmann [1999], used δ_{5-2} and *did* detect environmental dependence in Millennium for haloes with masses between 2×10^{12} and $10^{14}h^{-1}M_{\odot}$. Similarly, Hahn et al. [2008] used δ_5 and δ_{5-2} and found environmental dependence for haloes with masses between 2×10^{10} and $1.6 \times 10^{11}h^{-1}M_{\odot}$. We will show in §4.3.3 that $R = 7h^{-1}$ Mpc is an adequate choice that effectively characterises the environments of massive halos.

4.3.2 Disentangling Environment and Mass

Since the goal of this paper is to quantify the dependence of merger rates on halo environment as well as mass, it is essential for us to first examine the extent to which these two variables are independent measures of halo properties. This is particularly relevant considering that our measure of environment is based on the local mass density. We note that in the literature mass is often used loosely to refer to environment, e.g., clusters are considered denser environments than galaxies. This interpretation is valid for galaxy counts. We are concerned with FOF haloes (and not subhaloes or galaxies) here, however. As we will show, haloes of all masses can reside in a wide range of overdensities.

To study the relation between halo mass and environment, we present a scatter plot of the mass of every FOF halo (above 1000 particles $M > 1.2 \times 10^{12}M_{\odot}$) at $z = 0$ in the Millennium simulation versus its local $1+\delta$ in Fig. 4.1. Three definitions of local density are shown for comparison: all mass within a $7h^{-1}$ Mpc sphere (δ_7; left panel), all mass within $7h^{-1}$ Mpc excluding the central $2h^{-1}$ Mpc (δ_{7-2}; middle panel), and all mass within $7h^{-1}$ Mpc excluding the central FOF mass ($\delta_{7-\mathrm{FOF}}$; right panel).

Fig. 4.1 shows that galaxy-size haloes ($\sim 10^{12}M_{\odot}$) reside in a wide range of environmental densities from extreme underdense regions of $\delta \sim -0.8$ to regions with $\delta > 20$. The three panels show similar distributions of δ for these low mass haloes regardless of the definition of δ used. The high mass haloes, on the other hand, have very different distributions of δ. The rich statistics of the Millennium simulation allow us to study halo mass out to $2 \times 10^{15}M_{\odot}$, an order of magnitude higher than in previous studies. At $5 \times 10^{14}M_{\odot}$ and above, the spherical and shell measures of overdensity, δ_7 and δ_{7-2}, are seen to be tightly correlated with the halo mass (left and middle panels). In addition, all the points lie close to the line that represents the density in a $7h^{-1}$ Mpc sphere computed from the FOF halo mass *alone*, that is, $M/V_7/\bar{\rho}_m$, where V_7 is the volume of a sphere of radius $7h^{-1}$ Mpc. This trend clearly indicates that the central haloes are dominating the local overdensity at masses above $\sim 5 \times 10^{14}M_{\odot}$, and both δ_7 and δ_{7-2} are tracing the central halo mass rather than the overdensities in the neighbourhood *outside* the virial radius of the halo. Even though δ_{7-2} subtracts out the central $2h^{-1}$ Mpc regions, the tight residual correlation seen in the middle panel suggests that this quantity does not cleanly remove the contribution made by the central halo, probably because these massive haloes extend well beyond $2h^{-1}$ Mpc. We have also tested δ_{5-2}, the measure of environment used in Lemson and Kauffmann [1999], Harker et al. [2006], Hahn et al. [2007], and found nearly identical results as δ_{7-2}. This

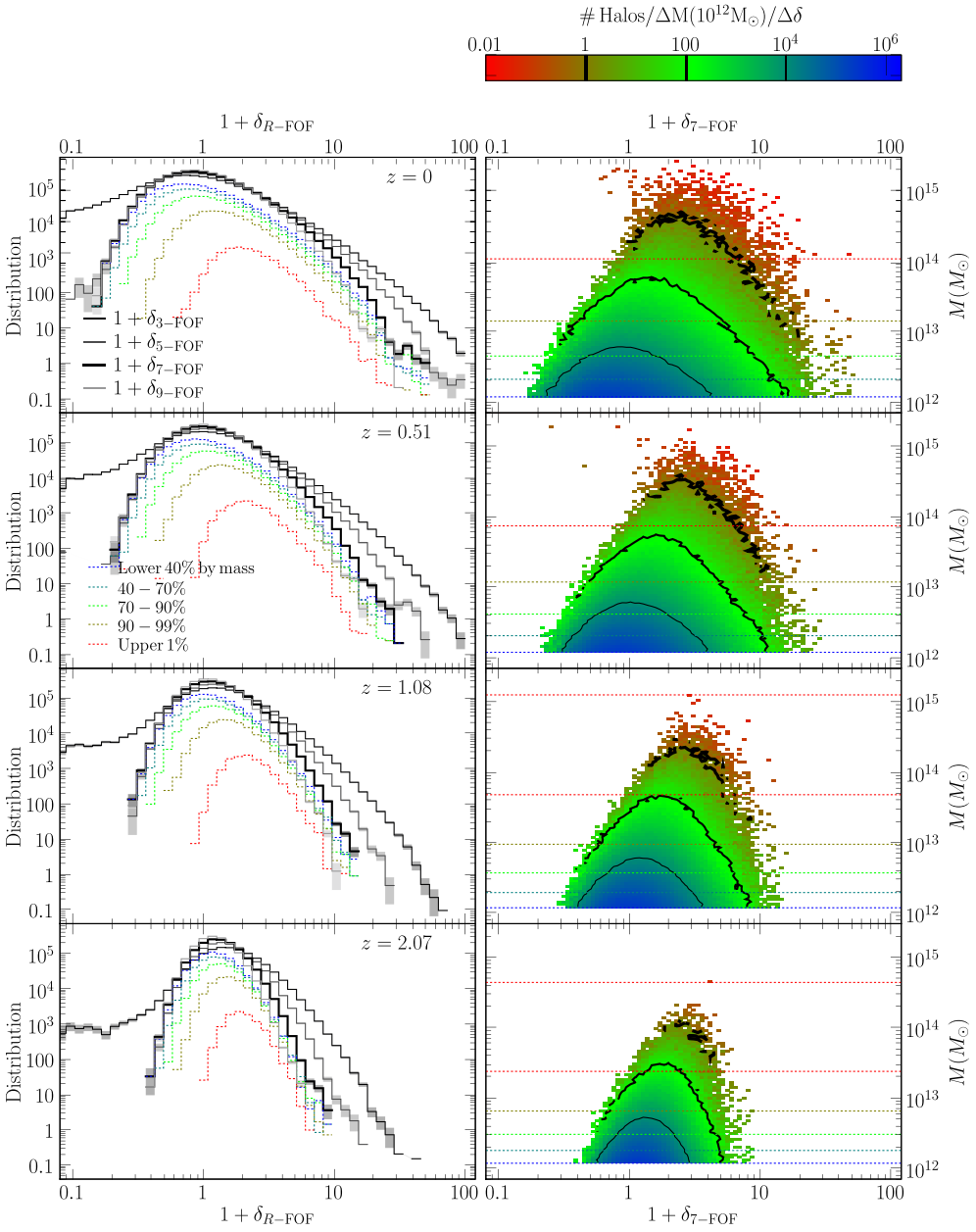

Figure 4.2: Left panels: Distribution of the environmental variable $1 + \delta_{R-FOF}$ defined in equation (4.3) at four redshifts $z = 0, 0.51, 1.08$ and 2.07 (top to bottom) computed from all haloes with $M > 1.2 \times 10^{12}$ (i.e. above 1000 particles) in the Millennium simulation. The broadening of the distribution with decreasing z illustrates the effect of gravitational instability. Within each panel, the four grey-scale histograms compare four smoothing radii R in h^{-1} Mpc: 3 (thin black), 5 (thin dark grey), 7 (thick black) and 9 (thin light grey); the five coloured dotted histograms compare the separate contributions to the δ_{7-FOF} distribution from haloes of different mass percentile bins: top 1% (red), 90-99% (olive), 70-90% (green) 40-70% (cyan), and bottom 40% (blue). Right panels: Similar scatter plot as the right panel of Fig. 4.1 but at four redshifts. The horizontal dotted lines mark the five mass percentile bins used in the left panels and listed in Table 4.1.

correlation may not be problematic for the results reported in these earlier papers, however, as these studies did not report results beyond $\sim 10^{14} M_{\odot}$.

The right panel of Fig. 4.1 shows that our third environmental variable, $\delta_{7-\mathrm{FOF}}$, in equation (4.3) is capable of disentangling the tight correlation between halo mass and density seen for δ_7 and δ_{7-2}; $\delta_{7-\mathrm{FOF}}$ is therefore a more robust measure of the environment *outside* of a halo's virial radius. It should, however, be kept in mind that haloes of different masses residing in the same 7 Mpc region will have the same δ_7 but different $\delta_{7-\mathrm{FOF}}$. A cluster-sized halo, for instance, will have a smaller value of $\delta_{7-\mathrm{FOF}}$ than a neighbouring galaxy-sized halo, and the difference between the two values of $\delta_{7-\mathrm{FOF}}$ will be the difference between the mass of the cluster and the galaxy (appropriately normalised). This caveat should be considered when interpreting values of $\delta_{7-\mathrm{FOF}}$ across different mass bins. The spherical measure δ_7, on the other hand, is simpler in this context. For this reason, we will report results using both $\delta_{7-\mathrm{FOF}}$ and δ_7 below.

4.3.3 δ Distributions

To gain further insight into the properties of the environmental measure $\delta_{R-\mathrm{FOF}}$ of equation (4.3), we plot in the left panels of Fig. 4.2 the distribution of $1 + \delta_{R-\mathrm{FOF}}$ centred at each halo for all haloes in the Millennium database with more than 1000 particles ($M > 1.2 \times 10^{12} M_{\odot}$) at four redshifts $z = 0, 0.51, 1.08$, and 2.07 (top to bottom). Within each panel, four choices of radii, $R = 3, 5, 7$, and $9\,h^{-1}$ Mpc, are shown for comparison (thin black, thin dark grey, thick black, and thin light grey). A comparison of the four left panels shows that the width of the $1+\delta_{R-\mathrm{FOF}}$ distribution becomes broader towards lower redshifts. This is a natural consequence of gravitational instability: denser regions become denser and vice versa as the universe evolves. We will explore the implications of this effect further in the appendix.

At a given redshift, as expected for a ΛCDM model, the overdensities computed using a larger smoothing radius R are generally smaller than those computed using a smaller R. In the voids, the distribution of $1 + \delta_{3-\mathrm{FOF}}$ is seen to have a low δ tail that extends down to unphysical (negative) $1 + \delta$; a faint remnant of this tail is also visible in $1 + \delta_{5-\mathrm{FOF}}$ at $z = 0$. This tail is due to a number of cluster-size haloes whose FOF member particles extend beyond 3 to 5 Mpc. We note that the virial radii of even the most massive halos ($10^{15} M_{\odot}$) do not extend beyond 3 Mpc; however we have found that the distance between the centre of an FOF's most massive subhalo and the furthest subhalo associated with said FOF can extend beyond 5 Mpc even for halos with masses of a few $\times 10^{14} M_{\odot}$. We therefore use $1+\delta_{7-\mathrm{FOF}}$ throughout this paper in order to better sample the environment surrounding these large haloes.

We break down the haloes represented by the thick black $1 + \delta_{7-\mathrm{FOF}}$ curve in the left panels of Fig. 4.2 into different mass bins and plot their δ-distributions using dotted colour histograms. Less massive haloes cover a broader range of $1 + \delta_{7-\mathrm{FOF}}$ than more massive haloes, and their distribution peaks at a lower value of $\delta_{7-\mathrm{FOF}}$. We note that even though the histograms for both the lower mass haloes and the total distribution at $z = 0$ are peaked

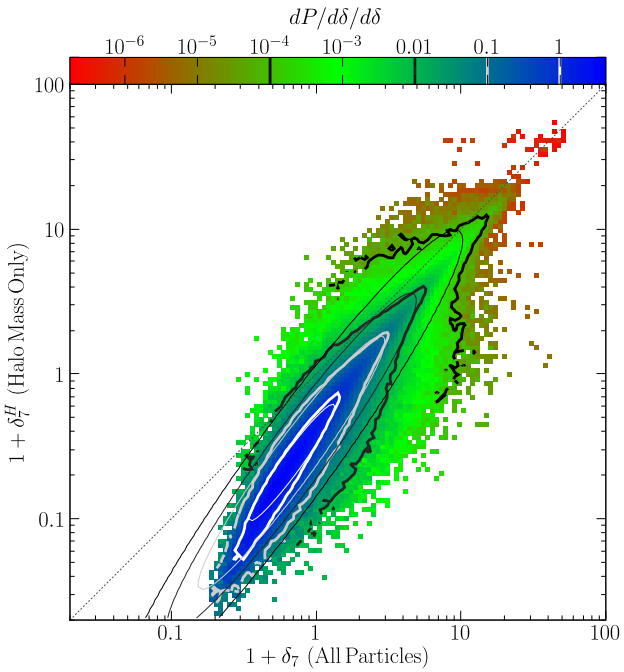

Figure 4.3: Scatter plot of two density variables: $1+\delta_7$ (see Fig. 1) computed from all dark matter particles centred within a $7h^{-1}$ Mpc sphere of each halo, and the mass-weighted halo counts $1+\delta_7^H$ computed by including only masses in haloes in the same sphere (down to halo mass of $1.2\times10^{12}M_\odot$). The contours are plotted at probability values of 10^{-4} (black), 0.01, 0.1 and 1 (white); the thick contours are for simulation data, the thin contours are from the fit in eq. (4.5). The grey dotted line is for $\delta_7=\delta_7^H$.

at a slightly negative value of δ, the mean value is in fact positive, e.g., $<\delta_7>=0.864$ and $<\delta_{7-\rm FOF}>=0.849$ for the lowest mass bin, and $<\delta_7>=1.04$ and $<\delta_{7-\rm FOF}>=0.956$ for all the halos. The value of $<\delta_{7-\rm FOF}>$ is only slightly smaller than $<\delta_7>$ because subtracting the FOF mass of the low mass haloes (which dominate the total distribution) makes little difference when δ is averaged over a sphere of radius as large as $7h^{-1}$ Mpc. The mean of δ is not zero here because the overdensities are not randomly sampled but are instead centred on haloes.

The right panels in Fig. 4.2 are scatter plots of each halo's local density $1+\delta_{7-\rm FOF}$ versus its mass at $z=0, 0.51, 1.08$, and 2.07 (top to bottom). (The top panel is a repeat of the right panel of Fig. 4.1.) The mass bins based on percentiles from Table 4.1 are marked by the horizontal lines. At high z the haloes cover a narrower range in both δ and M, but the tight correlation seen in Fig. 4.1 between the mass of the massive haloes and their local densities δ_7 and δ_{7-2} is removed at all redshifts when $\delta_{7-\rm FOF}$ is used.

4.3.4 Computing δ via Halo Counts

The local environmental measure δ_7 is convenient from a theoretical standpoint but is not easy to measure observationally as it demands accurate knowledge of the background dark matter distribution within a large (7 Mpc) radius of the halo in question. Here we consider a more observer-friendly quantity based on the mass-weighted halo counts (above a certain mass threshold):

$$1 + \delta_7^H(\mathbf{x}) \equiv \frac{\sum M_{\rm halo}}{V_7 \bar{\rho}_m} \tag{4.4}$$

where the sum is over all halos within a $7h^{-1}$ Mpc sphere centred at $\mathbf{x}$ above some minimum mass (we use 40 particles, or $4.7 \times 10^{10} M_\odot$), and V_7 is the volume of a sphere of radius $7h^{-1}$ Mpc. For a given halo in the Millennium simulation, we compute this quantity by summing over all haloes whose *centres* lie within the $7h^{-1}$ Mpc sphere centred on the halo in question. We do not account for the fact that halos near the boundary may only strictly contribute a fraction of their mass to the $7h^{-1}$ Mpc sphere.

The resulting mass-weighed halo counts $1 + \delta_7^H$ is plotted against $1 + \delta_7$ computed from the CIC density grid in Fig. 4.3. The 2d-histogram is normalised to have unit area and can be thought of as a bivariate probably distribution. We note that, while δ_7 is generally greater than δ_7^H as expected, there are regions with $\delta_7^H > \delta_7$, particularly in dense environments. This is due to the fact that a halo's entire mass contributes to δ_7^H if its centre lies within the $7h^{-1}$ Mpc sphere in question.

We approximate the distribution with a two-dimensional log-normal distribution. Since the variables are correlated, the fitting form has five parameters and is given by

$$\frac{dP}{d\delta_7 d\delta_7^H} = \frac{1}{2\pi x_1 x_2 \sigma_1 \sigma_2} \exp\left[-\frac{\ln(x_1)^2}{2\sigma_1^2} - -\frac{\ln(x_2)^2}{2\sigma_2^2} \right], \tag{4.5}$$

where $\ln(x_1)$ and $\ln(x_2)$ are uncorrelated variables that are simply linear combinations of $\ln(1 + \delta_7)$ and $\ln(1 + \delta_7^H)$ given by

$$\begin{bmatrix} \ln(x_1) \\ \ln(x_2) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \ln(1+\delta_7) - \mu_1 \\ \ln(1+\delta_7^H) - \mu_2 \end{bmatrix}. \tag{4.6}$$

Here μ_1 and μ_2 denote the mean values of $\ln(1+\delta_7)$ and $\ln(1+\delta_7^H)$ respectively, θ is an angle that quantifies the correlation between the two δs, and σ_1 and σ_2 are the standard deviations along the major and minor axes defined by θ. The best fit values for these five parameters are $\mu_1 = 0.210$, $\mu_2 = -0.549$, $\sigma_1 = 1.03$, $\sigma_2 = 0.141$, $\theta = 0.943$. The thin contours in Fig. 4.3 represent the resulting fit.

We have also computed a simpler power-law fit that can be used to approximate the mean relation between the two densities:

$$\ln(1 + \delta_7^H) = 1.28 \ln(1 + \delta_7) - 0.865\,. \tag{4.7}$$

Both fitting forms can be used to convert back and forth between δ_7 and δ_7^H.

4.4 Environmental Dependence

4.4.1 Halo Merger Rate

In FM08 we defined and computed the merger rate B/n as a function of progenitor mass ratio $\xi \equiv M_i/M_1$ (with $i \geq 2$), descendant mass M_0, and redshift z. The rate B/n is dimensionless and measures the mean number of mergers per halo per redshift interval per mass ratio. We found that in these units, the merger rate has a remarkably simple form and depends only weakly on mass and redshift. We proposed the fitting form

$$\frac{B(M_0,\xi,z)}{n(M_0,z)} = A\left(\frac{M_0}{\tilde{M}}\right)^{\alpha} \xi^{\beta} \exp\left[\left(\frac{\xi}{\tilde{\xi}}\right)^{\gamma}\right] \left(\frac{d\delta_c}{dz}\right)^{\eta}, \tag{4.8}$$

where $(\alpha,\beta,\gamma,\eta) = (0.083, -2.01, 0.409, 0.371)$, $A = 0.0289$, $\tilde{\xi} = 0.098$, $\tilde{M} = 1.2\times10^{12}M_{\odot}$, and $\delta_c(z) \propto 1/D(z)$ is the standard density threshold normalised to $\delta_c = 1.686$ at $z = 0$, with $D(z)$ being the linear growth factor. This fit is accurate to 10-20% over the mass range $10^{12} - 10^{15}M_{\odot}$ and redshift range $z < 6$.

The merger rates B/n at $z = 0$ for the five descendant halo mass bins in Table 4.1 are reproduced for reference in the top left panel of Fig. 4.4. As shown in FM08 and indicated by equation (4.8), the merger rate is approximately a power law in ξ in the minor merger regime and has a slight upturn in the major merger regime ($\xi \gtrsim 0.2$). All five curves are nearly on top of one another, reflecting the very mild mass dependence ($\alpha \sim 0.1$) in equation (4.8). We compute each curve by first selecting the descendant haloes in a given mass bin and computing the mass ratios ξ for the progenitors of these haloes. We then compute B/n by counting B, the number of progenitors (M_i with $i > 2$) that lie in a given mass ratio bin (ξ), and dividing by n, the total number of descendants in the mass bin in question. See FM08 for further details of this procedure and discussions of the results.

Equation (4.8) gives the global mean merger rate averaged over all halo environments. To investigate the correlation of B/n with environment, we divide each mass bin shown in the top left panel of Fig. 4.4 into five environmental bins and compute $B/n[\delta]$ using B *and* n in the given $\delta_{7-\mathrm{FOF}}$ bin. The remaining five panels in Fig. 4.4 show our results for the ratios of $B/n[\delta]$ to the global mean B/n, as a function of ξ, for each of the five descendant mass bins. Within each panel, the different curves are for different $\delta_{7-\mathrm{FOF}}$ bins for which there are sufficient halo statistics.

For descendant haloes of mass 10^{12} to $10^{13}M_{\odot}$, Fig. 4.4 shows a strong environmental effect with a *positive* correlation between merger rates and local density: haloes in the densest regions ($1 + \delta_{7-\mathrm{FOF}} > 7$; red curves) experience 1.5 to 2 times more mergers than the average, while haloes in underdense regions ($1 + \delta_{7-\mathrm{FOF}} < 0.7$; blue curves) experience fewer mergers (by a factor of 0.7 to 0.8) than average. For group and cluster scale haloes (10^{13} to $5\times10^{15}M_{\odot}$) in the bottom panels, only three curves are shown for the middle three $\delta_{7-\mathrm{FOF}}$ bins because massive haloes span a smaller range of $\delta_{7-\mathrm{FOF}}$, as shown in Figs. 1 and 2. For each $\delta_{7-\mathrm{FOF}}$ bin, the value of the merger rate ratio is quite similar for all five panels,

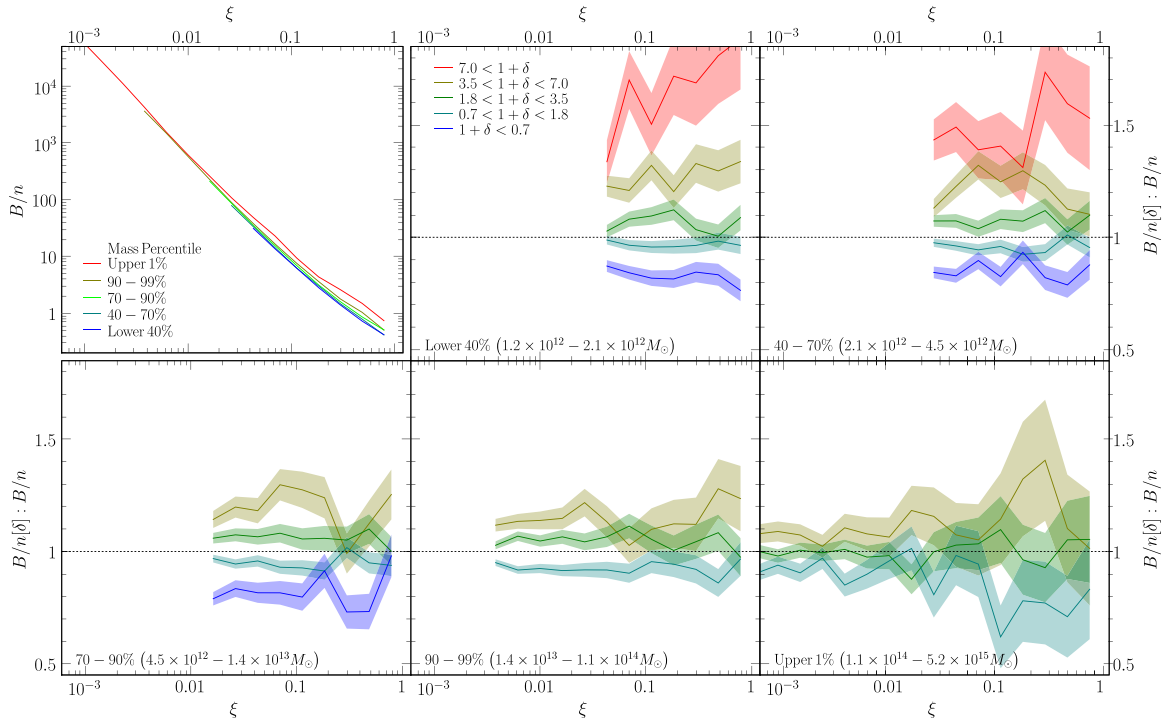

Figure 4.4: Halo merger rate and its environmental dependence on the local overdensity $\delta_{7-\mathrm{FOF}}$ measured in a $7h^{-1}$ Mpc sphere excluding the central FOF halo mass. Top left panel: The global mean merger rate B/n (in units of mergers per descendant halo per unit redshift per ξ bin) as a function of the progenitor mass ratio ξ for descendant haloes in five mass percentile bins (see Table 4.1). The results are computed using the $z = 0$ and 0.06 outputs from the Millennium simulation. The higher mass curves extend down to lower ξ because we have chosen a fixed minimal progenitor mass (40 particles) for all descendants. Other five panels: The ratio of the merger rate of haloes in a given environmental bin $B/n[\delta]$ to the global mean B/n as a function of ξ. Each panel is for a mass bin shown in the upper left panel. Within each panel, different colours show different $1 + \delta_{7-\mathrm{FOF}}$ bins (red for the densest regions; blue for the void regions), and the bands indicate the size of the Poisson errors. Note that the lower panels for the higher mass haloes have fewer δ curves since $\delta_{7-\mathrm{FOF}}$ for these haloes spans a narrower range. This figure clearly shows that the merger rate is higher in dense regions and lower in voids for all halo masses, and the boost or reduction factor is nearly independent of the progenitor mass ratios ξ.

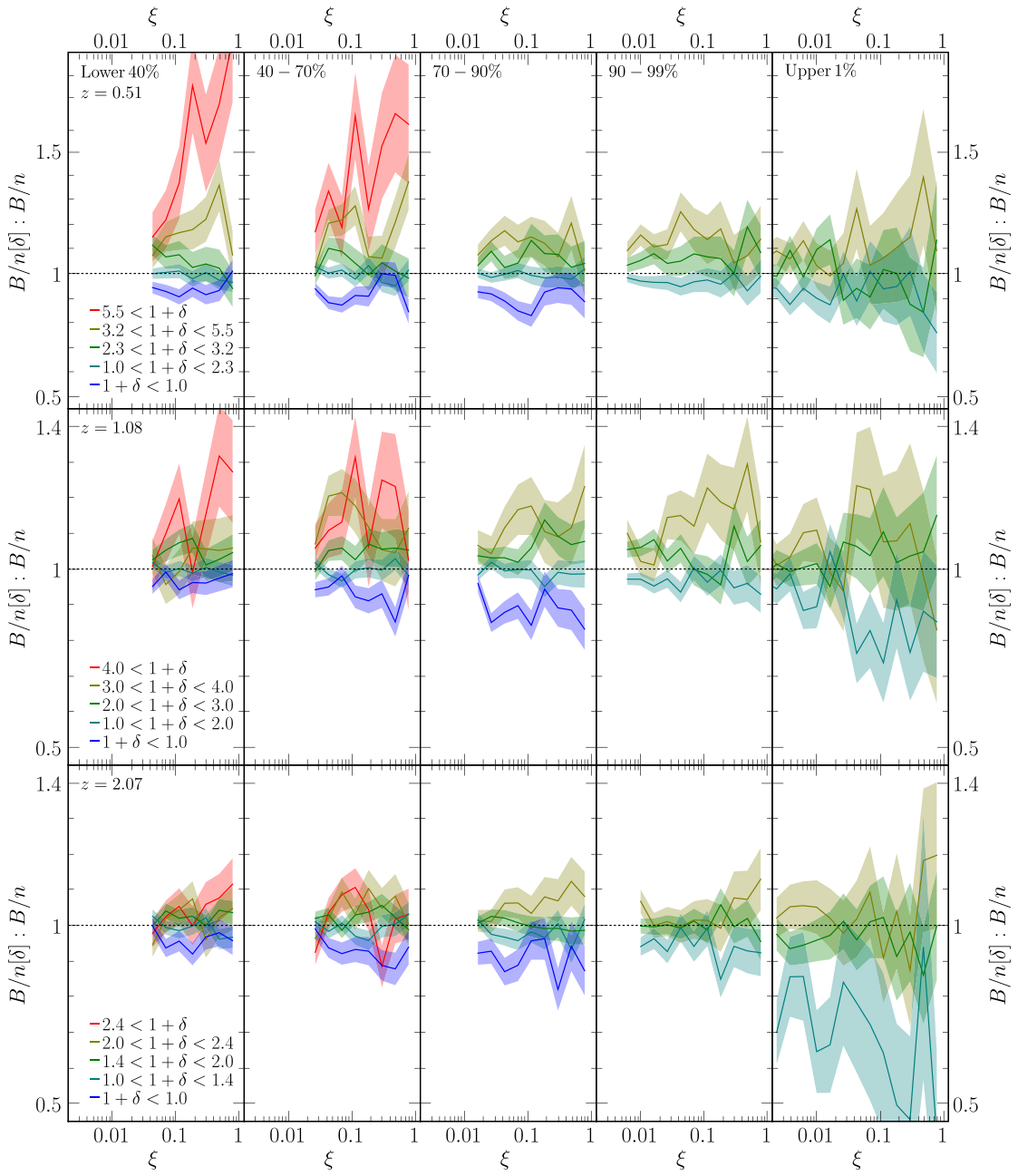

Figure 4.5: Same as the merger rate ratio plots in Fig. 4.4 except at higher redshifts: $z = 0.51, 1.08, 2.07$ (from top to bottom). The five columns correspond to the five mass percentile bins (see Table 4.1). Within each panel, the coloured curves show different $1 + \delta_{7-\mathrm{FOF}}$ bins (red for the densest regions; blue for the void regions), and the bands indicate the size of the Poisson errors. Note that since the distribution of $\delta_{7-\mathrm{FOF}}$ evolves with z, the corresponding δ bins (labelled in the leftmost columns) for the five coloured curves change with redshift. This figure shows that the positive correlation of the merger rate with $\delta_{7-\mathrm{FOF}}$ persists out to $z \approx 2$.

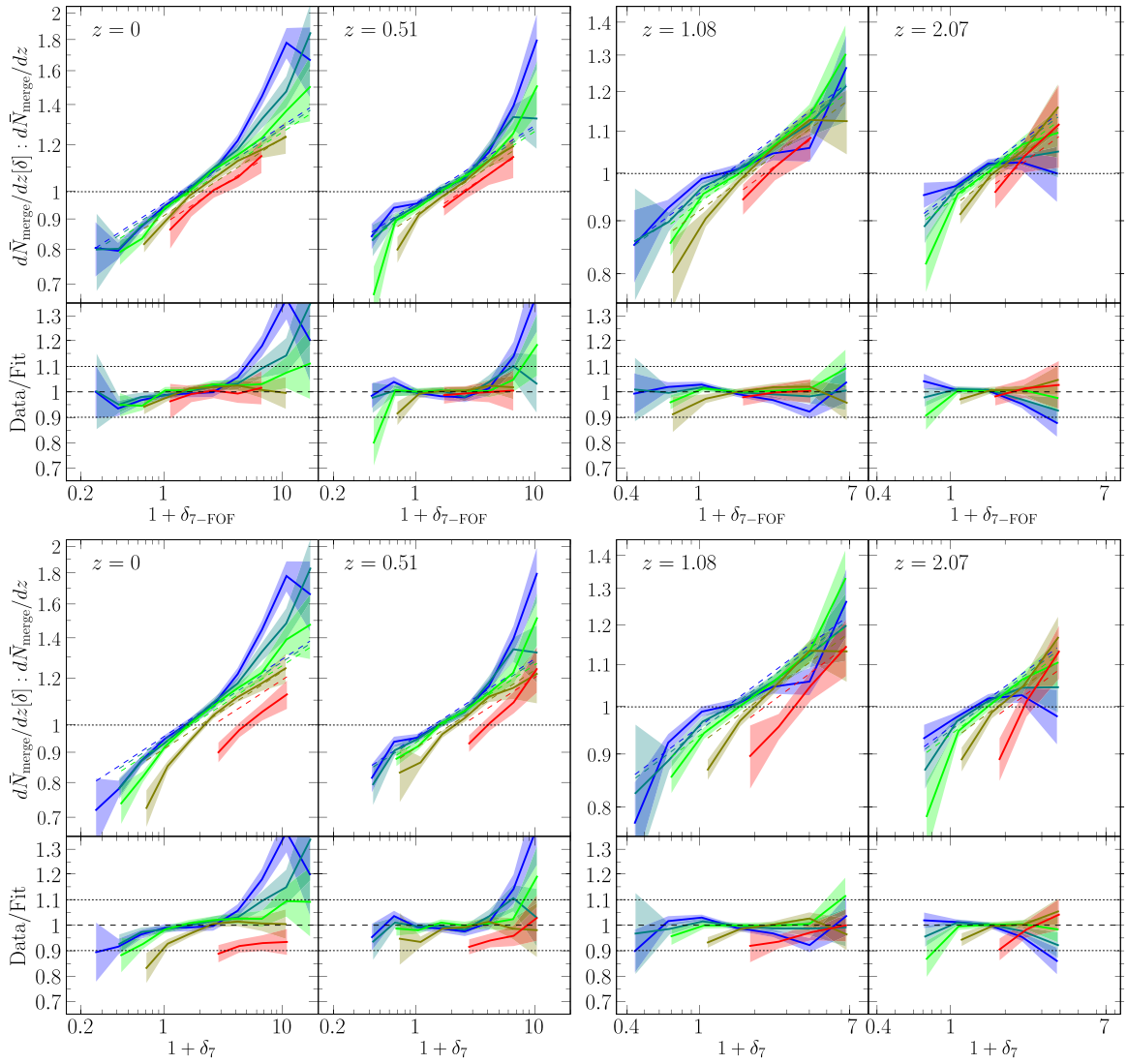

Figure 4.6: Dependence of the mean merger rate $dN_{\rm merge}/dz$ $(= \int B/nd\xi)$ on environmental variables $1 + \delta_{7-\rm FOF}$ (top figure) and $1 + \delta_7$ (bottom figure) at four redshifts $z = 0, 0.51, 1.08$,and 2.07 (left to right). Within each figure, the top panel shows the ratio of the mean merger rate $dN_{\rm merge}/dz[\delta]$ for haloes in a given environment to the global mean merger rate $dN_{\rm merge}/dz$. The bottom panel plots the ratio of the simulation results to the fits, showing that eq. (4.11) is generally accurate to within 10% (indicated by the dotted horizontal line). The colours correspond to the five mass percentile bins in Table 4.1 (red for the highest and blue for the lowest mass bin); the bands correspond to Poisson errors. This figure shows that the positive correlation of the merger rate with environmental density is present at all mass and redshift ranges probed.

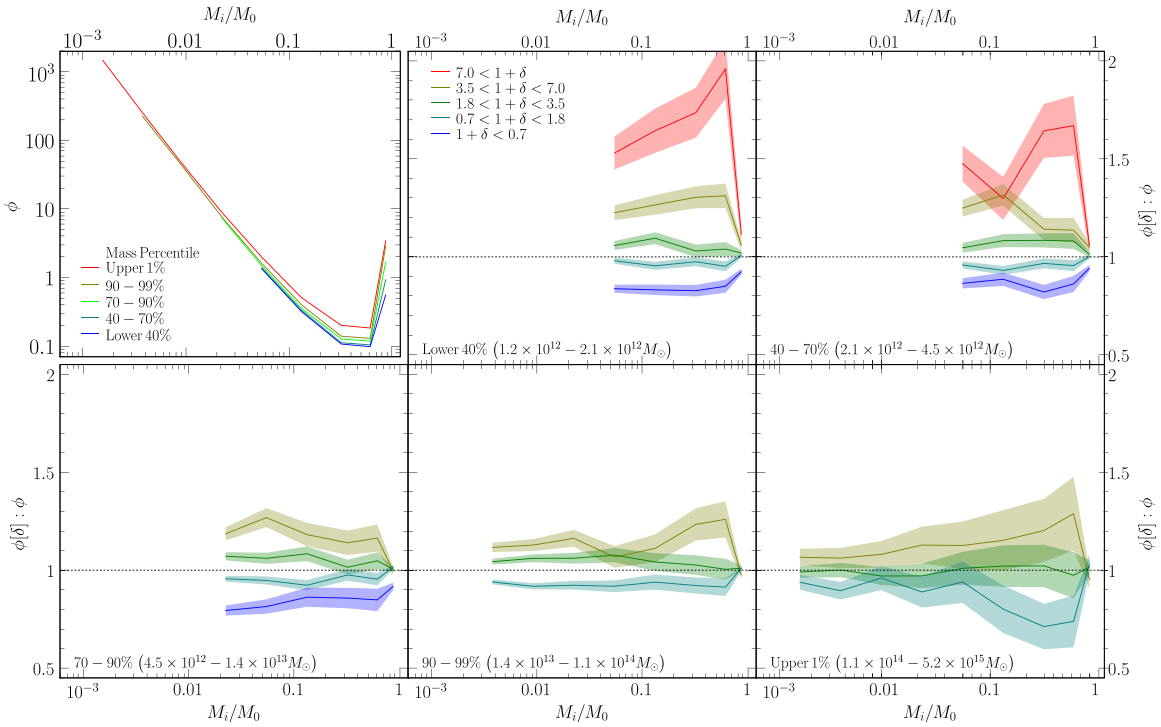

Figure 4.7: Same as Fig. 4.4 except the quantity shown is the progenitor (or conditional) mass function $\phi(M_1, z_1|M_0, z_0)$ instead of the merger rate B/n. The results are computed using the $(z_0, z_1) = (0, 0.06)$ outputs from the Millennium simulation. Similar to Fig. 4.4, we find that descendant haloes in the densest regions (red curves) have significantly more progenitors (by a factor of 1.5 to 2) than the global distribution of progenitors, whereas those in the voids (blue curves) have $\sim 20\%$ to 30% fewer progenitors.

indicating that the environmental effect, as measured by $\delta_{7-\mathrm{FOF}}$, depends very weakly on halo mass. We will quantify this statement using a fitting formula below.

Fig. 4.5 presents the same information as Fig. 4.4 but at higher redshifts ($z = 0.51, 1.08$, and 2.07 from top to bottom panels). The δ-bins for which the curves are too noisy are excluded. Since the distribution of $1 + \delta_{7-\mathrm{FOF}}$ narrows with increasing z, the $\delta_{7-\mathrm{FOF}}$ bins span a smaller range at $z = 2$ than at $z = 0$. Nonetheless, we see that the environmental dependence observed at $z = 0$ persists out to $z = 2$. Moreover, haloes with similar $1+\delta_{7-\mathrm{FOF}}$ experience similar amplifications or reductions in the merger rate regardless of mass and redshift.

An additional feature to note in Figs. 4.4 and 4.5 is that the curves are horizontal: environmental effect is therefore largely independent of the mass ratio ξ; that is, major and minor merger rates are boosted or dampened by a halo's environment by a similar factor. We can therefore integrate over the mass ratio parameter without diluting the environmental effect:

$$\frac{dN_{\mathrm{merge}}}{dz}(M, z) = \int_{\xi_{\mathrm{min}}}^{1} \frac{B(M, \xi, z)}{n(M, z)} d\xi \tag{4.9}$$

where dN_{merge}/dz is the mean merger rate per unit redshift per descendant halo with progenitor mass ratio above ξ_{min}.

The value of dN_{merge}/dz clearly depends on ξ_{min} and is larger when more minor mergers are included (see, e.g., Figs. 7 and 8 of FM08). For a fixed resolution mass (our choice is 40 particles or more for progenitor haloes), ξ_{min} extends down to lower values for higher mass descendants. For a fair comparison across halo mass bins, one should in principle use a fixed ξ_{min} for all mass bins at the expense of throwing out resolved progenitors for high mass descendant haloes. Since we plot ratios of the merger rates, however, the fact that more massive haloes are better resolved and have higher dN_{merge}/dz is normalised out. It is therefore possible to make a fair comparison across mass bins without throwing out any resolved progenitors.

Fig. 4.6 shows the ratio of the merger rate dN_{merge}/dz as a function of $1 + \delta$ at four redshifts ($z = 0, 0.51, 1.08, 2.07$ from left to right). For comparison, the results for two environmental measures are included: $\delta_{7-\mathrm{FOF}}$ (top figure) and δ_7 (bottom figure). Within each figure, the upper panel shows the simulation data and the lower panel compares the data to the analytic fitting formula discussed below. The five curves in each panel are for the five mass bins listed in Table 4.1. This figure shows the same trend as Fig. 4.4: haloes in the densest regions at $z = 0$ experience up to ~ 1.5 times as many mergers as the average halo, whereas the merger rate in the voids is 20 to 30% below the global average.

To quantify the dependence of the merger rate on δ, we introduce

$$\frac{dN_{\mathrm{merge}}}{dz}(\delta, M, z) \approx \frac{dN_{\mathrm{merge}}}{dz}(M, z) \times f(\delta, M, z)\,, \tag{4.10}$$

where we have made use of the fact that the environmental dependence is independent of ξ to define f, and $dN/dz(M, z)$ is the global merger rate from FM08. We provide two fitting

forms for f using $\delta = \delta_{7-\rm FOF}$ and δ_7, respectively. We find that a simple power-law and redshift-independent form works well:

$$f(\delta_{7-\rm FOF}, M) = 0.963\,(1+\delta_{7-\rm FOF})^{0.130}\left(\frac{M}{10^{12}M_\odot}\right)^{-0.0156}$$
$$f(\delta_7, M) = 0.968\,(1+\delta_7)^{0.135}\left(\frac{M}{10^{12}M_\odot}\right)^{-0.0252}. \quad (4.11)$$

The fits are performed using data from all four redshifts simultaneously (over much finer mass bins than those shown in Fig. 4.6). The reduced χ^2_ν for the two fits is 0.95 and 1.08, respectively. Errors are computed assuming Poisson statistics and are represented by the filled regions in Fig. 4.6. The resulting fit is shown as dashed curves in the upper row of each figure, and the ratio of the simulation data to the fits is shown in the lower rows. The fits are seen to be accurate to within 10% for a wide range of δ and M, except for low mass haloes with $\delta \gtrsim 5$ at $z = 0$ and 0.51, where the rates steepen suddenly.

As we will discuss in § 4.5, our extensive tests using various algorithms suggest that the merger rate in this particular parameter range (i.e. low mass, low z, high density) depends sensitively on the post-processing algorithm used to handle fragmentations in the merger tree, and variations of order 20% or more among different algorithms are observed. We therefore do not attempt to use a fitting form more complicated than equation (4.11) to get a better fit in this uncertain regime.

It is interesting to note that when δ_7 is used, instead of $\delta_{7-\rm FOF}$, as the environment variable, the only change in the fit in equation (4.11) is a stronger dependence on halo mass. This trend makes sense since the difference between δ_7 and $\delta_{7-\rm FOF}$ is $\delta_7 - \delta_{7-\rm FOF} = M/(V_7\bar{\rho}_m)$ (see eq. [4.3]). This difference is negligible for galaxy-scale haloes (e.g. $\delta_7 - \delta_{7-\rm FOF} \sim 0.01$ for $10^{12}M_\odot$) but becomes larger for more massive haloes, reaching $\delta_7 - \delta_{7-\rm FOF} \sim 10$ at $M \sim 10^{15}M_\odot$. The five curves for the five mass bins at a given z in Fig. 4.6 are therefore more spread out when δ_7 is used as the variable, resulting in a stronger mass dependence.

We have chosen to use halo mass and local density as variables in equation (4.11). It is interesting to ask if other choices of variables may lead to a more accurate fit across the wide ranges of halo masses, densities, and redshifts shown in Fig. 4.6. For instance, the variance of the linear density perturbation $\sigma(M)$ and the scaled density threshold $\nu(M,z) = \delta_c/\sigma(M)D(z)$ are commonly used to characterise mass and redshift dependence of halo properties (e.g., the mass function). We test these variables and describe the results in appendix 4.7. Our conclusion is that these alternative variables do not perform any better, and the $[M, 1+\delta_{7-\rm FOF}]$ pair shows the least systematic variation with redshift.

In summary, the simple parametrisation of the environmental dependence of the merger rate given by equations (4.10) and (4.11) can be used along with the fit for the global merger rate B/n in equation (4.8) to compute the merger rate in different environments at a variety of redshifts.

4.4.2 Progenitor Mass Function

For completeness and ease of comparison with analytic models, we present here the results for the environmental dependence of the conditional (or progenitor) mass function $\phi(M_1, z_1|M_0, z_0)$. This function gives the mean distribution of the progenitor masses M_i at redshift z_1 for a descendant halo of mass M_0 at redshift z_0. It is the key ingredient for the construction of Monte Carlo merger trees in the Extended Press-Schechter model.

The relation between ϕ and the merger rate B/n is discussed in Sec 3.3 of FM08. These two quantities are closely related but differ in two ways. First, ϕ is typically plotted vs M_1/M_0, while B/n is expressed in the mass ratio of the progenitors $M_i/M_1(i \geq 2)$ and the descendant mass M_0. Second, the conditional mass function $\phi(M_1, z_1|M_0, z_0)$ includes all progenitor halos at z_1 regardless of if a merger has occurred between z_1 and z_0, whereas the merger rate includes only descendant haloes with more than one progenitor. When the lookback time $z_1 - z_0$ is small, a large fraction of haloes in fact have only one resolved progenitor typically with a mass M_1 comparable to the descendant mass M_0. See the sharp rise in ϕ near $M_1/M_0 = 1$ in Fig. 4.7). No such peak is present in the merger rate in Fig. 4.4.

Fig. 4.7 shows that the progenitor mass function has a similar dependence on $1 + \delta_{7-\mathrm{FOF}}$ as the merger rate in Figs. 4.4-4.6. We have chosen to plot Fig. 4.7 in the same way as Fig. 4.4, where the upper left panel shows the global progenitor mass function ϕ at $z_1 = 0.06$ for five bins of descendant mass M_0 at $z_0 = 0$, and the other five panels show how ϕ for haloes in different $\delta_{7-\mathrm{FOF}}$ bins compare to the global mean ϕ. We see that, like B/n, the progenitor mass function has a noted dependence on environment. For galaxy-size descendant haloes, those in the overdense regions have ~ 1.5 times as many progenitor haloes as the mean, while those in the underdense regions have ~ 0.7 times as many progenitor haloes as the mean.

4.5 Alternative Algorithms for Post-Processing Halo Fragmentations

As we discussed in Sec. 2 and FM08, even though each subhalo in the Millennium tree is, by construction, identified with a single descendant subhalo, the resulting FOF tree can contain fragmentation events in which an FOF halo is split into two (or more) descendant FOF haloes. This fragmentation issue is not unique to the use of subhaloes in the Millennium simulation but occurs in merger trees in all prior studies that are typically constructed based on the FOF haloes rather than subhaloes. This problem arises because particles in a progenitor halo (or subhalo) rarely end up in exactly one descendant halo; a decision must therefore be made to select a unique descendant and there is no unique way to do this. The standard procedure to assign progenitor and descendant FOF haloes is the same as that applied to the *subhaloes* in Millennium: the descendant halo is the halo that inherits the most number of bound particles of the progenitor. We call this algorithm *snipping* since it effectively cuts off the ancestral link between a progenitor halo and its *subdominant*

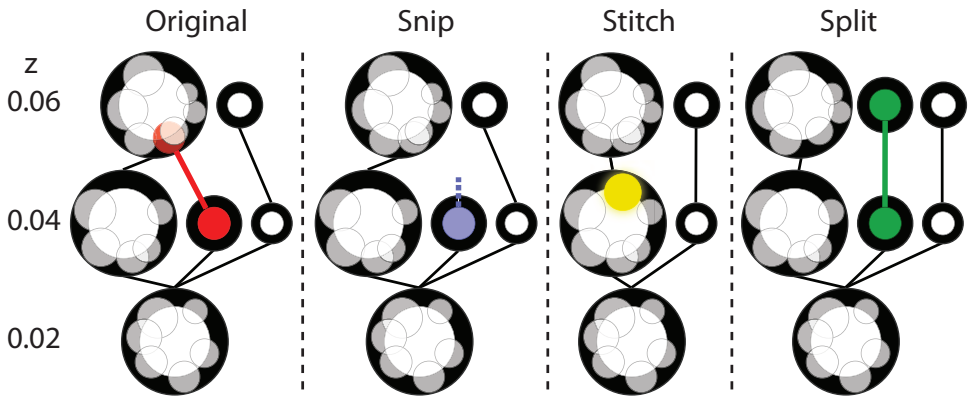

Figure 4.8: Example of a typical fragmentation event in the Millennium simulation. Black circles represent FOF haloes; white circles represent subhaloes. Circle radii scale with the logarithm of the (sub)halo mass. The left panel shows a fragmentation event occurring between $z = 0.06$ and 0.04 (red subhalo). The snip panel shows how the ancestral link between the fragmented halo and its progenitor is severed, producing a (blue) orphan halo. The stitch panel shows how the fragment is stitched back into the main branch at $z = 0.04$ (yellow subhalo). The split panel shows how the fragment's progenitor at $z = 0.06$ is split off from the FOF halo (green subhaloes).

descendant fragments, while leaving the halo masses unchanged (see Fig. 4.8).

In FM08, we explored a new method *stitching* for handling these fragmentation events. In this method (which we call *stitch-3* here), the fragmented haloes that remerge within 3 outputs after fragmentation occurs are stitched into a single FOF descendant; those that do not remerge within 3 outputs are snipped and become orphan haloes. We compared the two methods and showed that snipping inflates the merger rates by up to 10% in the major merger regime and 25% in the minor merger regime (Fig. 9 of FM08). This is not surprising since bound subhaloes are often on eccentric orbits that extend out to 2 to 3 virial radii of the main halo (see, e.g., Ludlow et al. 2008). The FOF finder can repeatedly disassociate and associate these subhaloes, leading to spurious fragmentation and remerger events.

In addition to the snipping and stitching algorithms, we examine a third method here that is complementary to stitching. We call this method *splitting* (see also Genel et al. [2009]). Our motivation for introducing this algorithm is the fact that fragmentations can be the result of either *false fragmentation* at the lower z_0, where physically bound subhaloes are broken up, or *false grouping* at the earlier z_1, where physically unbound subhaloes are falsely associated by the FOF finder. Even though multi-body subhalo encounters may unbind a subhalo, our visual inspections of a number of halo merger tracks indicate that such events are rare. Instead, most of the apparent fragmentations are due to the halo finder, which at an earlier output (z_1) may group subhaloes together, only to separate them at the next timestep ($z_0 < z_1$). A decision needs to be made about whether the falsely separated haloes at z_0 should be put back together (i.e. *stitching*), or the falsely grouped halo at z_1 should

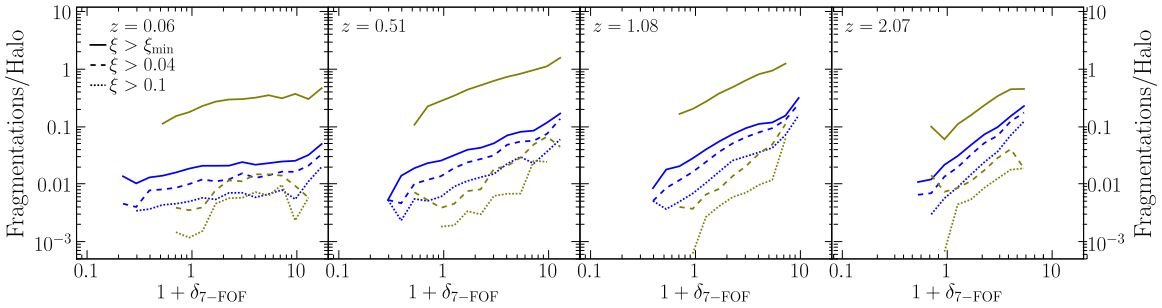

Figure 4.9: Fragmentation statistics (fraction of haloes that fragmented) as a function of environmental density $1+\delta_{7-\mathrm{FOF}}$ at redshifts $z = 0.06, 0.51, 1.08, 2.07$ (left to right). Within each panel, three fragmentation mass ratios ξ are shown: major fragmentations with $\xi > 0.1$ (dotted); those with $\xi > 0.04$ (dashed), and all fragmentations down to the resolution limit (40 particles; solid). The two colours are for different mass bins: 0 to 40% (blue) and 90-99% mass bin (red); see Table 4.1. The red solid curve is significantly higher than the blue solid curve because the fragments of higher mass haloes are better resolved (i.e. ξ_{min} is smaller). Minor fragmentations are seen to dominate, while only $\sim 1\%$ of haloes suffer major fragmentations ($\xi > 0.1$) in typical environments.

be broken up (i.e. *splitting*). (Note: *Snipping* effectively does nothing.)

Fig. 4.8 illustrates how each of the three algorithms – snip, stitch, and split – handles halo fragmentation. For completeness, we also explore a variation of stitch (and split), in which the number of outputs used to make the decision is altered. Instead of stitch-3 (or split-3), which only stitches (or splits) fragmented haloes that remerge within 3 time outputs, we consider stitch-∞ (or split-∞), which stitches (or splits) any fragmented haloes regardless of their future (or past) history. An important distinction between stitch-3 and stitch-∞ (and similarly for split-3 vs split-∞) is that the modifications to the haloes are confined to the three adjacent outputs in stitch-3 and split-3; haloes along the merger tree outside of this time range are unaltered. The modifications made in stitch-∞ and split-∞ however, propagate indefinitely either forward or backward along any tree branch where a fragmentation occurs. No algorithm is perfect, but any error made in stitch-∞ and split-∞ will affect the entire branch of the tree that contains a fragmentation event. By contrast, errors made in stitch-3 and split-3 are confined to the redshift at which the fragmentation occurs. Stitch-∞ and split-∞ are therefore extreme algorithms, which are included here for comparison purposes only.

Before comparing the algorithms, we first show the frequency of fragmentations in the Millennium FOF tree as a function of environment in Fig. 4.9. The fraction of haloes that experience fragmentations is seen to increase with $\delta_{7-\mathrm{FOF}}$, differing by a factor of ~ 3 at low z and by a factor of ~ 5 to 10 at $z \approx 2$. The fragments, however, are dominated by low-mass haloes: only $\sim 1\%$ of the haloes in typical densities have fragments of mass ratio ξ above 0.1, and this fraction is no larger than $\sim 10\%$ even in the densest regions. Most of the fragmentations are therefore minor.

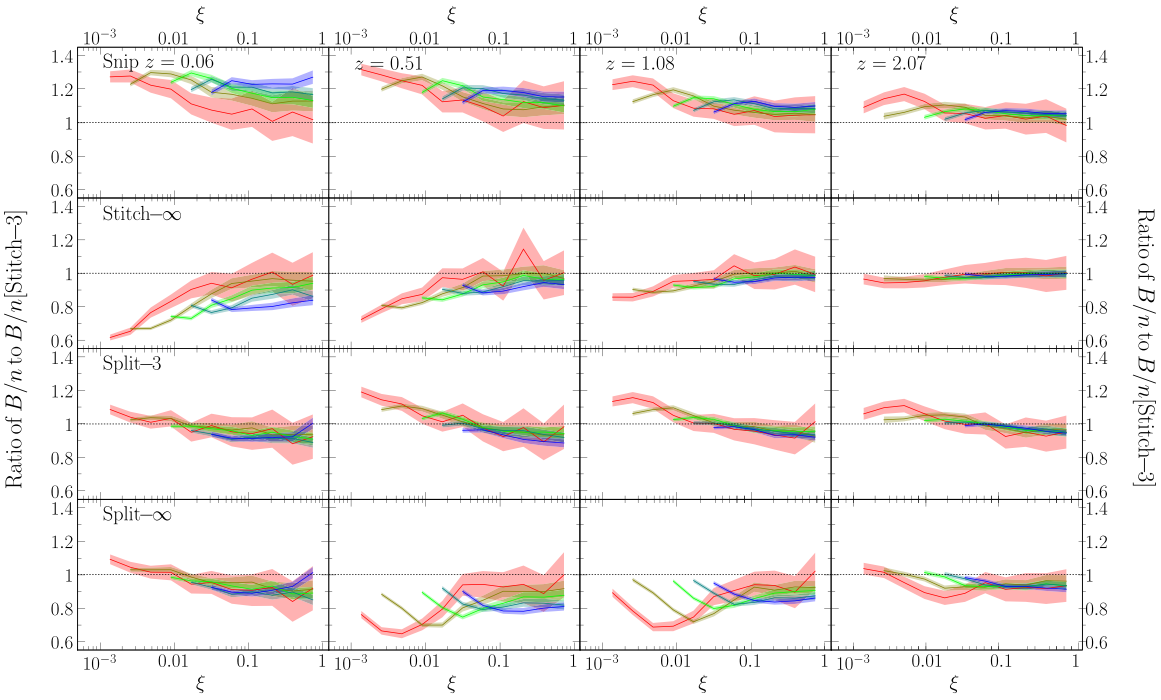

Figure 4.10: Comparison of the global merger rate B/n (including all environment) vs progenitor mass ratio ξ computed from five fragmentation algorithms at $z = 0.06, 0.51, 1.08$, and 2.07 (left to right). Since stitch-3 is the method used in FM08, we plot the ratio of B/n from the other four methods to B/n from stitch-3. Colours correspond to the mass bins in Table 4.1 (blue: lowest mass, red: highest mass). The shaded regions denote Poisson errors.

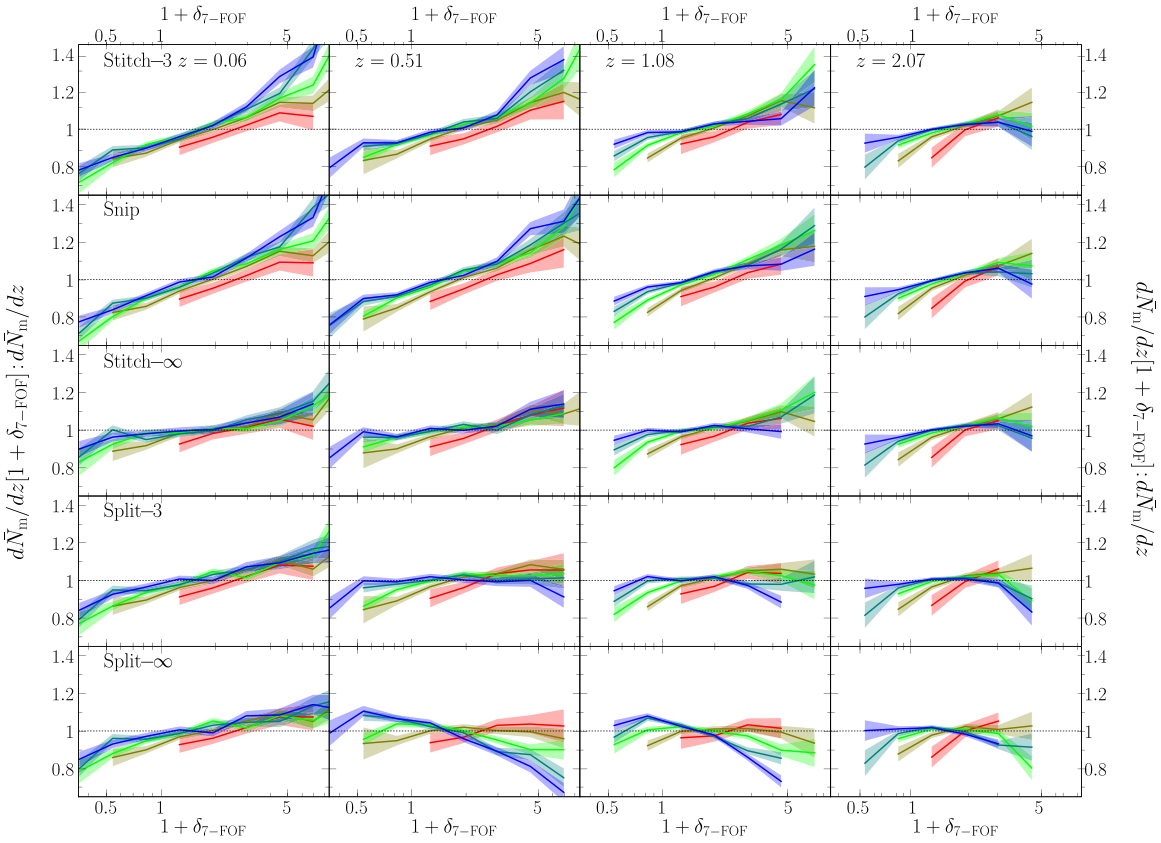

Figure 4.11: Comparison of the environmental dependence of the merger rate computed from five algorithms at $z = 0.06, 0.51, 1.08, 2.07$ (left to right). Similar to Fig. 4.6, we plot the ratio of the mean merger rate $dN_\mathrm{merge}/dz[\delta]$ for haloes in a given $\delta_{7-\mathrm{FOF}}$ bin to the global mean merger rate dN_merge/dz. The mass bins are shown in different colours (blue for the lowest and red for the highest bin in Table 4.1). Shaded regions indicate Poisson errors. This figure shows that stitch-3, snip, and split-3 have similar δ dependence. Split-∞, however, reverses the δ-trend for low mass haloes, which we believe is an artefact of the propagation of fragmentations up the tree (see text).

Fig. 4.10 compares the global mean merger rates B/n (i.e. including all environment) as a function of progenitor mass ratio ξ for the five algorithms (top to bottom) at four redshifts ($z = 0.06, 0.51, 1.08, 2.07$ from left to right). Since stitch-3 is the algorithm used in FM08, we show the ratio of each of the four alternative algorithms to stitch-3. Within each panel, the coloured curves show a variety of descendant mass bins (the bands show Poisson errors). As we have already seen in FM08, snipping (first row in Fig. 4.10) yields a higher merger rate (by $\sim 20\%$ at $\xi < 0.01$) due to the orphaned haloes, resulting in a steeper power-law dependence ξ^{β} ($\beta \sim -2.2$) than stitch-3 ($\beta \sim -2$). Stitch-∞ (second row), on the other hand, zips together all the fragments and reduces the number of minor mergers by as much as $\sim 40\%$ ($\beta \sim -1.8$) in comparison to stitch-3. Split-3 (third row) tends to raise the minor merger rate by up to $\sim 15 - 20\%$. Split-∞ (fourth row) has a feature in the low-ξ merger rate that breaks the power-law behaviour seen in the other trees. This feature is redshift dependent and drives the merger rate lower than in stitch-3.

We now examine how the environmental dependence of the merger rates is affected by the algorithm used for handling fragmentations. To do this, we integrate B/n over ξ and show the total rate, $dN_{\rm merge}/dz$, as a function of $1 + \delta_{7-\rm FOF}$ for the five algorithms (top to bottom) at four redshifts in Fig. 4.11. Similar to Fig. 4.6, the vertical axis shows the ratio of the merger rate in a $\delta_{7-\rm FOF}$ bin to the global rate, $dN_{\rm merge}/dz[\delta] : dN_{\rm merge}/dz$, computed with each method. This figure shows that the stitching (both stitch-3 and stitch-∞) and snipping algorithms produce very similar environmental dependence; though the extreme stitch-∞ yields a mildly weaker δ dependence. In contrast, split-∞ shows a sudden *reversal* in the δ-dependence at $z = 0.51, 1.08, 2.07$ in the three lower mass bins, with haloes in the densest regions experiencing *fewer* mergers. Split-3, on the other hand, shows a positive (albeit weak) correlation of merger rate with δ for all mass bins but the very lowest. We believe the difference between the split and stitch trees is due to an "unzipping" effect that is most pronounced in split-∞, in which splitting a fragmentation event at low z affects the entire branch above this redshift, resulting in the discrepantly low merger rates in high density regions seen in the last row of Fig. 4.11.

In summary, halo fragmentation is a generic feature of all merger trees. It occurs more frequently in dense regions than in voids, thereby prompting the detailed investigation in this section. Our tests of five algorithms show that the majority of tree-processing methods (stitch-3, stitch-∞, snip, and, to some extent, split-3) give very similar environmental dependence for the mean merger rate. In addition, the global merger rate (including all environment) is robust, differing by less than 10% for major mergers and less than 20% even in the very minor merger regime ($\xi < 0.01$) that is more prone to systematic effects. The split-∞ algorithm, on the other hand, appears to suffer from "non-local" effects that have propagated up the merger tree from the fragmentation point. In particular, the merger rate is greatly reduced in high density regions when split-∞ is used.

4.6 Conclusions and Implications

We have used the dark matter haloes and merger trees constructed from the Millennium simulation to quantify the dependence of halo merger rates on halo environment from redshift $z = 0$ to 2. A number of local mass density parameters centred at the haloes, both including and excluding the central halo mass itself, are tested as measures of environment. We have found that $\delta_{7-\mathrm{FOF}}$ defined in equation (4.3) is a robust measure of the surrounding environment outside of a halo's virial radius. It cleanly subtracts out the contributions to the local density from the central halo and thereby breaks the degeneracy between halo mass and environment for high mass haloes (see Figs. 4.1 and 4.2).

We have found strong and positive correlations in both the halo merger rate and the progenitor mass function with environmental densities. Figs. 4.4-4.7 present our main results, where haloes in the densest regions are seen to experience 2 to 2.5 times higher merger rates than haloes in the voids. Such a density dependence can be approximated analytically by multiplying our earlier fitting formula FM08 for the global merger rates (eq. 4.8) by an additional δ-dependent factor given by equation (4.11). This factor is a simple power-law in both the environmental density and halo mass, and it is redshift-independent. The mass dependence is quite weak, indicating that haloes with different masses but similar values of $1 + \delta_{7-\mathrm{FOF}}$ experience similar merger rate amplifications. This is intriguing in light of the fact, discussed in Section 4.3.2, that these haloes actually reside in different environments.

The strong correlations of the halo merger rate and progenitor mass function with environment discussed in this paper have important implications for the analytic Press-Schechter [Press and Schechter, 1974] and excursion set models [Bond et al., 1991, Lacey and Cole, 1993]. In this popular formalism, halo growth is modelled by the random walk trajectories of dark matter density perturbations smoothed at decreasing scales. Haloes are identified at scales at which these trajectories first cross some critical density threshold, and the Markovian nature of the model allows one to compute the distribution of these first crossings. This distribution is then mapped onto the number-weighted conditional mass function $\phi(M, z|M_0, z_0)$ discussed in Section 4.4.2 and plays an important role in the Monte Carlo construction of mock merger tree catalogues (see Zhang et al. 2008a and references therein).

It is generally assumed that the conditional mass function is independent of environment as the excursion set model is Markovian. The Markovian nature of the random walks, however, is not a prediction but rather an assumption resulting from the use of the k-space tophat window function to smooth the density perturbations. There have been recent attempts to weaken this assumption or to introduce environmental dependence into other parts of the model [Zentner, 2007, Sandvik et al., 2007, Desjacques, 2008], but these modifications thus far have not been able to reproduce the basic statistical correlation between halo clustering and formation time found in simulation studies: older haloes are more clustered [Gottlöber et al., 2001, Sheth and Tormen, 2004, Gao et al., 2005, Harker et al., 2006, Wechsler et al., 2006, Jing et al., 2007, Wang et al., 2007, Gao and White, 2007, Maulbetsch et al., 2007].

How do our environmental results for the merger rates tie in with these simulation and EPS studies? We have shown that the amplification of halo merger rates in denser regions

persists at all redshifts (up to at least $z = 2$). If mergers were the dominant channel for halo growth, our results would imply that for haloes of a *fixed* mass today, those in denser regions should have formed *more recently* than those in void regions. Interestingly, this is exactly opposite to the trend reported in many recent studies that have found older (i.e. earlier forming) haloes to be more clustered than younger haloes. As we will discuss in the next paper (Fakhouri & Ma 2008c), these two results are in fact not in conflict once the other important channel for halo mass growth – the "diffuse" accretion of non-halo material (either unresolved or stripped) – is taken into account. We will quantify the environmental dependence of this component and show that, when combined with the merger rate results presented in this paper, we recover the formation redshift dependence reported in prior simulation studies.

4.7 Appendix: Self-Similar Mass and Environment Variables

We have chosen to use the intuitive mass and environment variables M and $\delta_{7-\mathrm{FOF}}$ in the fitting form (eq. 4.11) for the environmental dependence of the merger rate for $0 \leq z \leq 2$. Here we investigate if other choices of mass and density variables may improve the fit. This is motivated by the well known property that the unconditional halo mass function is (almost) redshift-independent when the variable

$$\nu(M, z) = \frac{\delta_c}{\sigma(M)D(z)} \tag{4.12}$$

is used to characterise mass (see, for example, Jenkins et al. 2001); whereas when M is used as the mass variable, the halo mass function evolves significantly with redshift. Here $\delta_c = 1.686$ is the critical overdensity for spherical collapse, $\sigma(M)$ is the variance of the linear density perturbations evaluated at a scale corresponding to the halo mass M, and $D(z)$ is the linear growth function. On the other hand, as we discussed in the paper, the merger rate is more closely related to the *conditional* mass function than the *unconditional* mass function, and the redshift dependence of the former cannot be scaled out simply by using ν. Nonetheless, one can ask whether ν is the more appropriate variable for capturing the mass dependence of the environmental dependence of dN_{merge}/dz.

A similar question can be raised about $\delta_{7-\mathrm{FOF}}$. The overdensity $\delta_{7-\mathrm{FOF}}$ grows as a result of gravitational instability, leading to broader distributions of $\delta_{7-\mathrm{FOF}}$ towards lower redshifts as shown in Fig. 4.2[1]. We can scale out the growth of $\delta_{7-\mathrm{FOF}}$ in the linear regime by replacing $\delta_{7-\mathrm{FOF}}(z)$ by $\delta_{7-\mathrm{FOF}}(z)/D(z)$.

To incorporate the effects of nonlinear growth, we use the approach of Goldberg and Vogeley [2004] for underdensities and Peebles [1984] and Eke et al. [1996] for overdensities.

[1] We have also compared distributions across fixed ν bins and found equivalent changes in the distribution of $\delta_{7-\mathrm{FOF}}$.

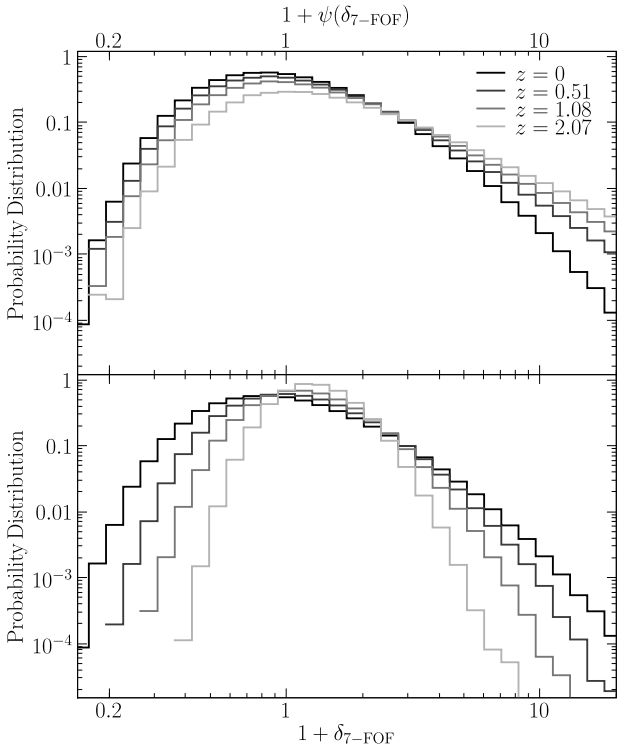

Figure 4.12: Distribution of scaled overdensity variable $1 + \psi(\delta_{7-\mathrm{FOF}})$ (top panel) vs the original $1 + \delta_{7-\mathrm{FOF}}$ (bottom panel) for $z = 0, 0.51, 1.08$, and 2.07. Each histogram is normalised to have unit area. The broadening of the $1 + \delta_{7-\mathrm{FOF}}$ distribution with decreasing z is a natural consequence of gravitational instability. This evolution is largely removed when $\psi(\delta_{7-\mathrm{FOF}})$ is used as the variable, suggesting that $\psi(\delta_{7-\mathrm{FOF}})$ is a more self-similar measure of halo environment.

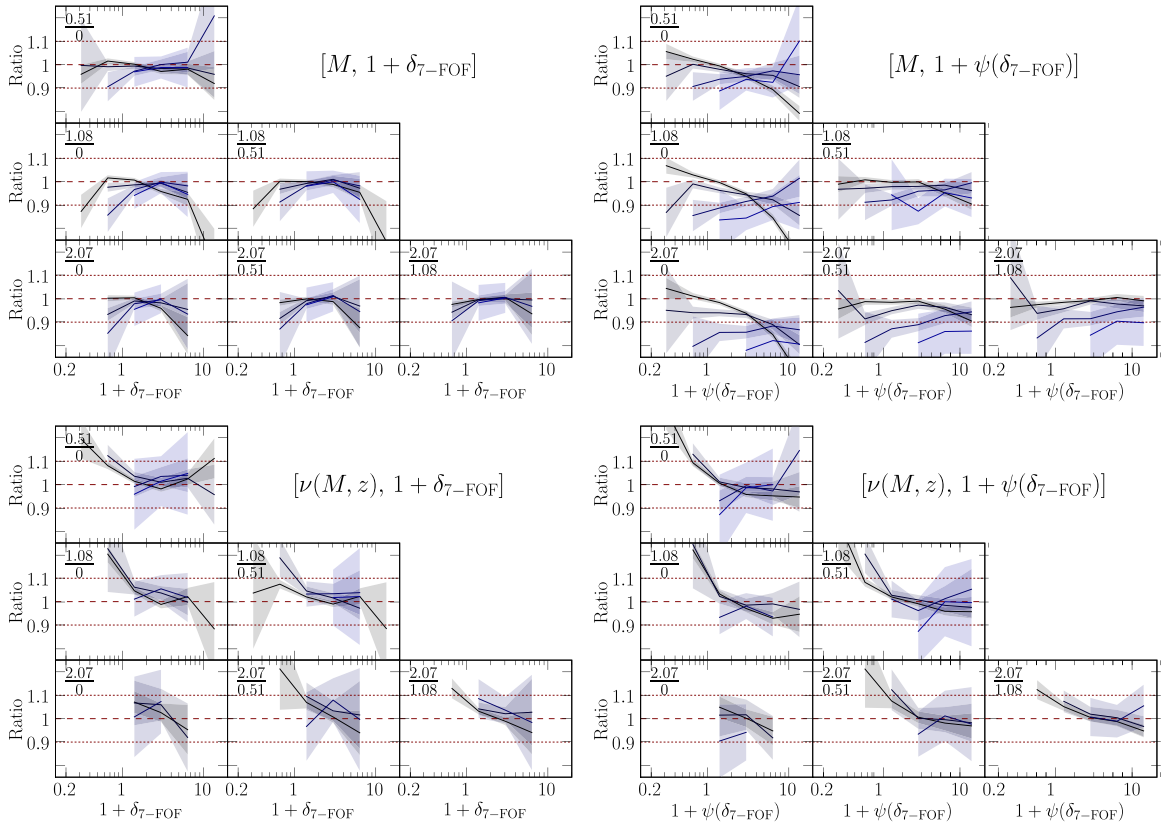

Figure 4.13: Ratios of the environmental dependence of the halo merger rates, $R(M, \delta, z) = \frac{dN_{\rm merge}}{dz}[M, \delta, z] / \frac{dN_{\rm merge}}{dz}[M, z]$, computed at different redshifts: $R(z_2)/R(z_1)$. Two mass and two environment variables are used, in clockwise order from the top-left matrix of subplots: $(M, \delta_{7-\rm FOF})$,$(M, 1 + \psi(\delta_{7-\rm FOF}))$,$(\nu(M, z), 1 + \psi(\delta_{7-\rm FOF}))$, $(\nu(M, z), \delta_{7-\rm FOF})$. Each matrix presents ratios of $R(z_2)/R(z_1)$ plotted as a function of environment. Each line represents a different mass bin with low mass in black and high mass in blue. The shaded regions represent Poisson errors. The redshifts used to compute $R(z_2)/R(z_1)$ are noted in the top left corner of each subplot. The top-left matrix $(M, \delta_{7-\rm FOF})$, the variables used throughout this paper, shows the least systematic dependence on mass and environment.

These authors assume that the under/overdense regions are spherically symmetric and apply Birkhoff's theorem, treating these regions as self-contained universes embedded within an expanding universe. The model cosmological parameters for these embedded cosmologies are computed from the density of the region under consideration, and the resulting Friedmann equation is solved to relate the densities at some redshift, $\delta_{7-\mathrm{FOF}}(z)$, to densities, $\psi(\delta_{7-\mathrm{FOF}}(z))$, at $z = 0$. When $\delta_{7-\mathrm{FOF}} \ll 1$, this procedure is in agreement with the linear relation $\psi(\delta_{7-\mathrm{FOF}}) \sim \delta_{7-\mathrm{FOF}}/D(z)$

A further complication for $\delta_{7-\mathrm{FOF}}$ is that the mass of the central object has been removed from δ_7. We have tested swapping the order of operation by first nonlinearly propagating $\delta_7 \rightarrow \psi(\delta_7)$ *then* subtracting the mass of the central object. We find that the resulting distributions of $\psi(\delta_7)_{-\mathrm{FOF}}$ are only slightly modified from the distributions of $\psi(\delta_{7-\mathrm{FOF}})$.

Fig. 4.12 compares the distribution of the scaled $1 + \psi(\delta_{7-\mathrm{FOF}})$ (top panel) with that of the original $1 + \delta_{7-\mathrm{FOF}}$ (bottom panel) for all haloes with $M > 1.2 \times 10^{12} M_\odot$ at $z = 0, 0.51, 1.08$ and 2.07 (black to light grey lines). The upper panel shows far less broadening with decreasing z than in the lower panel, indicating that $\psi(\delta_{7-\mathrm{FOF}})$ *does* remove much of the redshift evolution of $\delta_{7-\mathrm{FOF}}$. The mapping is clearly imperfect: The distributions at $z = 1.08$ and $z = 2.07$ have long positive density tails that are not present at $z = 0$. This is not surprising since the simple spherical approximation used for evolving the density cannot account for all non-linear effects such as the mergers of overdense and underdense regions.

To test if $\psi(\delta_{7-\mathrm{FOF}})$ and ν are more appropriate variables to use in the fitting formula than $\delta_{7-\mathrm{FOF}}$ and M, we show in Fig. 4.13 the ratio of $dN_{\mathrm{merge}}/dz[\delta]$ to the global mean dN_{merge}/dz for a number of log-spaced mass bins at $z = 0, 0.51, 1.08$, and 2.07. For brevity, let us refer to this ratio as $\Xi(M, \delta, z)$. We can compute Ξ using either M or $\nu(M, z)$ as the mass variable, and either $\delta_{7-\mathrm{FOF}}$ or $1 + \psi(\delta_{7-\mathrm{FOF}})$ as the environmental variable. The upper left set of plots in Fig. 4.13 uses M and $1 + \delta_{7-\mathrm{FOF}}$, which are the variables used throughout this paper; the upper right set uses M and $1 + \psi(\delta_{7-\mathrm{FOF}})$; the lower left set uses $\nu(M, z)$ and $1 + \delta_{7-\mathrm{FOF}}$; and the lower right set uses $\nu(M, z)$ and $1 + \psi(\delta_{7-\mathrm{FOF}})$.

Within each set of figures we plot a matrix of ratios of Ξ computed at different redshifts. The redshifts used to compute the ratios are noted in the upper left corner of each subplot. For example, the upper subplot is the ratio $\Xi(M, \delta, 0.51)/\Xi(M, \delta, 0)$ and is labelled "0.51/0". Each subplot contains five mass bins. Only points containing more than 40 haloes are plotted to minimise noise (this results in some mass bins being dropped). The variables that successfully capture the redshift evolution in the merger rate will show very little variation in $\Xi(M, \delta, z)$ with redshift and give ratios $\Xi(M, \delta, z_1)/\Xi(M, \delta, z_2)$ close to unity. Interestingly, the $[M, \delta_{7-\mathrm{FOF}}]$ pair used throughout this paper does the best job. The upper left matrix in Fig. 4.13 shows ratios of Ξ that tend to cluster around 1 and show few systematic trends with mass and environment.

Using $\nu(M, z)$ instead of M (lower panels) introduces a strong δ dependence in the ratios of Ξ: the δ-slope of dN_{merge}/dz flattens with increasing redshift. Similarly, using $1+\psi(\delta_{7-\mathrm{FOF}})$ instead of $\delta_{7-\mathrm{FOF}}$ (right panels) introduces a strong mass dependence when M is used as mass variable and does not improve on the δ dependence introduced by $\nu(M, z)$.

Thus, M and $1 + \delta_{7-\mathrm{FOF}}$ appear to be the optimal variables for capturing the environ-

mental dependence of dark matter halo merger rates for $0 \leq z \leq 2$.

Chapter 5

Dark Matter Halo Growth II: Diffuse Accretion and its Environmental Dependence

Fakhouri, Onsi; Ma, Chung-Pei — Februrary 2010
Monthly Notices of the Royal Astronomical Society, Volume 401, Issue 4, pp. 2245-2256

Dark matter haloes in ΛCDM simulations grow by mergers with other haloes as well as accretion of "diffuse" non-halo material. We quantify the mass growth rates via these two processes, $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$, and their respective dependence on the local halo environment using the $\sim 500,000$ haloes of mass $\sim 10^{12}$ to $10^{15} M_\odot$ in the Millennium simulation. Adopting a local mass density parameter as a measure of halo environment, we find the two rates to show strong but opposite environmental dependence, with mergers playing an increasingly important role for halo growths in overdense regions while diffuse accretion dominating the growth in the voids. For galaxy-scale haloes, these two opposite correlations largely cancel out, but a weak environmental dependence remains that results in a slightly lower mean total growth rate, and hence an earlier mean formation redshift, for haloes in denser environments. The mean formation redshift of cluster-mass haloes, on the other hand, shows no correlation with halo environment. The origin of the positive correlation of $\dot{M}_{\rm mer}$ with local density can be traced to the sourrounding mass reservoir outside the virial radii of the haloes, where more than 80% of the mass is in the form of resolved haloes for haloes residing in densest regions, while this fraction drops to $\sim 20\%$ in the voids. The negative correlation of $\dot{M}_{\rm dif}$ with local density, however, is not explained by the available diffuse mass in the reservoir outside of haloes, which is in fact larger in denser regions. The non-halo component may therefore be partially comprised of truly diffuse dark matter particles that are dynamically hotter due to tidal stripping and are accreted at a suppressed rate in denser regions. We also discuss the implications of these results for how to modify the analytic Extended Press-Schechter model of halo growths, which in its original form does not predict environmental dependence.

5.1 Introduction

The bottom-up growth of structure is a hallmark of hierarchical cosmological models such as the Λ cold dark matter (ΛCDM) model. In these universes, dark matter haloes of lower mass are expected to form earlier, on average, than massive haloes. In observations, theoretical studies, and semi-analytical modelling of galaxy formation (see Baugh 2006 for a review), the mass of a halo is the key variable upon which many properties of galaxies and their host haloes depend, e.g., formation redshift, galaxy occupation number, colour, morphology, star formation rate, and stellar feedback processes.

Recent work has shown that in addition to the mass, various properties of halo formation and evolution also depend on the environment within which the haloes reside. For instance, at a fixed halo mass, older haloes have been found to be more clustered than younger haloes, and the correlation between clustering strength and formation redshift is stronger for lower mass haloes [Gottlöber et al., 2002, Sheth and Tormen, 2004, Gao et al., 2005, Harker et al., 2006, Gao and White, 2007, Jing et al., 2007, Wechsler et al., 2006, Wang et al., 2007, Hahn et al., 2008, McBride et al., 2009]. In many of these studies, halo environment is characterised via the halo bias parameter, which is determined from the relative clustering strengths of haloes to the underlying dark matter distribution.

In Part I of this series (Fakhouri and Ma 2009; henceforth FM09), we have instead chosen to use local overdensities as a more direct measure of halo environment. The focus of the study was to quantify the environmental dependence of the merger rate of haloes. We compared a number of local overdensity variables (both including and excluding the halo mass itself), some of which had been used in earlier studies [Lemson and Kauffmann, 1999, Harker et al., 2006, Wang et al., 2007, Maulbetsch et al., 2007, Hahn et al., 2008]. The key finding in FM09 was that halo-halo mergers occur more frequently in denser regions than in voids, and that this environmental dependence is similar regardless of the merger mass ratio (e.g., minor vs major) or the descendant halo mass (galaxy- vs cluster-sized): we found mergers to occur about 2.5 times more frequently in the densest regions than in the emptiest regions. We provided an analytical formula as a function of local density to approximate this environmental trend. This expression can be used with the fit for the global mean rate of Fakhouri and Ma [2008] (henceforth FM08) to predict halo merger rates as a function of descendant mass, progenitor mass ratio, redshift, and environment over a wide range of parameter space.

In this paper, we build on our earlier merger rate studies by investigating the mass growth of haloes, and its environmental dependence, via two sources: growth from mergers with other haloes (a quantity closely related to the results of FM09), and growth from accretion of non-halo material, which we will refer to as "diffuse" accretion. Due to the finite resolution of the simulations, we expect a portion of this diffuse component to be comprised of unresolved haloes, which, in a higher-resolution simulation, should largely follow the merger physics and scaling laws of their higher mass counterparts in our earlier studies. The diffuse non-halo component, however, can in principal also contain truly diffuse dark matter particles that either were tidally stripped from existing haloes or were never gravitationally bound to any

haloes.

As reported below, we find that diffuse accretion plays an important role in contributing to halo mass growth in the Millennium simulation. Moreover, we will show that this growth component correlates with the local halo environment in an opposite way from the component due to mergers, with diffuse accretion playing an increasingly important role in halo growth in the voids, and mergers playing a more important role in the densest regions. This difference suggests that the diffuse component is not simply an extension of the resolved haloes down to lower masses, but rather that there is an intrinsic difference between the two components that results in the opposite environmental trends. An implication of this result is that Milky-Way size galaxies in voids and those that reside near massive clusters may have statistically distinguishable formation history, where the galaxies in voids acquire their baryons more quiescently via diffuse accretion, while those in dense regions assemble their baryons mainly via mergers. Such environmental effect can show up in galaxy properties such as the star formation rates, colors, and morphologies.

The environmental dependence of halo growths reported in this paper also has far-reaching implications for the much-used analytic theories for halo growth such as the Extended Press-Schechter (EPS) and excursion set models [Press and Schechter, 1974, Bond et al., 1991, Lacey and Cole, 1993]. These models assume that all dark matter particles reside in haloes, and halo growths depend only on mass and not environment. As we will elaborate on below, both assumptions are too simplistic and must be modified to account for the results from numerical simulations.

This paper is organised as follows. §5.2.1 summarises the various definitions of halo mass and environment used in our analysis, as well as the means by which we extract halo merger trees from the public data in the Millennium simulation. In Sec. 3 we discuss how the two mass growth rates due to mergers and diffuse accretion, $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$, are defined and computed. The distributions of the rates for the $\sim 500,000$ haloes at $z = 0$ and their redshift evolution are presented. The environmental dependence of halo growths is analysed in Sec. 4. The correlations of four halo properties with the local density parameter are investigated: the mass growth rates $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$ (§ 4.1), the fraction of a halo's final mass gained via mergers vs. diffuse accretion (§ 4.2), the formation redshift z_f (§ 4.3), and the composition of the surrounding mass reservoir outside of the virial radii of the haloes (§ 4.4). §5.4.5 discusses a test that we have performed to verify that the majority of haloes reside in a similar environmental region (e.g. overdense or underdense) throughout their lifetimes. In Sec. 5 we investigate further the nature of the "diffuse" component by varying the mass threshold used to define $\dot{M}_{\rm mer}$ vs $\dot{M}_{\rm dif}$. We then discuss the implications of our results for the analytic EPS model of halo growth, which is entirely independent of environment in its basic form and therefore must be modified to account for the various environmental trends reported in § 4.

Redshift	Mass Percentile (Masses in units of $10^{12} M_\odot$) 0-40%	40-70%	70-90%	90-99%	99-100%
$z = 0\,(\Delta z = 0.06, \Delta t = 0.83\,\mathrm{Gyr})$	$1.2 - 2.1$	$2.1 - 4.5$	$4.5 - 14$	$14 - 110$	> 110
$z = 0.51\,(\Delta z = 0.06, \Delta t = 0.38\,\mathrm{Gyr})$	$1.2 - 2.0$	$2.0 - 4.1$	$4.1 - 12$	$12 - 74$	> 74
$z = 1.08\,(\Delta z = 0.09, \Delta t = 0.34\,\mathrm{Gyr})$	$1.2 - 1.9$	$1.9 - 3.7$	$3.7 - 9.5$	$9.5 - 48$	> 48
$z = 2.07\,(\Delta z = 0.17, \Delta t = 0.25\,\mathrm{Gyr})$	$1.2 - 1.8$	$1.8 - 3.0$	$3.0 - 6.5$	$6.5 - 24$	> 24

Table 5.1: Mass bins used in this paper for redshifts $0, 0.51, 1.08$, and 2.07. The bins are computed assuming fixed mass-percentile bins (header row) and the mass boundaries are computed using the prescribed percentile. The high mass bins extend out to $5.2 \times 10^{15}, 3 \times 10^{15}, 1.3 \times 10^{15}$, and $4.4 \times 10^{14} M_\odot$ for $z = 0, 0.51, 1.08, 2.07$ respectively, though these exceptionally high mass objects are outliers and not particularly representative of the high mass range.

5.2 Merger Trees and Halo Environment in the Millennium Simulation

5.2.1 Merger Trees of the Millennium Haloes

The Millennium simulation [Springel et al., 2005] assumes a ΛCDM model with $\Omega_m = 0.25$, $\Omega_b = 0.045$, $\Omega_\Lambda = 0.73$, $h = 0.73$ and a spectral index of $n = 1$ for the primordial density perturbations with normalisation $\sigma_8 = 0.9$ [Springel et al., 2005]. The dark matter N-body simulation followed the trajectories of 2160^3 particles of mass $1.2 \times 10^9 M_\odot$ in a (685 Mpc)^3 box from redshift $z = 127$ to $z = 0$. A friends-of-friends group finder with a linking length of $b = 0.2$ is used to identify $\sim 2 \times 10^7$ dark matter haloes in the simulation down to a mass resolution of 40 particles ($\sim 4.7 \times 10^{10} M_\odot$). Each FOF halo thus identified is further broken into constituent subhaloes, each with at least 20 particles, by the SUBFIND algorithm that identifies gravitationally bound substructures within the host FOF halo [Springel et al., 2001a].

Even though the Millennium public database provides a catalogue of FOF haloes at each output, it does not give the merger trees for these haloes. Instead, it provides merger trees for the *subhaloes*, which are constructed by connecting the subhaloes across the 64 available redshift outputs. During this construction, a decision must be made about the ancestral relations of the subhaloes since the particles in a given subhalo may go into more than a *single* subhalo in the subsequent output. In this case, a subhalo is chosen to be the descendent of a progenitor subhalo at an earlier output if it hosts the largest number of bound particles in the progenitor subhalo. The resulting merger tree of subhaloes can then be processed further to construct the merger tree of the FOF haloes. This construction is non-trivial due to the fragmentation of FOF haloes; this is discussed at length in FM08 and FM09.

In FM08 and FM09, we proposed a variety of post-processing algorithms to handle FOF fragmentation. Three methods were compared in FM08: *snip, stitch-3*, and *stitch-*∞. The snipping method severs the ancestral relationship between halo fragments and their progen-

itors. This is the method commonly used in the literature and suffers from inflated merger rates due to the aberrant remerger of snipped fragments. The two stitching algorithms prevent halo fragmentation by "stitching" halo fragments together such that each FOF halo in the simulation has exactly one descendant. Stitch-∞ performs this procedure whenever fragmentation occurs, whereas stitch-3 only reincorporates halo fragments that are destined to remerge within 3 outputs of the fragmentation event. Both stitching algorithms lower the minor merger rate as they prevent spurious remergers of halo fragments. Naturally, this reduction is strongest for stitch-∞.

In FM09 we showed that the rate of halo fragmentation was a strong function of halo environment, with haloes residing in overdense regions undergoing fragmentations over three times as often as haloes in underdense regions. We showed that the choice of post-processing algorithm was important in determining the environmental dependence of halo merger rates in dense environments and introduced two new algorithms: *split-3* and *split-*∞. These algorithms remove halo fragmentations by splitting progenitor haloes that fragment into multiple haloes to ensure that each halo has exactly one descendant. As with stitch-3, split-3 only splits progenitors that are split within the past 3 outputs (looking backwards towards high z) of the fragmentation event. Split-∞ *always* splits fragmenting haloes. See also Genel et al. [2009] who introduced a similar algorithm.

We showed in FM09 that split-∞ suffered from an "unzipping" effect as halo fragmentations propagated up the tree, breaking the self-similarity in the merger rate, B/n, observed in FM08. Split-3, however, yielded very similar merger rates (within $\sim 10\%$) as stitch-3 (see Fig. 10 of FM09).

We have tested the various quantities presented in this paper using all five post-processing algorithms and have indeed found the differences not to be significant enough to warrant presentating plots for all methods. We will, however, briefly discuss any relevant differences in the text. The results presented throughout this paper will use the stitch-3 algorithm used in FM08 and FM09.

5.2.2 Halo Environment

To measure each halo's local environment, we will primarily use the overdensity variable $\delta_{7-\mathrm{FOF}}$ presented in FM09. This quantity computes the overdensity within a sphere of radius 7 Mpc, centred on each halo in the simulation, by adding up the mass contributions made by the particles within the sphere and subtracting the FOF mass of the central halo. The database provides the dark matter density on a 256^3 grid but not the full particle distribution itself. We have mapped this grid, which is given along a Peano-Hilbert space filling curve, to spatial coordinates and interpolated it to obtain δ at the position of each halo's centre.

We will also compare some of the results in this paper to the alternative environmental variable δ_7 discussed in FM09, where δ_7 is the overdensity surrounding a halo within a sphere of radius 7 Mpc (including the halo mass itself). Fig. 1 of FM09 showed that δ_7 was strongly correlated with mass for haloes above $10^{14} M_\odot$ when the central object's contribution to δ began to dominate at these mass scales. This correlation was absent when the variable

$\delta_{7-\mathrm{FOF}}$ was used. We will therefore use $\delta_{7-\mathrm{FOF}}$ as the environmental measure for most of the analyses below.

5.2.3 Halo Mass Definitions

The Millennium database provides a number of mass measurements for each identified FOF halo. Two simple mass measures are M_{FOF}, the total mass of the particles connected to an FOF by the group finder, and M_{SH}, the sum of the masses of the subhaloes that constitute the FOF halo. The latter mass includes only particles that are bound by the SUBFIND algorithm; thus $M_{SH} \leq M_{FOF}$ by definition.

In FM08 and FM09 we used M_{FOF}, the total mass of the particles that belong to an FOF halo by the group finder. Genel et al. [2009] have shown, however, that for the subset of haloes that are about to undergo minor mergers, the FOF mass of the lower-mass halo can be significantly higher (up to a factor of 1.4) than its subhalo mass (M_{SH}) due to deficiencies in the FOF halo-finding algorithm. Genel et al. [2009] showed that this effect is more severe for more minor mergers. In light of this study, we will present the majority of our results for both M_{FOF} and M_{SH} and comment on the difference.

Since the most massive haloes at $z = 0$ are more massive than the most massive haloes at higher redshifts, we will present some results below using mass bins defined by a fixed percentage of haloes rather than by a fixed absolute mass at different redshifts. Table 5.1 lists the corresponding halo masses at $z = 0$, 0.51, 1.08, and 2.07 for each of the five percentile bins. Note that even in the highest 1% mass bin, there are $\sim$ 5000 cluster-mass haloes available for our study.

5.3 Diffuse Accretion

5.3.1 Quantifying Halo Growth via Mergers vs Diffuse Accretion

The merger trees of FOF haloes constructed in Fakhouri and Ma [2008] provide, for a descendant halo of mass M_0 identified at redshift z_0, a list of the masses of its N_p progenitors at any chosen z_1. We label the progenitors M_i, with $i = 1, ..., N_p$, and $M_1 \geq M_2 \geq M_3$... etc. The number of progenitors, N_p, for a given descendant can range from 1 (i.e. no mergers in that timestep), 2 (binary mergers), to high values for massive haloes. The variable ξ is used to denote the mass ratio of a merging progenitor and the largest progenitor M_1: $\xi = M_{i>1}/M_1$.

Typically we find that the mass M_0 of a descendant halo is not equal to the mass in the progenitors $\sum_{i\geq 1} M_i$, leading us to define

$$\Delta M = M_0 - \sum_{i=1}^{N_P} M_i \,, \tag{5.1}$$

where ΔM quantifies the portion of the mass change that cannot be attributed to mergers of resolved haloes M_i, and ΔM can be negative. A non-zero ΔM can be due to the merging of unresolved progenitors below our minimum halo mass, accretion of dark matter that is not locked up in haloes, or mass loss processes such as tidal stripping. We collectively refer to ΔM as *diffuse accretion* throughout this paper.

To quantify the relative contributions to halo growth due to mergers vs diffuse accretion we define two mass growth rates for each halo of mass M_0 at some redshift z_0:

$$\begin{aligned} \dot{M}_{\rm mer} &\equiv \frac{\sum_{i\geq 2} M_i}{\Delta t}\,, \\ \dot{M}_{\rm dif} &\equiv \frac{\Delta M}{\Delta t}\,, \\ \dot{M}_{\rm tot} &= \dot{M}_{\rm mer} + \dot{M}_{\rm dif} = \frac{M_0 - M_1}{\Delta t}. \end{aligned} \tag{5.2}$$

where $\dot{M}_{\rm mer}$ is due to mergers and $\dot{M}_{\rm dif}$ is due to the accretion of diffuse material. Since $M_0 - M_1$ is the mass difference along the thickest branch of the merger tree, the sum, $\dot{M}_{\rm tot} = \dot{M}_{\rm mer} + \dot{M}_{\rm dif}$, is simply the net mass growth rate of the halo. A detailed study of $\dot{M}_{\rm tot}$ of the haloes in the Millennium simulation is given in McBride et al. [2009]; here our focus is on the separate contributions and their environmental dependence. The timestep Δt in equation (5.2) is the time interval corresponding to $\Delta z = z_1 - z_0$, the redshift spacing between two Millennium outputs used to calculate the rate. We use $\Delta z = 0.06$ for $z_0 = 0$ ($\Delta t = 0.83$ Myrs) and 0.51, $\Delta z = 0.09$ for $z_0 = 1.08$, and $\Delta z = 0.17$ for $z_0 = 2.07$, following the extensive convergence tests reported in Fakhouri and Ma [2008].

We also find it convenient to present the results for the mass growth rates in dimensionless units. When these occasions arise, we will use the fractional mass gain per unit redshift, $dM_{\rm mer}/dz/M_0$ and $dM_{\rm dif}/dz/M_0$, which are simply equal to the rates in equation (5.2) multiplied by $\Delta t/\Delta z/M_0$.

It is important to keep in mind that the distinction between mass growth from resolvable haloes and diffuse accretion depends on the mass threshold used to define these two components. In this study, we have made the conservative choice of including only descendant haloes with more than 1000 particles (with a corresponding mass of $1.2 \times 10^{12} M_\odot$) and progenitor haloes with more than 40 particles ($4.8 \times 10^{10} M_\odot$). A mass ratio of $\xi_{\rm min} = 0.04$ is therefore the resolution limit on progenitors for the most poorly resolved descendant haloes (1000 particles) in our sample. For these haloes, the mass contributed by progenitors above 40 particles is tagged as due to mergers, while any remaining mass contribution is tagged as due to diffuse accretion.

For more massive descendants, there are two natural choices for the mass threshold: one can include either all progenitors down to the 40 particle resolution limit as mergers, or all progenitors down to a fixed mass *ratio* of $\xi_{\rm min} = 0.04$ as mergers (any progenitors with $\xi < \xi_{\rm min}$ are then counted as "diffuse" material). The former has the advantage that 40 particles correspond to a smaller $\xi_{\rm min}$ for more massive descendants; we will therefore

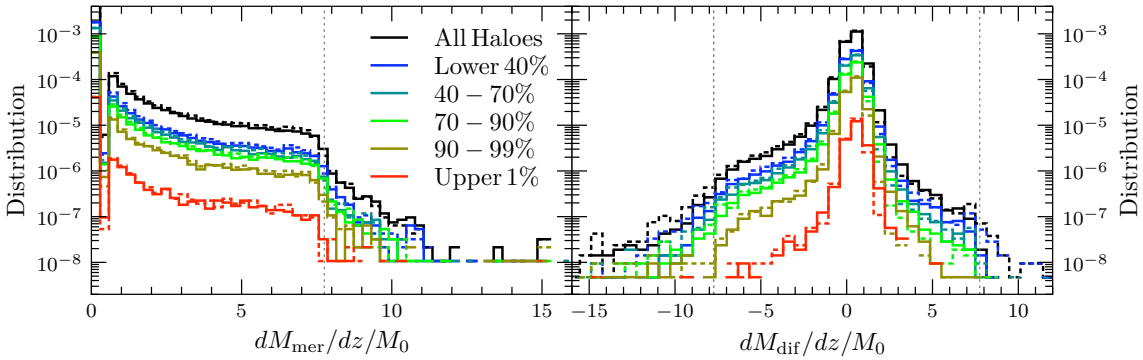

Figure 5.1: Distributions of the mass growth rates (per unit redshift) via mergers (left panel) and via diffuse accretion (right panel) for the $\sim 500,000$ haloes at $z = 0$ in the Millennium simulation. In each panel, the colours denote haloes of different mass percentiles (see Table 5.1), and two definitions of halo mass are shown (dashed for M_{FOF}; solid for M_{SH}).

have better statistics and dynamic range for more massive haloes. The latter choice has the advantage that mergers and diffuse accretion can be compared objectively across different descendant mass bins since the two components are defined with respect to the same $\xi_{\rm min}$ for all haloes and are, therefore, effectively equally well resolved. In Sec. 5.5.1 we will show that the basic behaviour of the environmental dependence presented in this paper does not depend on which definition is used. We generally favour the latter definition, and will use $\xi_{\rm min} = 0.04$ to separate mergers vs. diffuse accretion unless otherwise stated in this paper.

5.3.2 Distributions of $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$

Fig. 5.1 presents the distributions of the mass growth rates due to mergers (left panel) and diffuse accretion (right panel) at $z = 0$ for five halo mass bins. The two halo mass definitions discussed in Sec. 2.3, M_{FOF} (dashed) and M_{SH} (solid), are compared. In constructing these plots, we have used a $\xi > \xi_{\rm min}$ cut: mergers with mass ratios less than $\xi_{\rm min} = 0.04$ are counted as diffuse accretion.

We have also imposed a cut on the very negative and positive tails of the distributions. Although the majority of the haloes in the Millennium simulation have reasonable $dM_{\rm mer}/dz/M_0$ and $dM_{\rm dif}/dz/M_0$, about 0.3% of them more than doubled their mass via diffuse accretion between $z = 0$ and $z = 0.06$, and about 0.4% lost over half of their current mass via diffuse accretion. Our detailed examination of the merger histories of haloes in these tails of the distributions shows that these very large values of $|dM_{\rm dif}/dz/M_0|$ are unphysical and are due primarily to complex and spurious halo fragmentation. The algorithm used to construct the FOF merger trees, stitch-3, only stitches halo fragmentations that are

destined to remerge within 3 snapshots. Some fragmentations do not satisfy this condition.

This small set of haloes – a set too small to affect the merger statistics studied in our earlier work – can significantly affect the mean values of $dM_{\rm mer}/dz/M_0$ and $dM_{\rm dif}/dz/M_0$. This is especially true when these mean values are computed as a function of halo environment, since fragmentations occur more frequently in the messier dense environments. Moreover, haloes with artificially high growth rates have correspondingly artificially low formation redshifts (z_f). These events pollute the distribution of z_f and tend to underpredict the formation redshift, especially in overdense regions. We therefore implement a cut on $dM_{\rm mer}/dz/M_0$ and $dM_{\rm dif}/dz/M_0$ to remove this aberrant population. This is done by demanding that a halo never gains more than a fraction $f_+ = 1/2$, nor lose more than a fraction $f_- = 1/2$, of its *final* mass M_0 in either $dM_{\rm mer}/dz/M_0$ or $dM_{\rm dif}/dz/M_0$ between any two adjacent simulation outputs along the halo's main branch (dotted vertical lines in Fig. 5.1). This is a fairly stringent cut as it is applied to a halo's entire mass history. The resulting cut removes $\sim 3\%$ of the haloes in our sample. Varying the values of f_+ and f_- in the range of $f \in [0.2, 1.]$ and allowing $f_+ \neq f_-$ made insignificant changes and did not affect the overall trends reported in this work. A handful haloes with rates beyond these cuts remain in Fig. 5.1 because the cut is applied to the halo growth across *adjacent* outputs whereas the rates shown in Fig. 5.1 are computed across the three outputs between $z = 0$ and $z = 0.06$ (which straddle the $z = 0.02, 0.04$ outputs).

In the left panel of Fig. 5.1, the strong peak at $dM_{\rm mer}/dz/M_0 = 0$ is due to the fact that between $z = 0$ and 0.06 roughly 65% of haloes have only one progenitor and, therefore, experience no mergers. The remaining 35% of haloes have a distribution that falls off smoothly out to $dM_{\rm mer}/dz/M_0 \sim 8.3$ (dotted vertical line). Values exceeding 8.3 correspond to haloes that have gained more than half their mass between $z = 0.06$ and $z = 0$ (i.e., $\sum M_i/M_0/\Delta z > 1/2/0.06 = 8.3$); only $\sim 0.2\%$ of haloes are in this category. In the right panel of Fig. 5.1, the peak occurs at small positive values of $dM_{\rm dif}/dz/M_0$, though a significant set of haloes experience diffuse stripping.

To complement the differential distributions presented in Fig. 5.1, we show in Fig. 5.2 the *cumulative* distributions for the number density of haloes above a given mass growth rates in units of $M_\odot$ per year. The three panels are for three halo mass bins. The contributions from mergers (thick curves), non-halo material (thin curve), and the sum (black curves) are plotted separately. The curves for the total rate are identical to the solid curves ($z = 0$) in the bottom panels of Fig. 5 of McBride et al. [2009]. This figure shows that the haloes with high accretion rates in the simulation are undergoing mergers with other haloes, while the haloes with low accretion rates are mostly accreting non-halo mass.

5.3.3 Redshift Evolution of the Mean Growth Rates

Having examined the distributions of the growth rates at $z = 0$ in the last section, we now study the redshift dependence of the *mean* growth rates. The left panel of Fig. 5.3 shows the mean $\dot{M}_{\rm mer}$ (solid), $\dot{M}_{\rm dif}$ (dashed), and total growth rate $\dot{M}_{\rm mer} + \dot{M}_{\rm dif}$ (dotted) as a function of redshift for three mass bins: $10^{12} M_\odot$ (blue), $10^{13} M_\odot$ (green), and $10^{14} M_\odot$ (red),

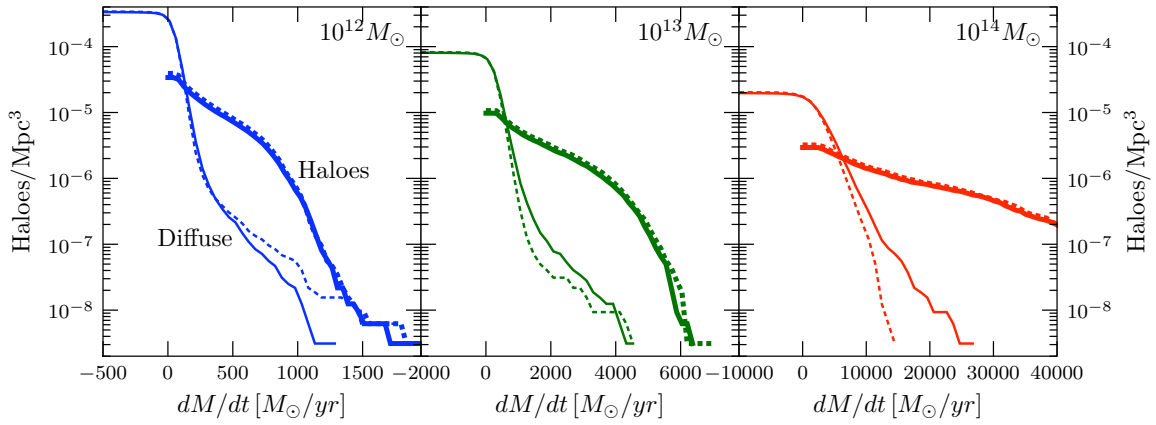

Figure 5.2: Cumulative distributions of the $z = 0$ mass growth rates via mergers with haloes (thick curves) and diffuse accretion (thin curves) for haloes in three mass bins (left to right panels). The axes are in physical units: halo number density (y-axis) and accretion rate in $M_\odot$ per year (x-axis). The total rate is plotted in black. The solid and dotted curves are for the two halo mass definitions M_{SH} and M_{FOF}. The high accretion rate events are dominated by mergers with other haloes.

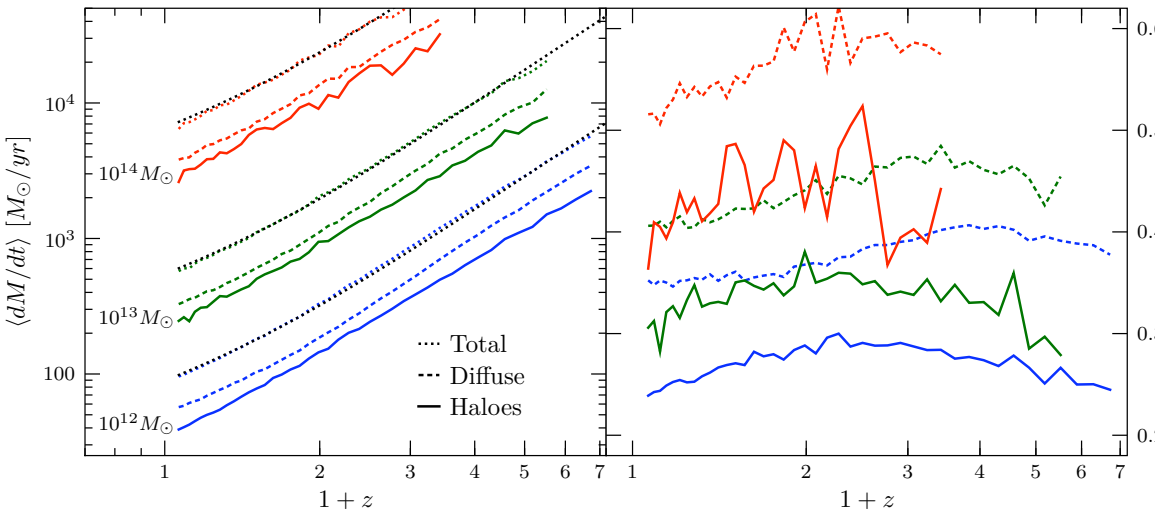

Figure 5.3: Redshift evolution of the mean halo mass growth rates due to mergers (solid) and diffuse accretion (dashed) for three mass bins: $10^{12}M_\odot$ (blue), $10^{13}M_\odot$ (green), and $10^{14}M_\odot$ (red). The rates in units of $M_\odot$ per year (left panel) are seen to rise with increasing z, while the rates in units of per redshift (right panel) is largely independent of z. We also plot the total merger rate (haloes + diffuses) with dotted lines and overlay the fit presented in equation 5.3 in black. Note that the fit is quite good and the total rate and fit are difficult to distinguish.

where the threshold $\xi_{\rm min} = 0.04$ is used to define $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$. The right panel shows the same information except the rates are expressed in the dimensionless units of $dM/dz/M_0$. The growth rates per year (left) are seen to increase rapidly with increasing z, while the rate per unit z (right) has a very weak dependence on z out to $z \sim 6$. As a function of halo mass M_0, the right panel shows that the haloes of higher mass experience somewhat higher *fractional* mass growth rates than haloes of lower mass. Similar dependence on time and mass was seen for the halo merger rates in Fakhouri and Ma [2008].

The left panel of Fig. 5.3 can be compared directly with the total mass growth rates shown in Fig. 4 of McBride et al. [2009], which is well approximated by equation (7) proposed there:

$$\left\langle \dot{M}_{\rm tot} \right\rangle = 42\, M_\odot {\rm yr}^{-1} \left(\frac{M}{10^{12} M_\odot}\right)^{1.127} \times (1+1.17z)\sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda}\,. \quad (5.3)$$

where Ω_m and Ω_Λ are the present-day density parameters in matter and the cosmological constant, and we have assumed $\Omega_m + \Omega_\Lambda = 1$ (used in the Millennium simulation). We have overlaid this fit using black dotted curves in Fig. 5.3. A comparison of the dashed and solid curves in Fig. 5.3 shows that the mean $\dot{M}_{\rm dif}$ is consistently $\sim 30\%$ higher than the mean $\dot{M}_{\rm mer}$ at all mass and redshifts. The functional form of the mass and redshift dependence in equation (5.3) is therefore also applicable for these separate rates; only the overall amplitude needs to be adjusted to obtain a fitting form for $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$.

The similarity in shape and mass dependence of the $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$ curves in Fig. 5.3 may lead one to conclude that the "non-halo" material contributing to $\dot{M}_{\rm dif}$ can be simply attributed to mergers with sub-resolution haloes. This interpretation is too simplistic, however, as we will discuss in Sections 4 and 5 below where their environmental dependence is analysed.

5.4 The Environmental Dependence of Halo Growth Rates and Histories

5.4.1 Halo Mass Growth Rate due to Mergers vs Diffuse Accretion

In Part I of this series [Fakhouri and Ma, 2009], we reported a strong positive correlation between the *number* of halo mergers and halo environment. We now turn to the environmental dependence of the rate of halo *mass* growth. Since haloes acquire mass through both mergers and diffuse accretion in the form of non-halo mass or unresolved haloes, we examine separately the two growth components defined in equation (5.2).

The minimum mass ratio $\xi_{min} = 0.04$ discussed in §5.3.1 is used as the threshold between $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$. We compute $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$ for each descendant halo and then bin the results by averaging over various ranges of mass and δ. For haloes with only one progenitor (above the mass threshold), there is no contribution from mergers by definition, and $\dot{M}_{\rm mer} = 0$.

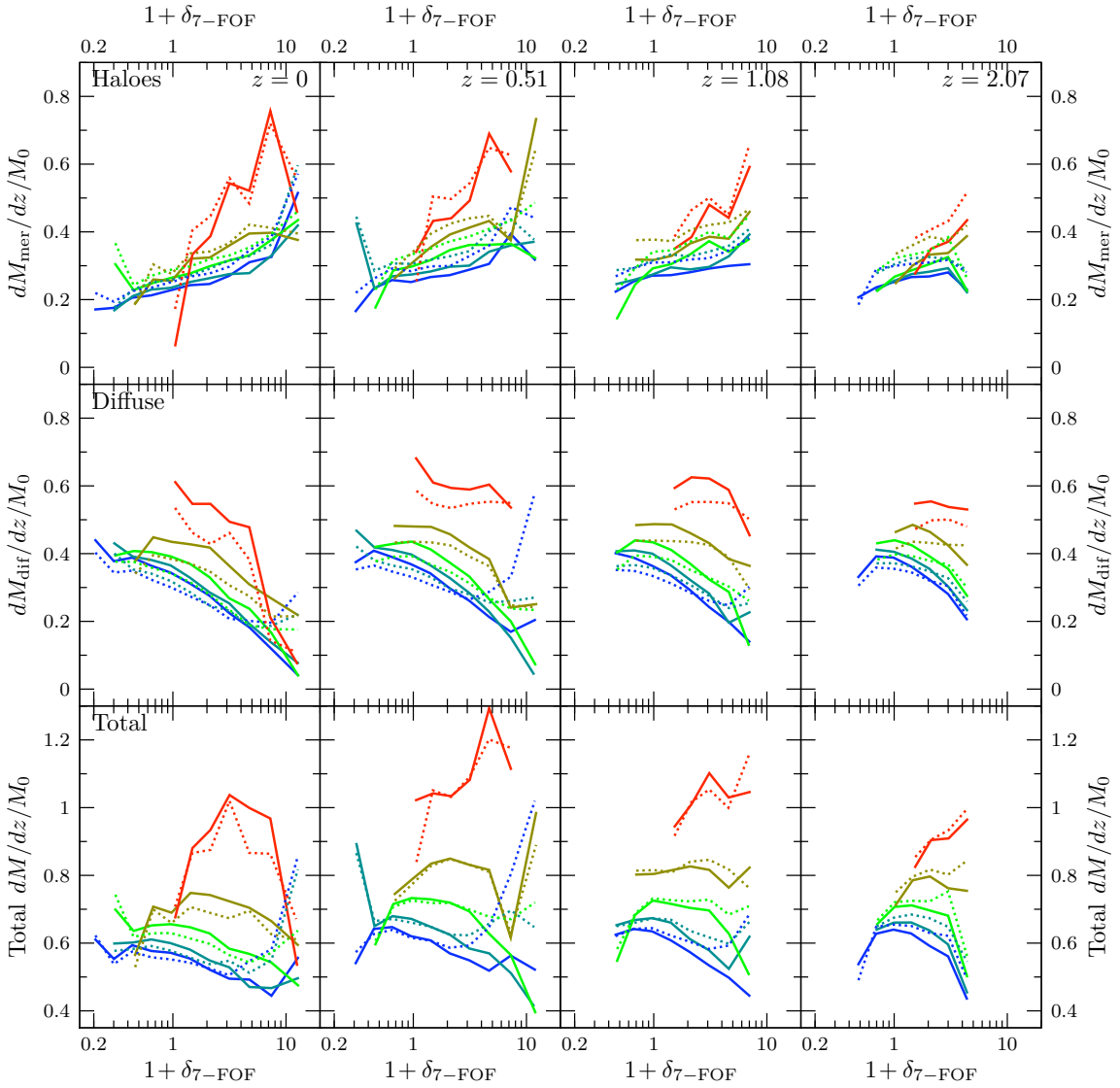

Figure 5.4: Environmental dependence of the halo mass growth rates due to contributions from merging progenitor haloes (top), diffuse accretion of non-halo mass (middle), and sum of the two (bottom). Four redshifts $z = 0, 0.51, 1.08, 2.07$ are shown for comparison (left to right). Within each panel, the five curves show five mass bins for the descendant haloes listed in Table 5.1 (blue: low mass, red: high mass) with solid lines computed using the subhalo mass definition and dashed lines computed using the FOF mass definition. On average, haloes in denser regions experience a higher mass growth rate from mergers and a lower rate from diffuse accretion than similar mass haloes in the voids. These two opposite trends with environment roughly cancel out to yield a weak, but net negative, dependence on $1 + \delta_{7-\mathrm{FOF}}$ for the total growth rate.

These events are included when averaging over $\dot{M}_{\rm mer}$ so that a fair comparison can be made to the mass growth rate due to diffuse accretion, $\dot{M}_{\rm dif}$.

Fig. 5.4 compares the environmental dependence of the mean halo mass growth rate (per unit redshift) due to mergers (top row), diffuse accretion (middle), and the total rate (bottom) computed from the Millennium simulation. The rates at four redshifts are shown: $z = 0, 0.51, 1.08$ and 2.07 (from left to right). Within each panel, different colours correspond to different mass bins (listed in Table 5.1), where more massive haloes have higher growth rates. The solid curves are computed using M_{SH}, the sum of all the subhalo masses in a given FOF halo, while the dotted curves are computed using M_{FOF} (see Sec. 2.3). Both mass definitions yield similar quantitative behaviour, though the subhalo mass definition yields slightly lower merger growth rates in accordance with the results in Genel et al. [2009].

The mass growth due to mergers (top row) is seen to increase in denser regions, in agreement with the higher merger rates in denser regions reported in FM09. This is to be expected as the merger rate B/n is related to the mass growth due to mergers $\dot{M}_{\rm mer}$ by

$$\dot{M}_{\rm mer} \approx \int_{0.04}^{1} M_0 \xi \frac{B}{n}\, d\xi \,. \tag{5.4}$$

As with B/n, the environmental dependence of $\dot{M}_{\rm mer}$ persists out to higher redshift.

By contrast, the middle row of Fig. 5.4 shows that the mass growth rate due to non-halo material has an opposite dependence on δ: the mean value of $dM_{\rm dif}/dz/M_0$ *decreases* with increasing δ for all halo masses at all redshifts. This trend is clean when M_{SH} is used as the mass definition (solid curves) but the curves show a sharp rise for low mass haloes when M_{FOF} is used (blue dotted curve). We believe this is due to the inadequacy of the FOF halo finder in dense regions. As Genel et al. [2009] have shown, the FOF mass for low mass haloes sometimes rise as they approach high mass haloes; this may contribute to the rise in FOF mass observed in the densest regions. Despite this fact, we emphasise that the discrepancies between the M_{FOF} rates and the M_{SH} rates only arise in the densest regions (beyond $1 + \delta_{7-\rm FOF} \sim 5$). Outside of these regions, $dM_{\rm mer}/dz/M_0$ still increases with δ and $dM_{\rm dif}/dz/M_0$ decreases with δ regardless of the halo mass definition.

The mean total mass growth rate is shown in the bottom row of Fig. 5.4. It shows that the strong but opposite dependence on environment found for the two growth rates act against each other, producing a relatively weak δ-dependence for the overall growth rate that declines with increasing δ for the three low mass bins. This dependence is not as clearly seen for the very highest mass bin, though the statistics are poor for the cluster-scale bin (red curves). Again, the trends are much cleaner for M_{SH} than for M_{FOF} in the high-δ regions. The mass dependence of the total growth rates, on the other hand, is more prominent than the environmental dependence, with massive haloes growing more rapidly than low mass haloes. Equation (7) of McBride et al. [2009] (repeated above in eq. (5.3) shows that this mass dependence is well approximated by a power-law: $\dot{M}_{\rm tot} \propto M^{1.127}$.

The curves in Fig. 5.4 are sensitive to the choice of cut used to remove the aberrant haloes (see Sec. 3.2) at the 20-50% level. The qualitative features of the plot, including the

markedly opposite correlation of the two rates with δ, and the negative, but relatively weak, δ dependence of the total rate, are insensitive to choices of f_+ and f_- in the cut and persist even if the cut is made asymmetric ($f_+ \neq f_-$).

It should be pointed out that although the two growth rates depend on the choice of ξ_{min} used to define resolvable haloes and diffuse material (0.04 in this case), the total growth rate is *independent* of this parameter. Thus, redistributing progenitors with $\xi_{min} < 0.04$ from $dM_{\rm dif}/dz/M_0$ to $dM_{\rm mer}/dz/M_0$ would not impact the bottom row of Fig. 5.4. We will return to this point in more detail in §5.5.1.

5.4.2 Fraction of Final Halo Mass Acquired from Mergers vs Diffuse Accretion

In addition to the *instantaneous* halo mass growth rates shown in Fig. 5.4, a related quantity of interest is the *integrated* contribution to a halo's final mass due to mergers vs diffuse accretion. Fig. 5.5 compares these contributions as a function of environment. The lowest set of curves plots the fraction of a halo's final mass gained from major mergers (defined to have progenitor mass ratio $\xi > 1/3$ at the time of merger) through the halo's entire history (out to $z = 6$); the set of curves above it shows the same quantity for more minor mergers with mass ratio $\xi > 1/10$. The five curves in each set show the five halo mass bins listed in Table 5.1. Similar to the top panels of Fig. 5.4, these two sets of curves exhibit a *positive* environmental dependence, where haloes in denser regions experience more mergers and have a higher fraction of their mass coming in from mergers. From voids to overdense regions, the average mass fraction of a halo due to major mergers ($\xi > 1/3$) increases from $\sim 20\%$ to 25%, and the mass fraction due to mergers with $\xi > 1/10$ increases from $\sim 30\%$ to 40%.

The upper two sets of curves show the analogous quantity for the fraction of a halo's final mass gained from material with $\xi < 0.04$ and $\xi < 0.1$. In contrast to the lower two sets of curves, the correlation of the mass contribution by non-halo matter with δ has the opposite sign, again consistent with the middle panels of Fig. 5.4.

A consistent picture is therefore emerging in that haloes in denser regions experience a *higher* rate of growth via mergers and a *lower* rate of growth via diffuse accretion than haloes of comparable mass in less dense regions. As a result, in dense regions, a higher fraction of a halo's final mass is gained through mergers than haloes in the voids. By contrast, diffuse accretion contributes less to halo mass growth in denser regions than in the voids. Overall, however, diffuse accretion (defined as $\xi < 0.04$) is an important component of a halo's final mass regardless of the environment: its contribution to the mass fraction is never lower than $\sim 40\%$ (second set of curves from the top in Fig. 5.5).

5.4.3 Halo Formation Redshifts

We follow the convention in the literature and define the formation redshift z_f for a halo to be the redshift at which the halo's mass first reaches $M_0/2$ (when traced backwards in

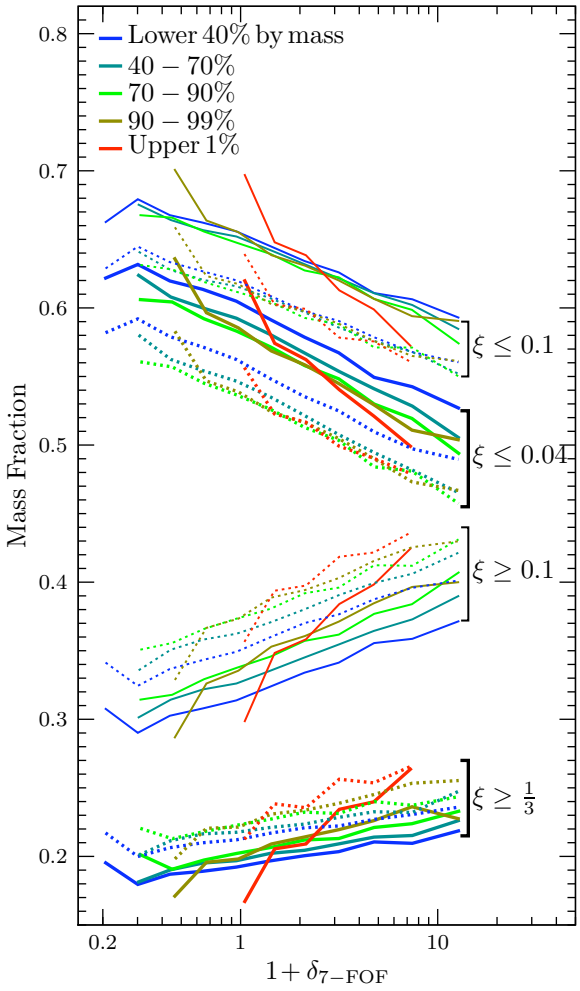

Figure 5.5: Environmental dependence of the fraction of the final halo mass M_0 gained by mergers vs diffuse accretion. The bottom two sets of curves show the rising mass fraction with $1 + \delta_{7-\mathrm{FOF}}$ for mass gained via mergers with progenitors of mass ratios $\xi > 1/3$ and $\xi > 0.1$. The top two sets of curves show the declining mass fraction with $1 + \delta_{7-\mathrm{FOF}}$ for the contributions from non-halo mass with $\xi < 0.04$ and $\xi < 0.1$. Within each set of curves, different colors represent different mass bins, and solid and dotted curves are for the two halo mass definitions M_{SH} and M_{FOF}, respectively. This figure illustrates the increasing importance of mergers to halo growth in overdense regions and importance of diffuse accretion in the voids.

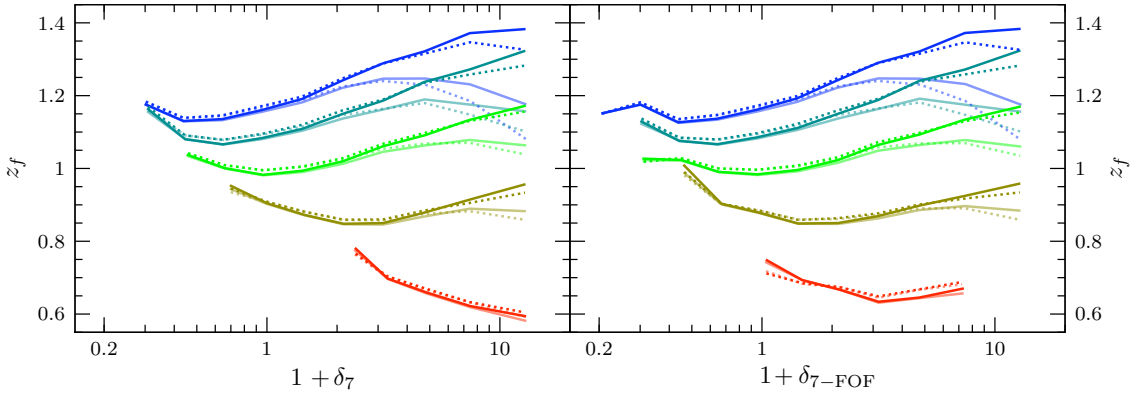

Figure 5.6: Environmental dependence of the mean formation redshifts for haloes in five mass percentile bins (same bins as Fig. 5.5) in increasing mass from top down. Two measures of environment are shown for comparison: $1 + \delta_7$ that includes the halo mass (left), and $1 + \delta_{7-FOF}$ that excludes the central halo. Solid lines are computed using M_{SH} and dotted lines using M_{FOF}. The darker colors are computed after the cut described in Section 5.3.2 is applied; the light colors are computed without the cut. This figure shows that more massive haloes (bottom curves) on average form more recently. Within a mass bin, galaxy-size haloes in denser regions form earlier than those in the voids, while cluster-size haloes show little environmental dependence in their formation time.

time), where M_0 is measured at $z = 0$. This is computed by tracing the mass of each of the $\sim 500,000$ haloes from low to high redshift in the Millennium outputs. A complementary z_f can be defined by going from high to low redshift; we have done this as well and found little difference in our results.

Fig. 5.6 shows the mean formation redshift for haloes in five mass bins (in increasing mass from top down) as a function of their environment at $z = 0$. The two panels compare two definitions of the local densities discussed in Fakhouri and Ma [2009]: $1 + \delta_{7-\mathrm{FOF}}$ used throughout this paper (right) and the simpler $1+\delta_7$ (left). The two measures of local densities give very similar results except in the most massive halo bin (red curves), where z_f shows little correlation with $\delta_{7-\mathrm{FOF}}$ but *decreases* with increasing δ_7. The negative correlation of z_f with δ_7 should not be interpreted as a true dependence on a halo's surrounding environment; instead, it is due to the tight correlation between δ_7 and the masses of massive haloes shown in Fig. 2 of FM09 and the fact that, on average, more massive haloes form earlier than less massive haloes. This degeneracy between mass and density of massive haloes is removed when the variable $\delta_{7-\mathrm{FOF}}$ is used (see Fig. 2 of FM09). The resulting z_f vs $\delta_{7-\mathrm{FOF}}$ shows almost no correlation in the right panel of Fig. 5.6.

The solid and dotted curves in Fig. 5.6 compare the mean values of z_f computed using the M_{SH} vs M_{FOF} definition of halo mass. As seen in Fig. 5.4, the two mass definitions yield similar results in underdense regions but differ in the densest regions. Fig. 5.6 also compares z_f computed with (darker shades) and without (lighter shades) the cut described in Section 5.3.2. The two z_f's are identical in underdense regions, but the sample without the cut has a lower z_f at high δ. This drop is unphysical and is due to the orhpaned halo fragments that suddenly appear at low z and thus bias z_f to artificially low values.

A number of conclusions can be drawn from Fig. 5.6. First, within a given environment, higher mass haloes on average form later (i.e. lower z_f). For instance, the mean formation redshift drops from $z_f \sim 1.1$ to 1.4 for $M \sim 10^{12} M_\odot$ haloes, to about 0.7 for $M > 10^{14} M_\odot$ haloes. This is a well known property of hierarchical cosmological models such as the ΛCDM. This mass dependence is also consistent with that shown in the bottom panels of Fig. 5.4, where more massive haloes have a larger mean mass growth rate and hence a later mean formation time.

The second feature to note is that galaxy-size haloes (top few curves) in denser regions form earlier than the same mass haloes in the voids, whereas z_f for cluster-size haloes (bottom curve) hardly correlates with environment (when $\delta_{7-\mathrm{FOF}}$ is used as a measure). The earlier formation of galaxy haloes in denser regions is consistent with the halo assembly bias discussed in a number of recent papers [Gao et al., 2005, Sheth and Tormen, 2004, Jing et al., 2007, Gao and White, 2007, Harker et al., 2006, Wechsler et al., 2006, Wetzel et al., 2007, Wang et al., 2007, Hahn et al., 2008, McBride et al., 2009]. Interestingly, Wechsler et al. [2006] and Jing et al. [2007] reported decreasing z_f with increasing halo bias for $10^{13}-10^{14} M_\odot$ haloes. This trend is similar to the trend for the cluster-mass haloes (red curve) in the left panel of our Fig. 5.6. Since the connection between δ and halo bias is complex, it is not immediately clear that their results are inconsistent with ours. Perhaps the negative correlation observed by these authors is also due to the fact that the bias of high mass haloes is

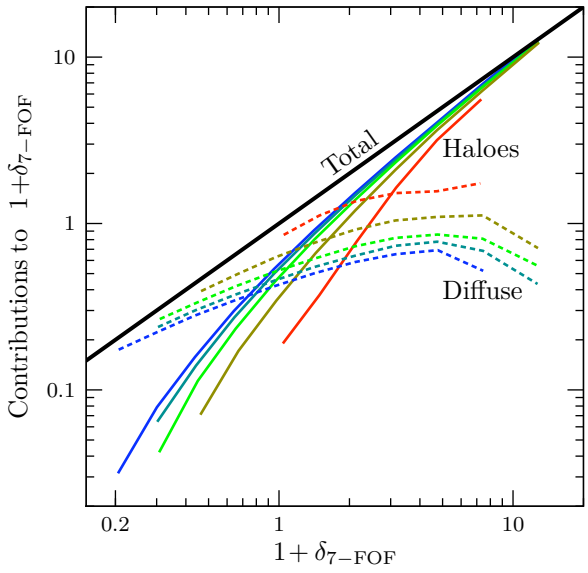

Figure 5.7: Composition of the mass reservoir in the region outside of the halo but within a 7 Mpc sphere centered at the halo. The vertical axis shows the contribution to $1 + \delta_{7-\mathrm{FOF}}$ by haloes (with mass above 4% of the central halo; solid curves) and the remaining material (dashed lines). By construction, the solid and dashed curves add up to the thick diagonal line. Different colours denote the same mass bins in earlier figures. The mass reservoir for haloes in overdense regions is seen to be dominated by resolved haloes, whereas the reservoir for haloes in the voids is dominated by non-halo material.

strongly correlated with mass and does not provide an independent measure of environment.

The third feature to note in Fig. 5.6 is the slight upturn of z_f in the most underdense void region. This trend makes intuitive sense since haloes in sufficiently underdense regions are deprived of fuel for growth, thereby growing more slowly and having larger formation redshifts. This trend does not seem to appear in any of the earlier literature except Desjacques [2008], which noticed a similar trend in Harker et al. [2006]. We suspect that the use of bias as a measure of halo environment in most of these earlier papers (as opposed to the local density used here and in Harker et al. 2006) is not a sensitive probe of the z_f statistics of haloes in the voids.

5.4.4 Mass Reservoir outside Haloes

In this subsection, we address the question: are the opposite environmental correlations of $\dot{M}_{\mathrm{mer}}$ and $\dot{M}_{\mathrm{dif}}$ with δ seen in Figs. 5.4 and 5.5 also present in the surrounding mass reservoirs outside of the haloes? If yes, then the different environmental dependence of the growth rates is originated from the mass reservoirs that are feeding the haloes. If not, then

different accretion timescales have to be operating for the haloes and non-halo material in overdense vs underdense regions. The answer to this question will therefore help explain the *origin* of the difference in the environmental correlations reported thus far.

To answer this question, we analyse the halo vs non-halo composition of the mass in the volume used to compute $\delta_{7-\mathrm{FOF}}$ for each halo at $z = 0$. Centered at each halo, we compute the mass within the 7 Mpc sphere (excluding the FOF mass itself) that resides in haloes with mass ratios exceeding $\xi > 0.04$ (relative to the central FOF mass) vs in all other matter. The results are shown in Fig. 5.7, which plots the mean values of these two components in the mass reservoir as a function of $1 + \delta_{7-\mathrm{FOF}}$ for five mass bins. The black diagonal line marks the sum of the two components, which by definition is simply $1 + \delta_{7-\mathrm{FOF}}$.

Both the halo and non-halo components in the 7 Mpc reservoir outside the haloes are seen to increase with $1 + \delta_{7-\mathrm{FOF}}$, but the halo component has a much steeper slope. The two components contribute a comparable amount for haloes in the cosmic mean density (i.e. $\delta_{7-\mathrm{FOF}} \sim 0$), but up to 80% of the mass reservoir is made of haloes in very dense regions, while this fraction drops to $\sim 20\%$ in voids, where the diffuse component dominates. Even though only a fraction of this mass reservoir is accreted onto the central haloes and contributes to actual halo growths, Fig. 5.7 suggests that the higher merger rates in denser environments are a direct consequence of the larger fraction of available haloes in the reservoir.

A somewhat subtle but important feature to note in Fig. 5.7 is that the *absolute* amount of diffuse mass in the reservoir *rises* with increasing $\delta_{7-\mathrm{FOF}}$ even though the *fractional* contribution due to this component drops. This rising trend with $\delta_{7-\mathrm{FOF}}$ of the diffuse component is in striking contrast to the middle panels of Fig. 5.4, where the accretion rates of diffuse material, $\dot{M}_{\mathrm{dif}}$, onto haloes are seen to *decrease* in denser regions. The naive interpretation that $\dot{M}_{\mathrm{dif}}$ decreases with increasing $\delta_{7-\mathrm{FOF}}$ is due to a dwindling reservoir of diffuse material outside of the haloes in denser regions is therefore invalid. On the contrary, at a fixed halo mass, there is more non-halo material available in the regions where $\dot{M}_{\mathrm{dif}}$ is lower.

Additional gravitational effects are likely to be at work in dense environments to explain why the halo mass growth rate due to diffuse material is stunted in denser regions while the supply of this mass is in fact larger. For instance, the non-halo component in the reservoir may be dynamically hotter than the resolved haloes due to tidal heating and stripping, thereby reducing $\dot{M}_{\mathrm{dif}}$ in higher density regions as shown in Fig. 5.4. Pieces of this picture have been suggested in the literature. Wang et al. [2007] noted that a halo's ability to consume the available fuel is as important as the amount of fuel available for growth. They found that haloes in dense regions accreted matter more slowly than expected, although they did not examine the separate $\dot{M}_{\mathrm{mer}}$ and $\dot{M}_{\mathrm{dif}}$ components as we do here. It has also been suggested that low-mass haloes form earlier because their late accretion is eliminated in a competition with massive neighbors [Zentner, 2007]. Fig. 5.4, however, shows that this competition idea is not strictly true: regardless of mass, all haloes in overdense regions have both increased growth rate $\dot{M}_{\mathrm{mer}}$ and decreased growth rate $\dot{M}_{\mathrm{dif}}$. The growth rates of low mass haloes are therefore not simply controlled by competition with nearby high-mass haloes in overdense regions. Alternatively, the diffuse material surrounding very dense regions may preferentially accrete onto the lower-density filaments and the haloes within them, which

then infall onto the clusters. Further analysis taking into account of filaments as well haloes would be required to test this idea.

5.4.5 Time Evolution of a Halo's Environment

Thus far we have used the mass densities computed at a halo's redshift to quantify its environment. Here we examine whether the environments of haloes evolve significantly over their lifetimes, that is, if the haloes in overdense regions today tended to have progenitors that also resided in overdense regions at earlier times.

To study such environmental evolution, we first group the haloes at $z = 0$ into mass bins. Within each mass bin, the haloes are further divided into five environmental bins, each corresponding to 20 percentile in the distribution of $1 + \delta_{7-\mathrm{FOF}}$. Each halo is therefore assigned to a mass and a $\delta_{7-\mathrm{FOF}}$ bin. For each halo at $z = 0$ we then identify its most massive progenitor at $z = 0.51$ and $z = 1.08$. We assign each progenitor a mass bin at its redshift and compute the $\delta_{7-\mathrm{FOF}}$ bin to which it belongs, using the density field at that given z.

We then compute the fraction of haloes at $z = 0$, as a function of mass bin, whose most massive progenitors reside in the *same* percentile δ bin at $z = 0.51$ and 1.08. Such haloes have not deviated from their local environment significantly over time. Since δ evolves with time, it is more meaningful to compare relative percentiles instead of absolute values of δ at different redshifts. We also compute the fraction of haloes at $z = 0$ that are at most one percentile bin away at $z = 0.51$ and 1.08 . These haloes have moved slightly outside of their $z = 0$ environmental percentile.

We find that $\sim 70\%$ of $M \lesssim 10^{13} M_{\odot}$ haloes do not change environmental bins between $z = 0.51$ and today, and 60% do not change environments between $z = 1.08$ and today. In addition, more than 95% of these haloes are within one δ-percentile bin of their final environment. The range of δ for cluster-mass haloes is too narrow (see, e.g., Fig. 6) for this analysis to be meaningful. Thus, the majority of haloes reside in the same environmental context across their lifetimes.

5.5 Discussion

5.5.1 Interpreting "Diffuse" Accretion

In Sections 3 and 4, we presented results for the mass accretion rates of dark matter haloes through mergers ($\dot{M}_{\mathrm{mer}}$) and diffuse non-halo material ($\dot{M}_{\mathrm{dif}}$), and their respective environmental dependence. As emphasised there, merging progenitors with a mass ratio of $\xi > \xi_{\mathrm{min}} = 0.04$ were counted towards $\dot{M}_{\mathrm{mer}}$, whereas the rest of the mass growth was counted towards $\dot{M}_{\mathrm{dif}}$.

To test the robustness of the results in Sections 3 and 4 with respect to the value of ξ_{min} used to define haloes vs non-haloes, we take haloes in one of the more massive bins (the 90-99% mass bin, corresponding to 1.4×10^{13} to $1.1 \times 10^{14} M_{\odot}$) that provides good statistics as

well as high mass dynamic range, and recompute the two rates using different values of $\xi_{\rm min}$. The results are shown in Fig. 5.8, which plots the environmental dependence of the two mass growth rates for three values of $\xi_{\rm min}$: 0.4 (black), 0.04 (dark grey; the value used in earlier figures), and 0.004 (light grey). As $\xi_{\rm min}$ is lowered, more of the mass is counted towards the merger component, so $dM_{\rm mer}/dz/M_0$ (solid curves) increases and $dM_{\rm dif}/dz/M_0$ (dashed curves) decreases, while the sum of the two (dotted curve) is, by construction, independent of $\xi_{\rm min}$. We note that even though the overall amplitudes of the two rates change with $\xi_{\rm min}$, $dM_{\rm mer}/dz/M_0$ remains positively correlated and $dM_{\rm dif}/dz/M_0$ remains negatively correlated with $\delta_{7-{\rm FOF}}$ regardless of $\xi_{\rm min}$. This test suggests that the environmental correlations shown in Figs. 5.4 and 5.5 are robust to the value of $\xi_{\rm min}$ used to define haloes vs non-haloes.

We explore further the nature of the diffuse non-halo component by examining the hypothesis that this component is comprised entirely of sub-resolution haloes such that in the limit of $\xi_{\rm min} \to 0$ (or more precisely, the true minimal halo mass, which is likely to be about an Earth mass [Diemand et al., 2005], but the difference is negligible), we would expect $\dot{M}_{\rm dif} \to 0$ and $\dot{M}_{\rm mer} \to M_0 - M_1$. To assess whether this hypothesis is plausible, we recompute $\dot{M}_{\rm mer}$ by extrapolating down to $\xi_{\rm min} = 0$, assuming that $\dot{M}_{\rm mer}$ maintains the same power-law dependence on ξ determined for a given mass bin in the simulation.[1]

The predicted mass growth rate under this assumption is plotted as a function of environment in Fig. 5.8 (top solid red curve). In comparison with the total rate $\dot{M}_{\rm mer} + \dot{M}_{\rm dif}$ (dotted curve) determined from the Millennium simulation, we observe a clear environment-dependent gap between these two curves that widens in the lower-density regions. This gap suggests two possible scenarios: (1) the power-law behaviour of $\dot{M}_{\rm mer}$ steepens towards low ξ (an effect that is not observed down to $\xi \sim 10^{-4}$ resolvable by current simulations), or (2) there exists a truly diffuse component of dark matter that does not belong to bound haloes. Moreover, both scenarios must be dependent on environment. For (1), the δ-dependent gap in Fig. 5.8 can only be closed if the power-law slope of $\dot{M}_{\rm mer}$ vs ξ for sub-resolution progenitors is a strong function of halo environment (being steeper in underdense regions). We cannot rule out this possibility without testing it in higher-resolution simulations, but all the halo properties that we have examined thus far in FM08 and FM09 do not exhibit such complex, non-universal behavior. On the contrary, down to the current simulation resolution, the halo-halo merger rates and mass growth rates as a function of progenitor mass ratio ξ were all well fit by a single power law, where the slope was entirely independent of the halo environment (see Figs. 4, 5, and 7 of FM09). We therefore favour scenario (2) that the diffuse non-halo mass has a truly diffuse component, and this component is more prominent in lower-density regions. New simulations with higher mass resolution such as the Millennium-II [Boylan-Kolchin et al., 2009b] will be used to explore further these two possibilities.

[1] We note that our best-fit slope for the ξ-dependence of $\dot{M}_{\rm mer}$ in Fakhouri and Ma [2008] was in fact slightly steeper than -2, which would give a divergent growth rate. Our goal there was to obtain a simple universal form. For better accuracy, here we fit $\dot{M}_{\rm mer}$ vs ξ for each mass bin independently, where the slope had a very mild mass dependence, ranging from -1.7 to -1.9.

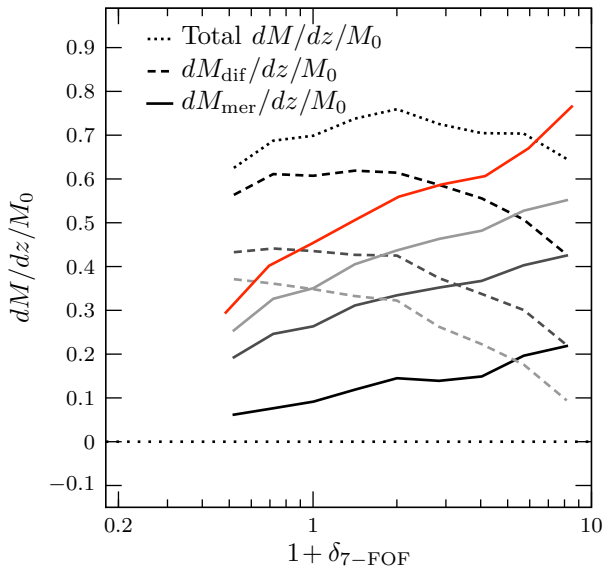

Figure 5.8: Dependence of the two growth rates $dM_{\rm mer}/dz/M_0$ (solid) and $dM_{\rm dif}/dz/M_0$ (dashed) on the mass threshold $\xi_{\rm min}$ used to define the halo and non-halo components. The sum of the two rates (dotted), by construction, is independent of $\xi_{\rm min}$. Haloes in the $90-99\%$ mass bin are used. As $\xi_{\rm min}$ is lowered from 0.4 (black), 0.04 (dark grey), to 0.004 (light grey), the amplitude of $dM_{\rm mer}/dz/M_0$ increases and that of $dM_{\rm dif}/dz/M_0$ decreases, but their respective correlations with $\delta_{7-\rm FOF}$ remain the same, suggesting that the environmental dependences reported in this paper are robust and do not depend on the exact values of $\xi_{\rm min}$ used to define haloes vs non-haloes. The top solid curve (in red) shows the hypothetical value of $dM_{\rm mer}/dz/M_0$ extrapolated down to $\xi_{\rm min}=0$ (see text for discussion). One explanation for the δ-dependent gap between this curve and the total rate (dotted) is that there exists a truly diffuse component of dark matter that does not belong to bound haloes, in particular in the lower-density regions.

5.5.2 Implications for the Extended Press-Schechter and Excursion Set Models

The Press-Schechter model [Press and Schechter, 1974] assumes that primordial Gaussian-distributed dark matter perturbations collapse into haloes of mass M when their linearly extrapolated overdensities, smoothed on a scale corresponding to a given mass, exceed a critical threshold computed using a simple spherical collapse model.

This framework was extended to address the problem of halo growth in the excursion set formalism (sometimes called the Extended Press-Schechter or EPS model) [Bond et al., 1991, Lacey and Cole, 1993]. Using a k-space tophat smoothing window, these authors showed that the density perturbation at a given point in space undergoes a random walk as the perturbation is smoothed on smaller and smaller scales. These perturbation "trajectories" obey the diffusion equation and are Markovian, i.e., the change in overdensity at a given scale is independent of the overdensities at other scales. The EPS model in its standard form therefore predicts no environmental dependence. For example, the conditional mass function $\phi(M_1, z_1|M_0, z_0)$ – the distribution of progenitor mass M_1 at redshift z_1 for a descendant halo of mass M_0 at z_0 – is predicted to be the same for all environments, in contrast to Fig. 7 of Fakhouri and Ma [2009] that showed descendant haloes in dense regions to have more progenitors than those in voids.

Recent attempts have been made to introduce environmental correlations into the EPS model. As emphasised in Zentner [2007], the Markovian nature of the random walks in the excursion set model is *not* a prediction but is an assumption originating from the use of the k-space tophat window function. This window function is traditionally chosen to reduce the time required for computing density perturbation trajectories since for other window functions, the perturbations must be computed at all scales simultaneously. When a Gaussian window function was used, Zentner [2007] indeed found an environmental dependence of the halo formation redshift, but the dependence was *opposite* to that seen in numerical simulations (including Fig. 5.6 of this paper), where older (i.e. earlier forming) haloes tended to be more clustered.

Desjacques [2008] instead chose to introduce environmental dependence into the threshold density for collapse in the ellipsoidal excursion set formalism. Overdense regions will tend to exert more tidal shear on collapsing haloes, an effect that causes the haloes to virialize earlier. Like Zentner [2007], however, Desjacques [2008] found the low mass haloes in denser regions to form later than similar mass haloes in emptier regions, again opposite to the effect seen in simulations. Desjacques [2008] did note that his results may be valid in the lowest-density region, where Fig. 5.6 here and Harker et al. [2006] both reported a subtle increase in formation redshift with decreasing δ.

Sandvik et al. [2007] carried out a multi-dimensional generalisation of EPS that allowed them to incorporate environmental effects by tracking the shapes of collapsing haloes and incorporating information about gravitational shear. They found z_f in this approach to depend very weakly on the environment; moreover, the δ dependence they did observe was stronger for more massive haloes, in contrary to the results of N-body simulations such as our

Fig. 5.6. They proposed, instead, that the environmental dependence of a halo's formation history may be related to the halo's progenitor pancakes and filaments. They found that haloes whose progenitors were the most massive pancakes or filaments were more clustered than halo's whose progenitors were the least massive pancakes or filaments, and that this clustering dependence was stronger for *low mass* haloes.

In summary, there is not yet a modification of the EPS model that can successfully predict the correct correlation of halo formation with environment seen in numerical simulations. We emphasize that in addition to the "assembly bias" for the formation redshift discussed in Sec 4.3 and shown in Fig. 5.6, an improved EPS model would also need to predict the strong but opposite environmental dependence of the growth rates shown in Fig. 5.4. A first step towards this goal would be to introduce into the model, and track, the contribution made by the non-halo material $\dot{M}_{\rm dif}$. The model would also need to be sophisticated enough to account for the opposite behaviours of $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$ as a function of the local density.

5.6 Conclusions

We have used the Millennium simulation [Springel et al., 2005] to investigate the dependence of halo mass growth rates and histories on halo environment. This paper complements our previous paper [Fakhouri and Ma, 2009] in which we reported and quantified how haloes in overdense regions experienced higher merger rates than haloes in underdense regions. Here, we have studied the mass growth rates due to both mergers with other resolved haloes, $\dot{M}_{\rm mer}$, and accretion from non-halo "diffuse" material, $\dot{M}_{\rm dif}$. Results for the distributions of $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$ at $z = 0$ and the redshift evolution of the mean $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$ are summarized in Figs. 5.1-5.3.

As a function of halo environment, we have found (see Fig. 5.4) that the growth rate $\dot{M}_{\rm mer}$ due to mergers, in agreement with the merger rates in Fakhouri and Ma [2009], is positively correlated with the local density, whereas $\dot{M}_{\rm dif}$ is negatively correlated. Consequently, in denser regions, mergers play a *relatively* more important role than diffuse accretion for the mass growth of haloes, and a higher fraction of a halo's final mass is acquired through mergers in these regions than in the voids (Fig. 5.5).

We have shown that the origin of the environmental dependence of growth rate due to mergers, $\dot{M}_{\rm mer}$, is directly linked to the mass reservoir immediately outside the virial radii of the haloes (see Fig. 5.7). The mass composition in these surrounding regions exhibits a strong positive correlation with environment such that more than 80% of the mass is in the form of resolved haloes for haloes residing in dense regions (for $1 + \delta_{7-\rm FOF} \gtrsim 5$), while only $\sim 20\%$ of the mass is in resolved haloes for haloes residing in voids (for $1 + \delta_{7-\rm FOF} \lesssim 0.4$). Even though only a fraction of these haloes enters the virial radii and contributes to the actual growth of the central halo, it is the environmental property of this mass reservoir that leads to the environmental dependence of $\dot{M}_{\rm mer}$. For the diffuse growth component $\dot{M}_{\rm dif}$, however, its negative correlation with local density observed in Fig. 5.4 is *not* explained by the environmental dependence of the available diffuse mass in the reservoir outside of the

haloes. In fact, more diffuse material is available in the reservoir in denser regions (dashed curves in Fig. 5.7). Why then is the diffuse accretion rate $\dot{M}_{\rm dif}$ *lower* in denser regions? We speculated that some of the diffuse mass may be dynamically hotter due to tidal stripping and therefore is harder to accrete. A careful analysis using the actual particle data from the simulation would be needed to investigate this issue further.

The halo growth rates $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$ together account for the overall halo mass accretion history (MAH) that we have studied in detail in a separate paper [McBride et al., 2009]. When combined, the opposite correlations of the two rates with the local density cancel each other to some extent, resulting in a weak environmental dependence for the formation redshift z_f (Fig. 5.6). At a fixed mass, galaxy-sized haloes in overdense regions on average form earlier than those in underdense regions, consistent with the assembly bias result reported in several recent papers [Gao et al., 2005, Sheth and Tormen, 2004, Jing et al., 2007, Gao and White, 2007, Harker et al., 2006, Wechsler et al., 2006, Wetzel et al., 2007, Wang et al., 2007, Hahn et al., 2008]. The z_f for cluster-sized haloes, on the other hand, show no dependence on environment. We note that had we ignored $\dot{M}_{\rm dif}$ and only taken into account the mass growth due to mergers, then the positively-correlated δ dependence of $\dot{M}_{\rm mer}$ would imply a later formation redshift (i.e., a smaller z_f) for haloes in denser regions, which is opposite to that shown in Fig. 5.6. The negatively correlated $\dot{M}_{\rm dif}$ with the local density therefore plays a crucial role in counterbalancing the positively correlated $\dot{M}_{\rm mer}$ so that the environmental dependence of the total rate is consistent with that of z_f.

We have emphasized throughout the paper that the values of $\dot{M}_{\rm mer}$ vs. $\dot{M}_{\rm dif}$ depend on the mass threshold $\xi_{\rm min}$ used to defined these two components. Despite this fact, our tests have shown (see Fig. 5.8) that the enviromental trends of $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$ in Figs. 5.4 – 5.7 are robust and are independent of the definition.

A number of interesting questions remain to be explored. For instance, what is the nature of the non-halo component $\dot{M}_{\rm dif}$? Besides tidal stripping, are there additional physical processes controlling the different environmental dependence of $\dot{M}_{\rm mer}$ vs $\dot{M}_{\rm dif}$? How do our results for $\dot{M}_{\rm mer}$ and $\dot{M}_{\rm dif}$ extend to higher-resolution simulations, e.g., will the power-law slope of $\dot{M}_{\rm mer}$ vs progenitor mass ratio ξ remain the same or steepen at smaller ξ, and will this change be δ-dependent? How does one introduce non-Markovian features into the random-walk picture in the much-used EPS model in order to reproduce the halo growth histories as a function of halo mass and halo environment reported in this paper? We expect to shed light on some of these questions by using the Millennium-II simulation [Boylan-Kolchin et al., 2009b] with ~ 1000 times better mass resolution in an upcoming paper.

Part V

Constructing Self-Consistent Theoretical Dark Matter Halo Merger Trees in the EPS Framework

Chapter 6

How to Grow a Healthy Merger Tree

Zhang, Jun ; Fakhouri, Onsi; Ma, Chung-Pei — October 2008
Monthly Notices of the Royal Astronomical Society, Volume 389, Issue 4, pp. 1521-1538

We investigate seven Monte Carlo algorithms – four old and three new – for constructing merger histories of dark matter halos using the extended Press-Schechter (EPS) formalism based on both the spherical and ellipsoidal collapse models. We compare, side-by-side, the algorithms' abilities at reproducing the analytic EPS conditional (or progenitor) mass function over a broad range of mass and redshift ($z = 0$ to 15). Among the four old algorithms (Lacey & Cole 1993, Kauffmann & White 1993, Somerville & Kolatt 1999, Cole et al 2000), we find that only KW93 produces a progenitor mass function that is consistent with the EPS prediction for all look-back redshifts. The origins of the discrepancies in the other three algorithms are discussed. Our three new algorithms are designed to generate the correct progenitor mass function at each timestep. We show that this is a necessary and sufficient condition for consistency with EPS at any look-back time. We illustrate the differences among the three new algorithms and KW93 by investigating two other conditional statistics: the mass function of the $i_{\rm th}$ most massive progenitors and the mass function for descendants with N_p progenitors.

6.1 Introduction

In the hierarchical structure formation scenario, dark matter halos grow by accreting and merging with other halos. Statistically modelling halo merger histories is important for understanding a diverse spectrum of astrophysical processes ranging from galaxy formation, the growth of super-massive black holes, to cosmic reionization.

Numerical simulations aside, the most frequently used theoretical framework for studying the build up of dark matter halos is the Press-Schechter (PS) model [Press and Schechter, 1974]. This framework is further developed in the so-called extended Press-Schechter (EPS) model (Bond et al. 1991, Lacey and Cole 1993, Mo and White 1996, Sheth et al. 2001, Sheth and Tormen 2002). For a descendant halo of a given mass at redshift z_0, the EPS model

predicts the average mass spectrum of its progenitors at a higher redshift z_1 (the *conditional* or *progenitor* mass function).

The EPS model provides only statistical information about halo mergers and does not specify how progenitor halos are to be grouped into descendant halos. However, it is often useful, particularly in semi-analytial modelling, to have actual realisations of the merging history for a large set of haloes. A number of Monte Carlo algorithms have been proposed for this purpose (see, *e.g.* , Lacey and Cole 1993, Kauffmann and White 1993, Sheth and Pitman 1997, Sheth and Lemson 1999, Somerville and Kolatt 1999, Cole et al. 2000, 2008, Moreno et al. 2008, Neistein and Dekel 2008a). These algorithms allow one to produce realisations of halo merger trees stretching back to high redshifts in a fraction of the time that is required for performing and analysing cosmological N-body simulations of comparable resolution.

Thus far, most of the commonly used Monte Carlo methods are based on the spherical EPS theory. In Lacey and Cole 1993 (also see Bond et al. 1991), halo mergers at each time step are assumed to be binary: one of the progenitor masses is randomly drawn from the conditional mass function, and the other progenitor mass is defined by the difference between the descendant halo mass and this first progenitor mass. Though this seems to be the most natural way to generate halo merger histories, it has been pointed out by several authors that the binary picture does not reproduce the EPS progenitor abundance at earlier times (see, *e.g.* , Somerville and Kolatt 1999). Moreover, this problem does not disappear when the time step is greatly reduced. This fact has led to the investigation of alternative Monte Carlo algorithms with different recipes for building halo merger trees in the spherical EPS framework. For example, Somerville & Kolatt (1999) find that if the binary assumption is relaxed while taking into account the contribution of mass from continuous accretion then the progenitor abundance at large look-back times is better reproduced. Cole et al. (2000), on the other hand, include diffuse accretion but preserve the assumption of binary mergers. More recently, partially due to the rapid advances in N-body simulation, various other algorithms have been proposed that are either designed to fit N-body results (*e.g.* , Parkinson et al. 2008, Cole et al. 2008, Neistein and Dekel 2008b) or are based on the spherical [Neistein and Dekel, 2008a] or ellipsoidal [Moreno et al., 2008] excursion set model. The presence of these numerous Monte Carlo algorithms suggests that building a Monte Carlo algorithm that is fully consistent with the underlying EPS model is not unique and can be non-trivial.

We were motivated to write this paper for a number of reasons. First, this is a sequel to our previous work [Zhang et al., 2008b], which presented an accurate analytic formula for the conditional mass function for small timesteps in the ellipsoidal EPS model. This formula is particularly useful as an input for high-resolution Monte Carlo simulations of halo merger trees. Earlier formulae (e.g. Sheth and Tormen 2002) were accurate only for larger look-back redshifts ($z_1 - z_0 \gtrsim 0.1$). Taking such a large timestep would limit the dynamic range in both the progenitor mass and redshift that can be covered in a Monte Carlo simulation. In addition, until recently, all previous Monte Carlo algorithms were studied in the framework of the spherical EPS model, which is well known to produce inaccurate total (i.e. unconditional) halo mass function when compared with simulations. This paper will

investigate the algorithms in the ellipsoidal model using the formula in Zhang et al. [2008b].

Second, as we began to investigate the various Monte Carlo algorithms proposed in the literature, we were frustrated by the lack of direct comparison among the different methods, each of which has its own range of validity and own set of assumptions about how to group progenitors into descendants (e.g. binary vs multiple progenitors; how the mass in progenitors below mass resolution is treated). Moreover, it was not always clear why a given algorithm succeeded or failed. In this paper, we examine closely the four most frequently used algorithms – Lacey and Cole 1993, Kauffmann and White 1993, Somerville and Kolatt 1999, Cole et al. 2000 – and compare their predictions for the conditional mass function over a wide range of progenitor masses and look-back redshift (e.g., down to 10^{-6} of descendant mass and up to redshift 15). We find that only Kauffmann and White [1993] is fully consistent with the EPS model at all look-back time steps. The limitations and causes of discrepancies in the other three methods are discussed.

Third, in light of the discrepancies in earlier algorithms, we investigate three new Monte Carlo algorithms that are all constructed to reproduce accurately the EPS predicted conditional mass function at any look-back redshift. We present a consistency criterion that is useful as a general guide for building Monte Carlo algorithms: If an algorithm reproduces the EPS progenitor mass function for a sequence of simulation timesteps between z_i and z_{i+1} (where $i = 0, N$), then it is guaranteed to reproduce the EPS progenitor mass function at any z_j for descendants at any later z_k (where $j, k = 0, N$). This is a necessary and *sufficient* condition.

Fourth, the EPS model is an incomplete theory that predicts only a subset of statistical properties of halo mergers. It therefore leaves one with much freedom in how to assign progenitors to descendants in a given Monte Carlo algorithm. For instance, it is possible to construct different consistent Monte Carlo algorithms that predict different statistical merger quantities beyond the conditional mass function. Our three new algorithms and KW93 are four examples that are degenerate in the conditional mass function but are different in other progenitor statistics. In this paper we illustrate the differences among the models with two such statistics: the mass function of the i^{th} most massive progenitors and the mass function of progenitors for descendant halos with N_p progenitors. Results from N-body simulations will be needed to constrain these higher-moment statistics. Since computing the statistics of progenitor dark matter halos in simulations is by itself a major independent project, we will focus on the EPS theory and Monte Carlo algorithms in this paper and leave the comparison with N-body results to a subsequent paper (Zhang, Fakhouri & Ma 2008, in prep).

The paper is structured as follows. The EPS formalism based on both the spherical and ellipsoidal gravitational collapse models is reviewed in §6.2. In §6.3 we discuss three ingredients for how to grow an accurate Monte Carlo merger tree: the consistency criterion for reproducing EPS (§6.3.1), the asymmetry in the EPS progenitor mass function and the necessity of non-binary mergers in an algorithm (§3.2), and the role of mass resolution and diffuse accretion for progenitor mass assignment (§3.3). Details of the four old and three new algorithms are discussed in §4 and §5, respectively. Whenever possible, the resulting progenitor mass functions from different algorithms are shown on the same plots for ease of

comparison. §6 compares the two new progenitor statistics that can be used to discriminate among the Monte Carlo algorithms that are consistent with EPS. We summarise our findings in §6.7, with a discussion of some recent work in this field.

The calculations in this paper assume a ΛCDM model with $\Omega_m = 0.25$, $\Omega_b = 0.045$, $h = 0.73$, $\Omega_\Lambda = 0.75$, $n = 1$, $\sigma_8 = 0.9$. This is the same cosmology used in the Millennium simulation (Springel et al. 2005).

6.2 An Overview of EPS

In this section we present a brief overview of the EPS theory based on both the spherical and ellipsoidal gravitational collapse models. We often refer to the two models in parallel as the spherical and ellipsoidal EPS models, with the understanding that the ellipsoidal version is based on the excursion set formalism of Bond et al. [1991]. The emphasis here is on the conditional mass function, which is the main statistical quantity used to generate progenitors in merger tree algorithms. For a more complete and pedagogical review of EPS, see Zentner [2007] and references therein.

6.2.1 EPS Based on the Spherical Collapse Model

The Press-Schechter (PS) model provides a framework for identifying virialized dark matter halos. It is assumed that the seed density perturbations that grow to form these halos are characterised by an initially Gaussian random density field with larger fluctuations on smaller spatial scales. This latter assumption allows one to use $S(R) = \sigma^2(R)$, the variance of the linear density fluctuations[1] smoothed over spatial scale R, as a proxy for the spatial scale R. Moreover, since a given spatial scale is related to a unique mass scale $M(R)$ via the mean density of the universe $\bar{\rho}$, one can use R, M, and S interchangeably as measures of scale.

The density field smoothed over a given scale M is given by $\delta_M = \rho_M/\bar{\rho} - 1$ where ρ_M is the average density within the smoothing scale R. In the EPS model, the linear density field centred at a given point in the initial Lagrangian space traces out a random walk (referring to a Markovian process)[2] as the smoothing scale is reduced. Starting from a large smoothing scale, a virialized dark matter halo is assumed to form at the given spatial coordinate when the linear δ_M crosses a critical value for the first time; the mass of the halo is determined by the smoothing scale at first-crossing. In the spherical EPS model, the critical over-density is given by the spherical collapse model and is a constant $\delta_c = 1.69$ independent of mass scale.

In the above description, as a result of gravitational instability, the density field grows with time as a linear function of its initial value, *i.e.* , $\delta_M(z) = \delta_M(0)D(z)$, where $D(z)$ is the standard cosmology-dependent linear growth factor satisfying $D(z = 0) = 1$. In practice,

[1] In this paper, the variance of the density fluctuation is calculated using the fitting formula of the linear mass power spectrum from Eisenstein and Hu 1998

[2] Strictly speaking, this is only true when the smoothing window function is a top-hat in Fourier space.

one usually fixes the value of δ_M at some reference time (*e.g.* today: $\delta_M(0)$) and evolves the critical over-density to identify virialized halos at earlier redshifts. We denote this time-dependent critical over-density by $\omega(z) = \delta_c/D(z)$. Note that a lower redshift corresponds to a smaller $\omega(z)$, implying that larger halos form at later times, in accordance with the hierarchical structure formation scenario.

Under the assumption of Gaussian statistics, the EPS framework allows one to compute the first crossing distribution $f(S(M_1), z_1|S(M_0), z_0)$. Of the set of random walks that begin at $\delta_{M_0} = \omega(z_0)$, the first crossing distribution is the fraction of these random walks that *first* cross the critical over-density $\omega(z_1)$ at scale $S(M_1)$, where $z_1 > z_0$ and $S(M_1) > S(M_0)$ (*i.e.* $M_1 < M_0$). It can be shown [Lacey and Cole, 1993] that the first crossing distribution in the spherical EPS model has the form

$$\begin{aligned} & f(S(M_1), z_1|S(M_0), z_0)d\Delta S \\ = \quad & \frac{1}{\sqrt{2\pi}} \frac{\Delta\omega}{\Delta S^{3/2}} \exp\left(-\frac{\Delta\omega^2}{2\Delta S}\right) d\Delta S \end{aligned} \tag{6.1}$$

where $\Delta\omega = \omega(z_1) - \omega(z_0)$ and $\Delta S = S(M_1) - S(M_0)$.

The first crossing distribution can be reinterpreted as the conditional mass function $P(M_1, z_1|M_0, z_0)$, which is defined to be the mass fraction of a descendant halo of mass M_0 at redshift z_0 that originates from a progenitor halo of mass M_1 at redshift z_1:

$$P(M_1, z_1|M_0, z_0)dM_1 = -f(S(M_1), z_1|S(M_0), z_0)d\Delta S \tag{6.2}$$

Note, in particular, that $P(M_1, z_1|M_0, z_0)$ is the *mass-weighted* conditional mass function as it represents the merging history of a unit of mass. The average number of progenitors of mass M_1 at z_1 associated with the formation of *each* descendant halo of mass M_0 at z_0 is given by the *number-weighted* conditional mass function $\phi(M_1, z_1|M_0, z_0)$, which is simply related to the mass-weighted conditional mass function $P(M_1, z_1|M_0, z_0)$ by

$$\phi(M_1, z_1|M_0, z_0) \equiv \frac{M_0}{M_1} P(M_1, z_1|M_0, z_0)\,. \tag{6.3}$$

For brevity, we often refer to the number-weighted conditional mass function $\phi(M_1, z_1|M_0, z_0)$ as the *progenitor* mass function, **and denote it simply as $\phi(M_1|M_0)$ with z_0 and z_1 specified elsewhere in the rest of the paper**. This quantity is sometimes denoted as $dN(M_1, z_1|M_0, z_0)/dM_1$ in the literature.

6.2.2 EPS Based on the Ellipsoidal Collapse Model

The original Press-Schechter theory was based on the spherical collapse model. The unconditional mass function in this model is well known to have an excess of small halos and a deficit of massive halos in comparison with simulation results (*e.g.* , Lacey and Cole 1994, Gelb and Bertschinger 1994, Ma and Bertschinger 1994, Tormen 1998, Sheth and Tormen

1999). This discrepancy arises because halo collapses are generally triaxial rather than spherical (Bardeen et al. 1986, Sheth et al. 2001, Sheth and Tormen 2002). In the spherical collapse picture, the virialization of a dark matter halo is purely determined by the density-contrast on the scale of the halo mass. This assumption is too simplistic because dark matter halos generally have non-zero ellipticity and prolateness, and the condition for virialization should be determined by both the density-contrast and the halo shape parameters.

By assuming that a dark matter halo virializes when its third axis collapses, Sheth et al. [2001] find a new criterion for virialization that depends on the ellipticity and prolateness of a dark matter halo in addition to its density contrast. In practice, this condition can be simplified either by averaging over its dependence on the shape parameters, or by fixing the shape parameters at their most likely values for a given over-density. By doing so, these authors obtain a fitting formula for the scale-*dependent* critical over-density, or barrier, in contrast to the scale-independent δ_c of the spherical collapse model. In this ellipsoidal collapse model, the scale-dependence is such that the formation of small halos is delayed, thereby reducing their abundance and providing closer agreement with the unconditional mass function in simulations than the spherical model.

To compute the *conditional* mass function in the ellipsoidal EPS model, one would need the equivalent of the first-crossing distribution eq. (6.1). The exact analytical form of eq. (6.1), unfortunately, is valid only for the scale-independent constant barrier δ_c of the spherical EPS model. Sheth and Tormen [2002] have presented a Taylor-series-like approximation for the ellipsoidal model, but Zhang et al. [2008b] show that this form works well for large $z_1 - z_0$ but is invalid for small $z_1 - z_0$. As the construction of an accurate ellipsoidal Monte Carlo merger tree algorithm requires accurate knowledge of the ellipsoidal progenitor mass function at small timesteps, it is crucial that this matter be resolved.

This was done in Zhang et al. [2008b]. Using the scale-dependent critical over-density of Sheth and Tormen [2002] and the technique of Zhang and Hui [2006], Zhang et al. [2008b] derived an accurate form for the progenitor mass function of ellipsoidal EPS model for small time steps ($\Delta z \lesssim 0.1$), which can be written as:

$$\begin{aligned}\phi(M_1|M_0) &= \frac{M_0}{M_1}\frac{dS(M_1)}{dM_1}\frac{B_0}{\Delta S\sqrt{2\pi\Delta S}} \\ &\times \left\{\exp\left[-\frac{(B_0+B_0'\Delta S)^2}{2\Delta S}\right] + \exp\left(-\frac{B_0'^2\Delta S}{2}\right)\right. \\ &\times \left. 0.0373\nu_0^{-0.115}\left(\frac{\Delta S}{S(M_0)}\right)^{3/2}\left[1+\frac{B_0'\sqrt{\Delta S}}{\Gamma(3/2)}\right]\right\}\end{aligned} \tag{6.4}$$

where

$$\begin{aligned} \Delta S &= S(M_1) - S(M_0), \\ B_0 &= 0.866[\omega(z_1) - \omega(z_0)](1 - 0.133\nu_0^{-0.615}), \\ B_0' &= 0.308\nu_0^{-0.115} S(M_0)^{-0.5}, \\ \nu_0 &= \omega^2(z_0)/S(M_0). \end{aligned} \tag{6.5}$$

Note that unlike eq. (15) of Zhang et al. [2008b], we have not neglected the small B_0 in the exponent because it is important for tracing the massive progenitors (small ΔS). The coefficients in eq. (6.4) are closely related to the scale dependence of the critical over-density of Sheth and Tormen [2002]. The details can be found in Zhang et al. [2008b]. Eq. (6.4) provides a closer match to the merger rates determined from N-body simulations [Zhang et al., 2008b], but the agreement was not perfect, perhaps due to the non-Markovian nature of numerical simulations.

6.3 Ingredients for Growing Healthy Merger Trees

As discussed in the introduction, the EPS model only provides a subset of *statistical* information about dark matter halo merger histories. For example, the EPS progenitor mass function $\phi(M|M_0)$ (eq. 6.3 for spherical and eq. 6.4 for ellipsoidal) gives the *average* mass spectrum of the progenitors for the descendant halos. However, it is often useful, especially in semi-analytical modelling, to have an actual Monte Carlo realisation of the formation history for a large set of halos. Of particular interest is the merger tree of individual halos, which provides the hierarchical links among the progenitors and their descendants. Since the EPS model itself does not specify explicitly how to group progenitors into descendants, in each timestep in a Monte Carlo algorithm, assumptions must be made about the number of progenitors and their mass distributions to be assigned to a given descendant.

The earlier Monte Carlo algorithms (e.g., Lacey and Cole 1993, Kauffmann and White 1993, Somerville and Kolatt 1999, Cole et al. 2000) for merger tree constructions share a similar overall structure: A descendant halo of mass M_0 at some redshift z_0 (typically $z_0 = 0$) is chosen. The EPS progenitor mass function, $\phi(M|M_0)$, is then used to generate a set of progenitors at some earlier redshift, using the rules of the given algorithm. In the next timestep, these progenitors become descendants, and each is assigned its own set of progenitors at an earlier redshift using $\phi(M|M_0)$. This process is repeated out to some early redshift and for a (typically large) number of halos of mass M_0 at the starting z_0.

The existence of a number of diverse Monte Carlo algorithms (see further discussion in §6.4) in the literature implies that the above process is, in fact, not unique and can be quite subtle. We now explore some of these subtleties and the key ingredients for constructing a healthy merger tree.

6.3.1 A Criterion for Consistently Reproducing the EPS Progenitor Mass Function

We consider a Monte Carlo algorithm to be *consistent* with EPS if the merger trees it produces can reproduce the EPS progenitor mass function $\phi(M_1, z_1|M_0, z_0)$ exactly for *any* set of $\{M_1, z_1, M_0, z_0\}$ regardless of the number or size of the simulation timesteps between z_0 and z_1.

Clearly, to be consistent with EPS, a Monte Carlo algorithm must *necessarily* reproduce the EPS-predicted $\phi(M|M_0)$ exactly at *adjacent* time steps. We now show that this is also a *sufficient* condition for the Monte Carlo method to reproduce $\phi(M|M_0)$ exactly at *any* look-back time regardless of the number, or width, of intervening timesteps. This condition is important because it simplifies the analysis of Monte Carlo algorithms: the failure of a given algorithm to reproduce faithfully the EPS $\phi(M|M_0)$ at a particular redshift or mass range necessarily implies that the algorithm fails to reproduce the progenitor mass function (in either amplitude or shape or both) across a single time step.

We start with the first crossing distribution eq. (6.1) and note that due to the continuous nature of the random walk, it obeys the following identity at different look-back times:

$$\begin{aligned} & f(S(M), z|S(M_0), z_0) \\ = & \int_{S(M_0)}^{S(M)} dS' f(S(M), z|S(M'), z')\, f(S(M'), z'|S(M_0), z_0) \end{aligned} \tag{6.6}$$

for any $z_0 < z' < z$. This relationship is true in both spherical and ellipsoidal EPS models because both variants are based on barrier crossings of random walks. Note that in the ellipsoidal model, eq. (6.6) is a property of only the exact first-crossing distribution, which is well represented by eq. (6.4) for small look-back times but not the Taylor-series-like approximation of Sheth and Tormen [2002].

Using eqs. (6.2) and (6.3) to relate f to the progenitor mass function ϕ, we then obtain

$$\begin{aligned} & \phi(M, z|M_0, z_0) \\ = & \int_{M}^{M_0} dM' \phi(M, z|M', z')\, \phi(M', z'|M_0, z_0)\,. \end{aligned} \tag{6.7}$$

Setting $z' = z_0 + \Delta z$ and $z = z_0 + 2\Delta z$, we see that eq. (6.7) implies that if a Monte Carlo method generates progenitors *exactly* according to the progenitor mass function of EPS at each time step Δz, then the Monte Carlo progenitor mass function should agree with the EPS prediction at any look-back time $z - z_0$. We stress that $\phi(M|M_0)$ must be reproduced exactly, that is, in both the overall shape and normalisation of $\phi(M|M_0)$. This consistency condition is both necessary and sufficient.

An additional feature to note is that consistency is possible in the presence of a mass resolution limit $M_{\rm res}$ (discussed further in §6.3.3). Eq. (6.7) shows that $\phi(M, z|M_0, z_0)$ does not depend on masses outside of the range $[M, M_0]$. Thus if a Monte Carlo algorithm reproduces $\phi(M|M_0)$ for all $M > M_{\rm res}$ in single timesteps, it will consistently reproduce $\phi(M|M_0)$ at $M > M_{\rm res}$ for any $z - z_0$.

6.3.2 The Asymmetry of EPS and Binary Mergers

The simplest way to group progenitors into descendants in a Monte Carlo algorithm is through mass conserving binary mergers, *i.e.* , each descendant halo of mass M_0 is composed of two progenitors of mass M_1 and $M_0 - M_1$. This assumption is used in, e.g., Lacey and Cole [1993] and Cole et al. [2000]. This simple scenario, however, will necessarily fail to reproduce both the spherical and ellipsoidal EPS progenitor mass functions. This is because if all descendants were the products of binary mergers, then $\phi(M|M_0)$ would be symmetric about $M_0/2$. This is simply not the case in EPS even for infinitesimal Δz.

We illustrate the asymmetry of the EPS $\phi(M|M_0)$ in Fig. 6.1 for a descendant halo of mass $10^{13}M_\odot$ at $z_0 = 0$ and a look-back time of $z = 0.02$ (which is the typical timestep used in our Monte Carlo simulations; see Sec. 4). The solid black curve shows the total $\phi(M|M_0)$, while the red dashed curve shows the symmetric part $\phi_{\rm sym}(M|M_0)$ defined by

$$\phi(M|M_0) = \phi_{\rm sym}(M|M_0) + \phi_{\rm asym}(M|M_0)\,, \tag{6.8}$$

where the left side ($M \leq M_0/2$) of $\phi_{\rm sym}(M|M_0)$ is defined to be identical to $\phi(M|M_0)$ and the right side is defined to be simply the reflection of the left half about the mid point $M_0/2$. The second term $\phi_{\rm asym}(M|M_0)$ is then the residual of ϕ after subtracting out $\phi_{\rm sym}$. The figure illustrates that it is not possible for all progenitors with $M > M_0/2$ to have binary-paired progenitors of mass $M_0 - M < M_0/2$. In particular, we find that for sufficiently small look-back times (e.g. $z = 0.02$ used in Fig. 6.1), $\phi(M|M_0) > \phi_{\rm sym}(M|M_0)$ when $M_0/2 \leq M \lessapprox 0.97M_0$ and $\phi(M|M_0) < \phi_{\rm sym}(M|M_0)$ when $M \gtrapprox 0.97M_0$ (see the popup in Fig. 6.1). That is, there are slightly fewer progenitors with masses below $M_0/2$ than above, except near the end points (below $0.03M_0$ and above $0.97M_0$) where the trend is flipped.

Even though the asymmetry is typically small ($\phi_{\rm asym} \lesssim 0.1\,\phi_{\rm sym}$ out to $M_0 - M_{\rm res}$), an accurate algorithm must include non-binary progenitor events. These can be descendants with either a single progenitor or multiple ($N_p > 2$) progenitors, as will be seen in the new algorithms discussed in §6.5 below. This fact was emphasised by Neistein and Dekel [2008a]. These authors construct a mass conserving consistent Monte Carlo algorithm that produces a large number of non-binary descendants. However, one intuitively expects that more mergers will be binary as $z_1 - z_0 \to 0$. This intuition is supported by results from the Millennium simulation [Fakhouri and Ma, 2008], which show that the binary assumption becomes increasingly valid down to smaller $M_{\rm res}$ as $z_1 - z_0$ is made smaller. This result suggests that the Markovian nature of the standard EPS model with a tophat smoothing window may need to be modified to account for the correlated sequences of mergers occurring in simulations [Neistein and Dekel, 2008a, Zentner, 2007].

6.3.3 Mass Resolution, Diffuse Accretion, and Mass Conservation in Monte Carlo Algorithms

In the EPS model, all the mass in the universe is assumed to be in dark matter halos. Although the mass-integral of the (unconditional) mass function in this model is finite, the

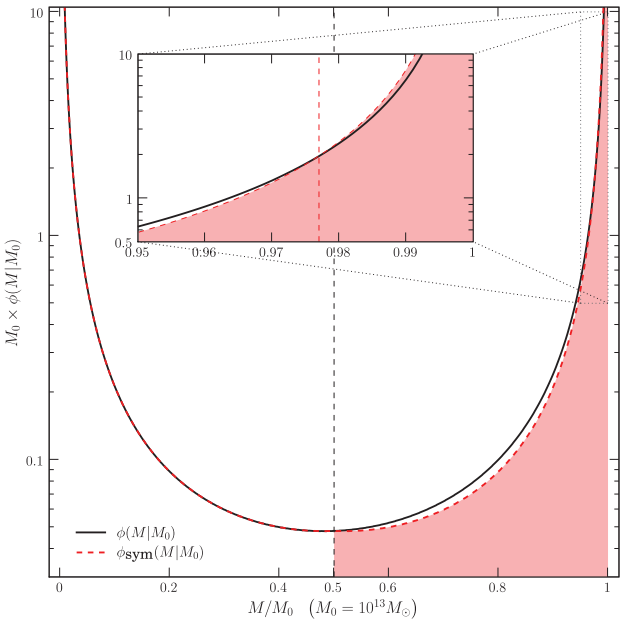

Figure 6.1: An illustration of the asymmetry in the number-weighted conditional (or progenitor) mass function $\phi(M, z|M_0, z_0)$ of the spherical EPS model for a descendant halo of mass $M_0 = 10^{13} M_\odot$ at $z_0 = 0$ and a look-back redshift of $z - z_0 = 0.02$. The red dashed curve shows the symmetric part of $\phi(M|M_0)$, $\phi_{\rm sym}(M|M_0)$, whose right side is simply the reflection of the left side. The figure indicates that some progenitors of masses larger than $M_0/2$ do not have companions in the simplest binary scheme. The popup is a zoom-in on the right-most part of the plot and illustrates that the red dashed curve exceeds the solid curve at $M \approx 0.977 M_0$.

number-integral is unbounded; that is, EPS predicts a preponderance of very low mass halos. Thus, any practical Monte Carlo algorithm must necessarily assume a lower mass cutoff, the mass resolution $M_{\rm res}$.

For a nonzero $M_{\rm res}$, a halo's merger history at each time step can be thought of consisting of mass in the form of resolvable progenitor halos and a reservoir of mass due to "diffuse" accretion that is the aggregate contribution from all subresolution progenitors. This technical distinction is introduced for ease of implementing the Monte Carlo methods. It will, however, play a more physical role when we compare the results with N-body simulations, which has its own mass resolution as well as a possibly physical diffuse component consisting of tidally stripped dark matter particles. In this paper we use ΔM to denote this diffuse accretion component, which we define to be

$$\Delta M = M_0 - \sum_i M_i \,, \tag{6.9}$$

where M_i are the masses of the progenitors above $M_{\rm res}$ and M_0 is the mass of the descendant.

We call a Monte Carlo algorithm *mass conserving* if each descendant and its progenitors produced by the algorithm satisfies $\sum_i M_i \leq M_0$. Monte Carlo algorithms are generally expected to be mass conserving, but we note that this is not a necessary condition for reproducing the EPS progenitor mass function because the latter is a statistical measure of merger properties. In two of our new algorithms below (methods A and B in §6.5), a small fraction of the descendants can have $\sum_i M_i > M_0$. We allow this to simplify the description and implementation of our algorithms. We have experimented with redistributing these excess progenitors among other descendant halos in a mass-conserving manner and found it not to modify significantly the resulting merger statistics. In addition, it may appear that $\sum_i M_i > M_0$ is unphysical. We have found, however, that a non-negligible fraction of halos in N-body simulations in fact have $\Delta M < 0$, perhaps as a result of tidal stripping. This point will be discussed in greater detail in our next paper.

We note that for a Monte Carlo algorithm that is consistent with EPS, the mean value of ΔM per descendant halo of mass M_0 (i.e., averaged over all descendants in a given timestep) is, by construction, related to the mass resolution by

$$\langle \Delta M \rangle = \int_0^{M_{\rm res}} M \phi(M|M_0) dM \,, \tag{6.10}$$

For a given $\phi(M|M_0)$ and $M_{\rm res}$, $\langle \Delta M \rangle$ is therefore specified. The *distribution* of ΔM, however, can differ greatly among different algorithms; that is, there is much freedom in how to assign the amount of diffuse accretion to individual descendants in a given timestep. For instance, Cole et al. [2000] assumes a delta-function distribution for ΔM (see §6.4.4 for details), while most of other methods, including our new methods discussed in §6.5, have broader distributions.

6.4 Comparison of Four Previous Monte Carlo Algorithms

In this section we examine four existing Monte Carlo algorithms for generating merger trees: Lacey and Cole [1993] [LC93], Kauffmann and White [1993] [KW93], Somerville and Kolatt [1999] [SK99], and Cole et al. [2000] [C00]. This set is by no means complete, but these are four of the most frequently used algorithms in the literature. The purpose here is to compare these well known algorithms side-by-side and to illustrate the mass and redshift ranges for which each method succeeds and fails in matching the spherical EPS model. Comparison of KW93 with the ellipsoidal model will be presented in Sec. 5.

We review each algorithm in a subsection below and compare the resulting progenitor mass functions $\phi(M|M_0)$ with the spherical EPS prediction for look-back redshifts ranging from 0.24 to 15. In Figs. 6.2-6.4 we plot the progenitor mass functions produced by all four methods, along with the analytical EPS prediction, on log-log plots for three descendant masses ($10^{12}, 10^{13}, 10^{14} M_\odot$) and four look-back times ($z_1 - z_0 = 0.24, 2.07, 7$ and 15). To ease comparison, we also plot the ratio between each Monte Carlo result and the EPS prediction on a linear-log plot. As Figs. 6.2-6.4 clearly show, of the four algorithms, only KW93 is able to match the spherical EPS $\phi(M|M_0)$ for *all* $z - z_0$. We will explore why each algorithm fails below and discuss the care that must be taken when implementing KW93. A summary of the four algorithms, their discrepancies, and the causes of the discrepancies is given in Table 6.1.

In our Monte Carlo simulations, we generally keep track of all progenitors down to $0.001 M_0$ at each time step for a descendant halo of mass M_0. This large dynamic range allows us to predict reliably the progenitor abundance even for a very large look-back time ($z_1 - z_0 \sim 15$). To speed up the algorithm, we take each time step to be a constant difference in the barrier height $\Delta\omega(z) = \omega(z + \Delta z) - \omega(z)$ (where $\Delta\omega \approx \Delta z$ at low z), which is chosen to be about 0.02 for LC93, KW93, SK99, and 0.003 for C00 at $z = 0$. The progenitor mass function of a given descendant halo mass is then identical for each time step and does not have to be recomputed. Numerical convergence is tested by changing the timesteps used in the simulation: our results do not change.

6.4.1 Lacey & Cole (1993)

The algorithm proposed by LC93 makes two important assumptions: all mergers are binary (before mass resolution is imposed), and the descendant mass M_0 is the sum of the two progenitor masses M_1 and M_2 (where $M_1 \geq M_2$ in our convention). For each small look-back time step and for each descendant, a progenitor mass is randomly chosen according to the mass-weighted conditional mass function eq. (6.2), and the mass of the other progenitor (which can be larger or smaller) is simply set to be the difference between M_0 and the first chosen progenitor mass. If the less massive progenitor M_2 falls below a chosen mass resolution $M_{\rm res}$, or equivalently, $M_1 > M_0 - M_{\rm res}$, then M_1 is kept but M_2, being a sub-

	Overview	Discrepancy in progenitor mass function $\phi(M\|M_0)$	Reasons for Discrepancies
LC93	Binary and 1-to-1 $\Delta M \leq M_{\rm res}$	Overestimates $\phi(M\|M_0)$ by large factors when the look-back time is large, *i.e.* , $z_1 - z_0 \gtrsim 1$	Binary assumption fails to reproduce EPS $\phi(M\|M_0)$ asymmetry.
KW93	Multiple mergers $\Delta M \neq 0$	None	
SK99	Multiple mergers $\Delta M \neq 0$ (can be bigger or smaller than $M_{\rm res}$)	Typically over-predicts the abundances of small progenitors ($\lesssim$ 10% of the descendant halo mass) by a factor of ~ 2 for $z_1 - z_0 \lesssim 1$. This discrepancy propagates to smaller mass scales for larger look-back times.	Truncation of $\phi(M\|M_0)$ fails to reproduce its shape exactly.
C00	Binary and 1-to-1 ΔM is a constant given by equation (6.12)	Works reasonably well for a large range of the look-back time but significantly underestimates $\phi(M\|M_0)$ at high mass ends, particularly when the look back time is large ($z_1 - z_0 \gg 1$).	Binary assumption fails to capture asymmetry of EPS $\phi(M\|M_0)$; fixed ΔM yields 1-to-1 events that do not accurately reproduce the high mass end of $\phi(M\|M_0)$.

Table 6.1: A scorecard for the four old Monte Carlo algorithms discussed in §6.4. We note that the 1-to-1 events in LC93 and C00 are actually binary mergers involving a secondary progenitor with mass below $M_{\rm res}$. Since these progenitors are below the resolution limit they are not counted as progenitors but as diffuse mass ΔM.

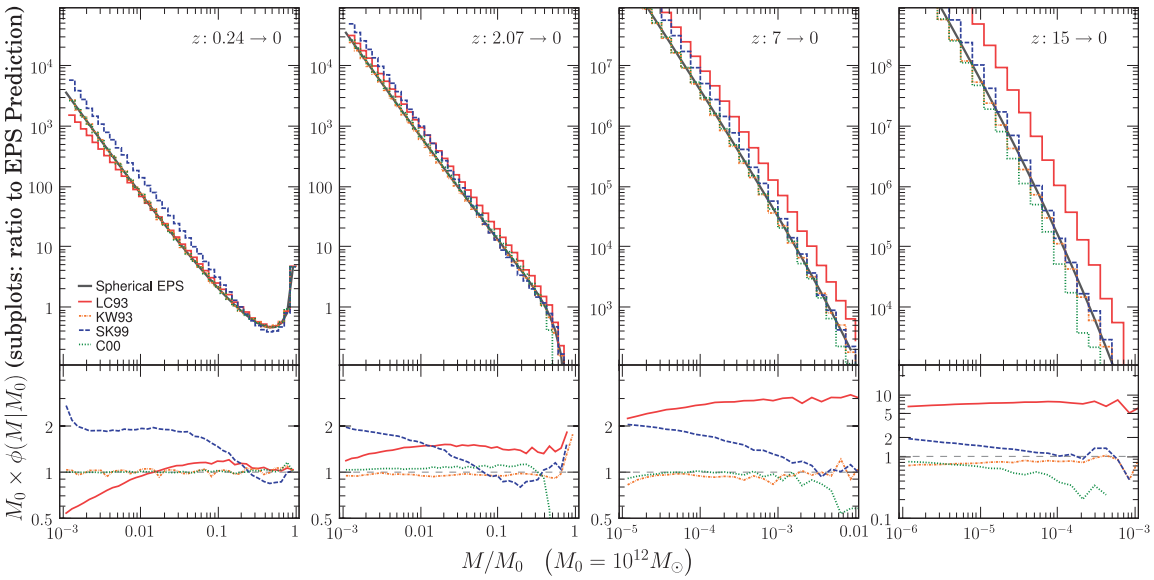

Figure 6.2: Comparison of the progenitor (or conditional) mass functions $\phi(M, z|M_0, z_0)$ that we generated using the four previous Monte Carlo algorithms by LC93 (red solid), KW93 (orange dot-dashed), SK99 (blue dashed), and C00 (green dotted), and the predictions of the analytic spherical EPS model (black solid). The four panels show four look-back redshifts ($z - z_0 = 0.24, 2.07, 7$ and 15) for a descendant halo of $M_0 = 10^{12} M_\odot$ at $z_0 = 0$. For clarity, we plot in the sub-panel below each panel the ratios of the Monte Carlo result and the EPS prediction. One can see that KW93 is the only accurate algorithm for all z. Note that different ranges of M/M_0 are shown in each panel since the progenitors have progressively smaller masses at higher redshifts.

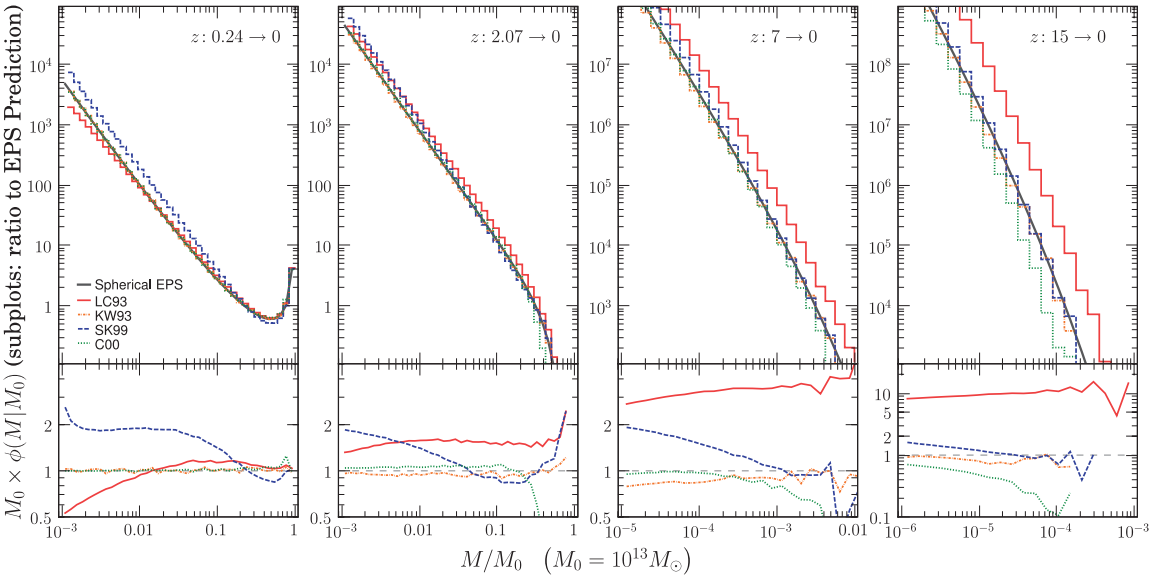

Figure 6.3: Same as Fig. 6.2, but for a descendant halo of $10^{13}\mathrm{M}_\odot$.

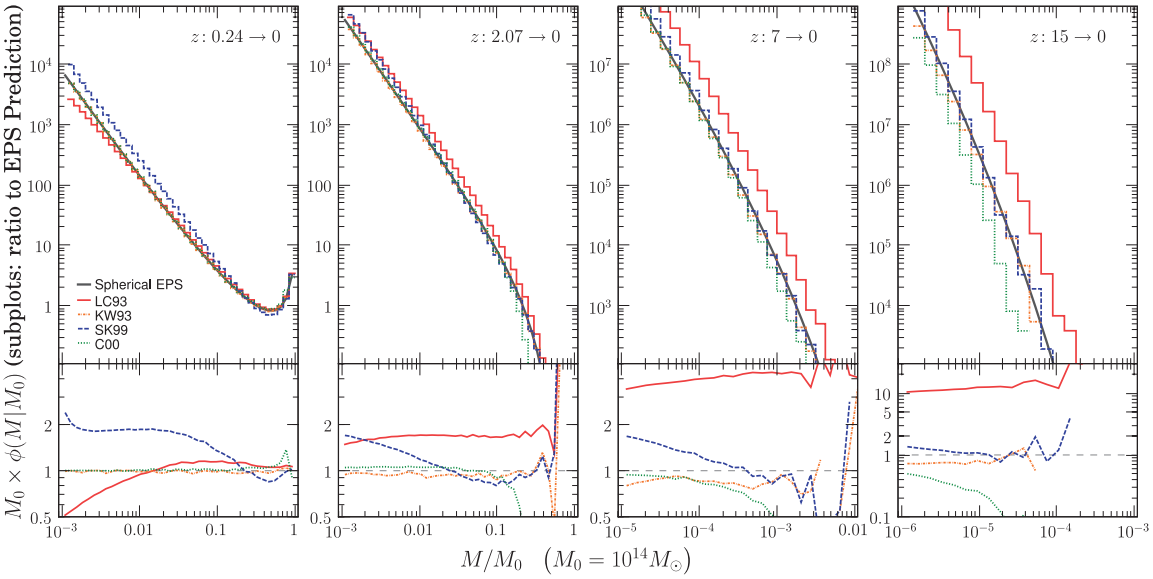

Figure 6.4: Same as Fig. 6.2, but for a descendant halo of $10^{14}\mathrm{M}_\odot$.

resolution progenitor, is discarded. This results in single-progenitor halos which we label as "$1 \rightarrow 1$" events. In this notation, binary mergers are "$2 \rightarrow 1$" events. When a smaller timestep is used in LC93, the ratio of $2 \rightarrow 1$ to $1 \rightarrow 1$ events decreases.

We find that random progenitor masses can be easily generated using the parameter transformation:

$$x = \mathrm{erf}\left\{\Delta\omega / \sqrt{2[S(M_1) - S(M_0)]}\right\} . \tag{6.11}$$

The parameter x has a uniform probability distribution between 0 and 1 and can be quickly generated using any random-number generator. A simple inversion then yields progenitors distributed according to the mass-weighted conditional mass function.

The red solid histograms and curves in Fig. 6.2 – 6.4 compare the progenitor mass functions generated using the LC93 algorithm with the predictions of the spherical EPS model (solid black curves). For all three descendant halo masses shown ($10^{12}, 10^{13}$ and $10^{14} M_{\odot}$), we see close agreement for small look-back times such as $z_1 - z_0 = 0.24$, but LC93 produces an excess of progenitors at larger look-back times, and the discrepancy worsens, reaching an order of magnitude by $z_1 - z_0 = 15$. We believe this discrepancy is due to the binary nature of LC93: the number of progenitors with mass M is equal to the number of binary companions of mass $M_0 - M$. Thus the LC93 Monte Carlo algorithm generates a progenitor mass function after one time step that is symmetric in the left and right sides, which will not match the asymmetric nature of the EPS $\phi(M|M_0)$ discussed in Sec. 6.3.2 and shown in Fig. 6.1. This discrepancy is amplified after many timesteps when the look-back time becomes large.

Finally, we note that the authors of LC93 also consider another way of drawing the first progenitor mass from the mass-weighted conditional mass function, which is to draw it from the mass range of $[M_0/2, M_0]$ instead of $[0, M_0]$. In practice, we find that this slightly modified version of LC93 generates very similar results, and our above discussion is valid.

6.4.2 Kauffmann & White (1993)

For each timestep in the KW93 algorithm, a large number of progenitors are generated across many progenitor mass bins for a fixed number of descendant halos of the same mass. The number of progenitors in each mass bin is determined by the progenitor mass function of the descendant halo mass, and rounded to the nearest integer value. These progenitors are then assigned to the descendant halos in order of decreasing progenitor mass. The target descendant halo is chosen with a probability proportional to its available mass (i.e. the mass not yet occupied by progenitors), and with the restriction that the total mass of the progenitors in a descendant halo cannot exceed the descendant mass. This procedure allows one to work out all the merger configurations and their frequencies for one time step and for different descendant halo masses. This information is then stored and used repeatedly for determining the progenitors of a halo at each time step.

To speed up the implementation of KW93, we divide the look-back time into steps with equal spacing in the barrier height $\Delta\omega$ as discussed earlier. The progenitor mass function

for a fixed descendant halo mass is then identical for every time step and only has to be calculated once. We store the ensemble of progenitors and their merger configurations for each descendant halo mass bin. In a Monte Carlo simulation, we randomly select one merger configuration from the many stored ones for a descendant halo at each time step.

In practice, we find that extreme care must be taken to avoid numerical problems in KW93. First of all, this algorithm requires a large number of progenitor mass bins in the neighbourhood of M_0 because $\phi(M|M_0)$ is sharply peaked near $M_1 \sim M_0$ for small timesteps. Interestingly, we find that if the mass range of $[M_{\rm res}, M_0]$ is simply divided into evenly-spaced logarithmic bins, this method is not accurate even when the number of mass bins is as large as 2000, which already requires more than ~ 50000 descendant halos to guarantee that the integer rounding does not introduce a significant error to the progenitor number in each bin. As a result, a large amount of computer memory is necessary to repeat this procedure for descendant halos of different masses. The improved mass bin configuration that we end up using will be introduced in §6.5. Using that setup, we find that only 200 bins are required to reproduce accurately the EPS progenitor mass function over large look-back times.

The second problem is that KW93's scheme for assigning progenitors to descendant halos is somewhat ambiguous and does not guarantee that all the progenitors can be assigned. Fortunately, we find that this problem usually does not arise when the ensemble of progenitors is large. For each descendant halo mass, we use ~ 8000 descendant halos to determine the merger configurations of the progenitors.

The orange dash-dotted curves in Fig. 6.2 - 6.4 compare the progenitor mass functions generated using the KW93 algorithm with the predictions of the spherical EPS model (black). The results show very good agreement. Since KW93 reproduces the exact EPS progenitor mass function at every timestep, it is expected to be consistent with EPS at any $z_1 - z_0$ according to the discussion in §6.3.1.

6.4.3 Somerville & Kolatt (1999)

Somerville & Kolatt (1999) [SK99] point out that the assumptions of binary mergers and $M_0 = M_1 + M_2$ made in LC93 lead to an overestimate of the progenitor abundance at high redshift. They first attempt to remedy this problem by preserving the binary assumption while allowing the mass below the resolution limit $M_{\rm res}$ to be counted as diffusely accreted mass ΔM (see §6.3.3). They show, however, that this "binary tree with accretion" method fails in the opposite direction, *underproducing* the progenitor mass function relative to the spherical EPS prediction. This discrepancy arises partly because whenever two progenitors are chosen in this method, the remaining mass is assigned to ΔM regardless of whether it is above or below $M_{\rm res}$. Thus the EPS $\phi(M|M_0)$ is not faithfully reproduced: the binary tree with accretion method yields an excess of accreted mass and a corresponding shortage of low-mass halos.

SK99 then consider a natural extension of this method, in which both assumptions made in LC93 are relaxed. In this "N-branch tree with accretion" algorithm, each descendant halo is allowed to have more than two progenitors for every simulation timestep. To guar-

antee that the total mass of the progenitors does not exceed that of the descendant, each subsequent progenitor mass is randomly chosen from the mass-weighted conditional mass function truncated to the maximally possible progenitor mass. This procedure is repeated until the descendant halo cannot contain any more progenitors with masses above $M_{\rm res}$, and the remaining mass deficit is assigned to diffuse accretion ΔM.

The parameter transformation of eq. (6.11) is also applicable for SK99. The probability distribution of x is still uniform, but the upper limit of x can now take on any value between 0 and 1 depending on where the conditional mass function is truncated.

The blue dashed curves in Fig. 6.2 - 6.4 compare the progenitor mass functions generated using the N-branch tree algorithm of SK99 with the predictions of the spherical EPS model (black). It is interesting to note that the sign of the discrepancy is now opposite to that of LC93: SK99 produces an excess of low-mass ($\lesssim 0.1M_0$) progenitors by up to a factor of ~ 2 for small look-back times, but it does a better job than LC93 at high redshifts. However, it is noteworthy that even at high redshifts, discrepancies of up to a factor of ~ 2 are still present for small progenitor masses.

We believe that the use of a truncated progenitor mass function in SK99 is at least a partial cause for the over-prediction of small progenitors. Since the distribution of progenitors (in particular, the upper limit for the progenitor mass) depends on the sum of the masses of the progenitors already picked out for the current halo, the *order* in which progenitor halos are randomly pulled out matters in this method. Halos more massive than the truncation limit are effectively discarded instead of being randomly selected and placed in, for example, new descendant halos. This procedure tends to preferentially skew the progenitor mass function at small time steps towards more low mass progenitors and fewer high mass progenitors.

6.4.4 Cole et al. (2000)

Similar to SK99, Cole et al. [2000] [C00] treats the mass in progenitors smaller than the mass resolution $M_{\rm res}$ in the Monte Carlo simulation as accreted mass, but unlike the N-branch tree model in SK99, only a maximum of two progenitors are allowed per descendant. The amount of accreted mass gained in one timestep, ΔM, is fixed to a single value and is calculated by integrating the mass-weighted conditional mass function from 0 to $M_{\rm res}$:

$$\Delta M = \int_0^{M_{\rm res}} M\phi(M|M_0)dM\,, \tag{6.12}$$

where M_0 is the descendant mass. The progenitors are drawn from the *lower* half of the progenitor mass function between $M_{\rm res}$ and $M_0/2$ according to the average number of progenitors in that range:

$$p = \int_{M_{\rm res}}^{M_0/2} \phi(M|M_0)dM\,. \tag{6.13}$$

The simulation timestep is chosen to be small enough so that $p \ll 1$ (note that it is for this reason that we use $\Delta z = 0.003$ when implementing C00).

The C00 merger tree is generated with the following steps: A random number x between 0 and 1 determines whether a descendant halo has one progenitor (if $x > p$) or two progenitors (if $x \leq p$). In the case of a single progenitor, its mass is $M_1 = M_0 - \Delta M$. In the case of two progenitors, the mass of the smaller progenitor, M_2, is chosen randomly between $M_{\rm res}$ and $M_0/2$ according to the progenitor mass function. The larger progenitor is then assigned a mass of $M_1 = M_0 - M_2 - \Delta M$. Since $p \ll 1$, most descendants form via $1 \to 1$ events rather than $2 \to 1$ events. To improve the speed of this algorithm, we precompute and store the binary merger rates and diffuse accretion mass fractions for a single time step for different descendant mass bins.

The green dotted curves in Fig. 6.2 - 6.4 compare the progenitor mass functions generated using the C00 algorithm with the predictions of the spherical EPS model (black). The agreement is noticeably better than LC93 and SK99. The largest discrepancy occurs at the high mass end at large $z_1 - z_0$, where C00 *underpredicts* the progenitor number at z_1 by more than a factor of two for group-to-cluster size descendants at z_0 with $M_0 \gtrsim 10^{13} M_\odot$.

At least two problems contribute to this discrepancy: (i) Since ΔM is fixed to one value (eq. 6.12), the mass of the progenitor for $1 \to 1$ descendants is also a fixed value: $M_1 = M_0 - \Delta M$. This is an over-simplification that compresses the high mass end of $\phi(M|M_0)$ into a delta function. (ii) For descendants with binary progenitors, C00 uses the spherical EPS conditional mass function only in the lower mass range $[0, M_0/2]$ to generate the progenitor abundance. By construction, then, the shape of the progenitor mass function in the upper mass range, $[M_0/2, M_0]$, is symmetric with the lower half and fails to match accurately the asymmetric EPS $\phi(M|M_0)$.

6.5 Three Consistent Monte Carlo Algorithms

In this section, we present three Monte Carlo algorithms that all satisfy the criterion for consistency discussed in §6.3.1 and will therefore accurately reproduce the EPS progenitor mass function $\phi(M|M_0)$. We introduce the common setup for our methods in §6.5.1 and discuss in detail how each method assigns the ensemble of progenitors to descendants in §6.5.2 – 6.5.4.

To help the reader follow our discussions, we provide a summary of the breakdown of the merger configurations for the three new algorithms in Table 6.3 and the accompanying Fig. 6.6.

Although the standard practice in the community has been to generate merger trees using the spherical EPS model, we emphasise that the Monte Carlo algorithms can be applied to the ellipsoidal EPS model as well. In fact, since the ellipsoidal model matches the unconditional mass function in simulations better than the spherical model, we would expect the ellipsoidal EPS to also match better the progenitor statistics in simulations. We will therefore present our results for both the spherical and ellipsoidal EPS models in parallel below.

6.5.1 The Common Setup

Basic Features

Our Monte Carlo algorithms for growing consistent merger trees all share the following implementation framework. We begin at redshift 0 and build the merger tree backwards in cosmic time. We typically choose a large descendant halo mass range ($M_0 = [10^6 M_\odot, 10^{15} M_\odot]$) and a small simulation timestep ($\Delta z \approx 0.02$ at low z; see discussion below) to achieve a high resolution tree and a large dynamic range in the progenitor mass. For a given descendant halo, we first compute which mass bin it belongs to, and then obtain its progenitors across a single timestep using the distribution of merger configurations specific to each algorithm (described in the next three subsections). The progenitors then become descendants in the next timestep, and this process is repeated to build up the higher tree branches.

To be specific, a merger configuration here is defined as a set of progenitor masses that form a descendant halo of a given mass in one timestep. For example, for a descendant halo of mass M_0, one merger configurations may include only two progenitors of mass $0.6M_0$ and $0.4M_0$, while another may contain three progenitors of mass $0.4M_0$, $0.3M_0$, and $0.2M_0$. Note that the sum of the progenitor mass in each configuration need not equal the descendant mass, and the deficit, ΔM, is implicitly attributed to sub-resolution progenitors (see §6.3.3). Different Monte Carlo algorithms have different distributions of merger configurations and progenitor multiplicities for each descendant bin. For convenience, we call the most massive progenitor in a merger configuration the *primary* progenitor, and the rest of the progenitors the *secondary* progenitors.

Our basic implementation is applicable to both the spherical and ellipsoidal EPS models. We find a particularly efficient choice of timestep to be the one corresponding to a constant difference in the barrier height $\Delta\omega(z) = \omega(z + \Delta z) - \omega(z)$, as is used in §6.4 for the four old algorithms. For the spherical case, the progenitor mass function eq. (6.3) depends on time only through $\Delta\omega(z)$ and is therefore identical for all redshifts when the same $\Delta\omega(z)$ is used. Thus we only have to generate the merger configurations in the spherical case across a single timestep once. For the ellipsoidal case, however, the progenitor mass function eq. (6.4) not only is a function of $\Delta\omega(z)$ but also depends explicitly on z. For each Monte Carlo algorithm, it is therefore necessary to generate and store the merger configurations and their probabilities for both descendant halos of different masses and several redshift bins. In practice, since the redshift dependence of eq. (6.4) is weak, typically fewer than ~ 20 redshift bins are required.

Important Progenitor Mass Scales

A number of natural mass boundaries play critical roles in the construction of our algorithms. These mass scales demarcate the regions with different progenitor multiplicities, as illustrated in Fig. 6.6 and discussed in detail in the next three subsections.

(i) The resolution scale $M_{\rm res}$ and its complement $M_0 - M_{\rm res}$ are two obvious boundaries, as is the half descendant mass $M_0/2$ discussed in the context of binary mergers in Sec. 6.3.2.

We generally choose a small $M_{\rm res}$ (typically $M_{\rm res} = 0.001M_0$) for numerical precision and keep track of all the progenitors down to this limit at each timestep.

(ii) The mass αM_0 given by

$$\int_{\alpha M_0}^{M_0} \phi(M|M_0)\, dM = 1 \tag{6.14}$$

defines the range of progenitor mass over which every descendant halo is guaranteed to have one progenitor with $M \in [\alpha M_0, M_0]$. Table 6.2 lists the values of α for both the spherical and ellipsoidal progenitor mass functions for three descendant masses; α is seen to range from 0.361 to 0.448.

(iii) The mass μM_0 demarcates where the asymmetric progenitor mass function self-intersects: $\phi(\mu M_0|M_0) = \phi(M_0 - \mu M_0|M_0)$ with $\mu > 0.5$. For binary merger configurations of the form $M_0 = M_1 + M_2$, $\phi(M_1|M_0) > \phi(M_2|M_0)$ when $M_1 < \mu M_0$ and $\phi(M_2|M_0) > \phi(M_1|M_0)$ when $M_1 > \mu M_0$. This mass scale is illustrated in the pop-up in Fig. 6.1. Table 6.2 shows that $\mu \approx 0.956$ to 0.977.

Fig. 6.5 shows α and μ as functions of the look-back time Δz for three descendant halo masses ($10^{12}M_\odot$, $10^{13}M_\odot$, $10^{14}M_\odot$) at redshift zero. According to the figure, α and μ have well defined constant values when Δz is less than about 0.05, a natural upper limit of time stepsize for a Monte Carlo simulation to achieve convergence in both the spherical and ellipsoidal EPS models.

Mass Bins

To help the reader reproduce our Monte Carlo algorithms, we discuss our distribution of mass bins.

We divide the descendant mass range $10^6 \leq M_0 \leq 10^{15}M_\odot$ into ~ 100 logarithmic descendant bins. Halos that fall into the same descendant bin are assumed to have the same distribution of single-timestep merger configurations that are computed using the central (in logarithmic scale) value of the bin as the descendant mass. The progenitor masses in a merger configuration are recorded in the form of ratios to the descendant halo mass, instead of their absolute masses. This allows us to correct for the (small) difference between the descendant halo in question and the central mass of its bin.

For a given descendant mass M_0, its progenitor mass range $[M_{\rm res}, M_0]$ is divided into a certain number of mass bins to facilitate the process of forming merger configurations. Interestingly, we note that simply dividing the whole progenitor mass range into evenly spaced logarithmic bins is not accurate, as discussed in §6.4.2. This is because the simplest logarithmic binning assigns very few bins to the mass range of $[M_0/2, M_0]$, which requires many mass bins to sample accurately the shape of the sharply peaked (at around M_0) progenitor mass function for a small timestep. To give the peaked region more fine structures, we find a simple way: the mass range of $[M_{\rm res}, M_0/2]$ is divided into evenly spaced logarithmic bins, and its reflection about the mid point $M_0/2$ determines the binning on the right side of the mid point. Mathematically, it can be stated as follows: The progenitor mass range

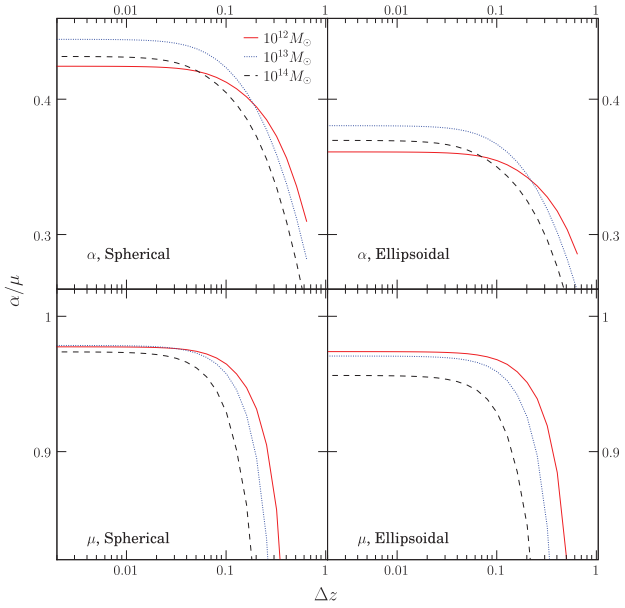

Figure 6.5: α and μ as functions of the look-back time Δz at redshift zero. The red solid, blue dotted, and black dashed curves are for descendant halos of $10^{12}M_\odot$, $10^{13}M_\odot$, and $10^{14}M_\odot$ respectively. The label in each plot indicates the quantity (α or μ) shown and the EPS model (spherical or ellipsoidal) used.

$[M_{\rm res}, M_0]$ is divided into $2N+1$ logarithmic mass bins. The i^{th} $(i = 0, 1, 2, ..., 2N)$ bin spans $[M^{i+1}, M^i]$, where M^i is defined as follows:

$$M^i = \begin{cases} M_0 & \text{if } i = 0; \\ \exp\left[\ln M_{\rm res} + \Delta \times (2N + 1 - i)\right] & \text{if } i \geq N + 1; \\ M_0 - M^{2N+2-i} & \text{if } 1 \leq i \leq N. \end{cases}$$

and $\Delta = (\ln(M_0/2) - \ln M_{\rm res})/N$.

The average number of progenitors (per descendant halo) in the i^{th} bin is called N^i, which is equal to $\int_{M^{i+1}}^{M^i} \phi(M|M_0)dM$. Note that N^i is not an integer. For $i \geq 1$, we choose the mean mass $\bar{M}^i$ of the i^{th} bin to be $\sqrt{M^i M^{i+1}}$. The progenitor mass function often changes rapidly across the 0^{th} bin so we do not assign it a mean mass. Instead, whenever a progenitor of the 0^{th} bin is needed, we generate a probabilistic progenitor mass according to the progenitor mass function inside this bin.

6.5.2 Method A

We first attempt to resolve the asymmetry problem in the EPS progenitor mass function $\phi(M|M_0)$ by assuming that the primary progenitors in the symmetric part $\phi_{\rm sym}$ in eq. (6.8) are paired up with secondary progenitors to form binary mergers such that $M_0 = M_1 + M_2$.

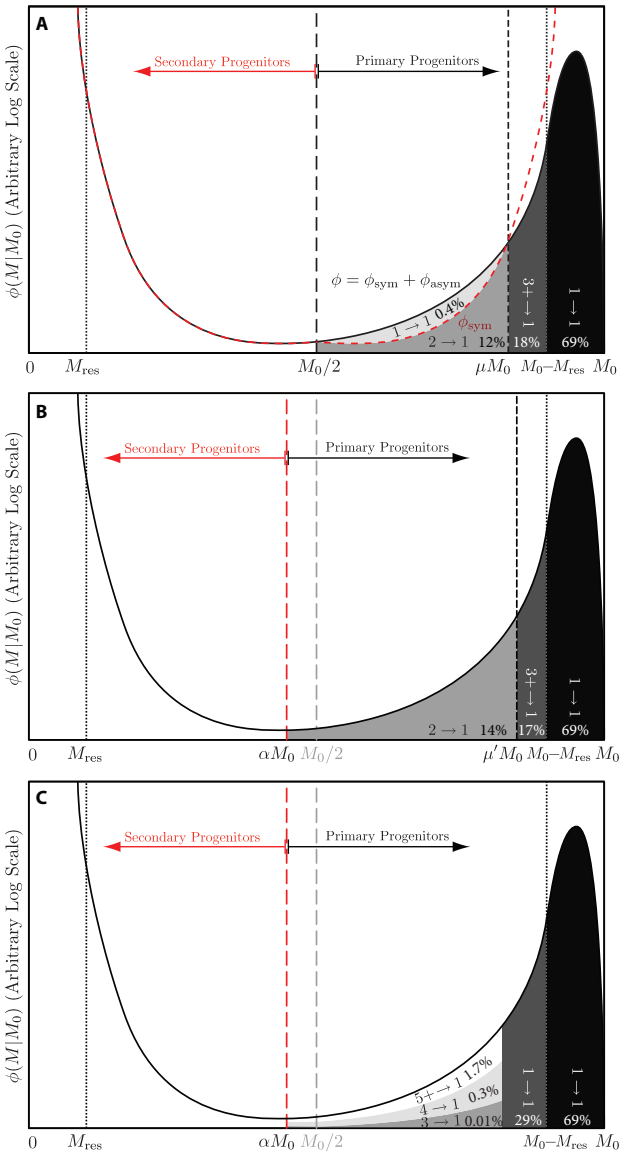

Figure 6.6: A schematic summary of how the three new algorithms proposed in this paper assign progenitors to descendants in a single timestep (see §5). The regions are shaded according to the progenitor multiplicity (marked by $N_p \to 1$) and the mass ranges. See Table 6.3 for a description of each shaded region and the fraction of descendants that belongs to each region. The numbers quoted in this plot are from the ellipsoidal EPS model. The axes are in arbitrary units, though the horizontal axis is drawn to be symmetric about $M_0/2$ and the vertical axis is assumed to be logarithmic. Important characteristic progenitor masses are labelled on the horizontal axis (see §5.1.2 for discussion). The dashed line in panel A plots $\phi_{\rm sym}$, the reflection of the left side of $\phi(M|M_0)$.

	Spherical EPS			Ellipsoidal EPS (z=0)		
M_0 ($M_\odot$)	10^{12}	10^{13}	10^{14}	10^{12}	10^{13}	10^{14}
α	0.421	0.448	0.435	0.361	0.384	0.372
μ	0.977	0.977	0.970	0.974	0.970	0.956

Table 6.2: Values of the progenitor mass scales α and μ discussed in §6.5.1 for the spherical and ellipsoidal EPS models for three descendant masses (10^{12}, 10^{13}, and $10^{14} M_\odot$) and $\Delta z = 0.02$, where αM_0 is defined such that $\int_{\alpha M_0}^{M_0} \phi(M|M_0)\, dM = 1$ and μM_0 is defined such that $\phi(\mu M_0|M_0) = \phi(M_0 - \mu M_0|M_0)$ with $\mu \neq 0.5$.

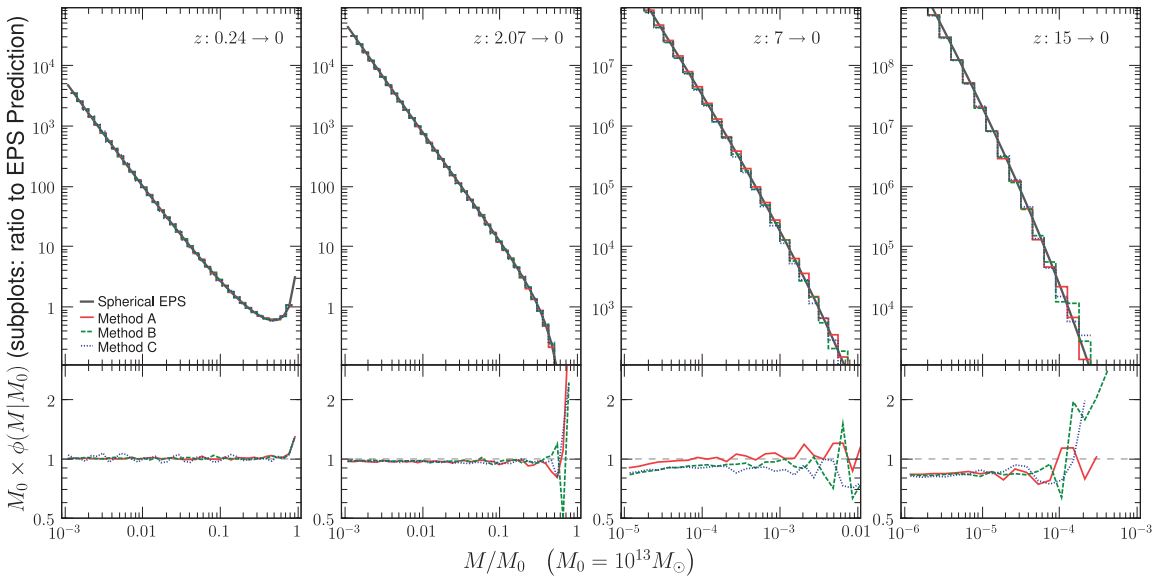

Figure 6.7: Comparison of the progenitor (or conditional) mass functions $\phi(M, z|M_0, z_0)$ generated using the three new Monte Carlo algorithms introduced in §6.5: A (red solid), B (green dashed), and C (blue dotted), and the predictions of the spherical EPS model (black solid). The four panels show four look-back redshifts ($z - z_0 = 0.24, 2.07, 7, 15$) for a descendant halo of mass $10^{13} M_\odot$ at $z_0 = 0$. For clarity, we plot in the sub-panel below each panel the ratios of the Monte Carlo result and the EPS prediction. All three algorithms are seen to match very closely the spherical EPS prediction at all redshifts.

Method	N_p	% Desc. (spher.)	% Desc. (ellip.)	Description	Key
A	$0 \to 1$	0.3%	0.4%	Descendants with no progenitors because $\int_{M_0/2}^{M_0} \phi \, dM < 1$	N/A
	$1 \to 1$	60%	69%	$M_1 \in [M_0 - M_{\rm res}, M_0]$: binary-turned-singles due to $M_2 < M_{\rm res}$	A ■
	$1 \to 1$	0.8%	0.4%	$M_1 \in [M_0/2, \mu M_0]$: $\phi_{\rm sym} < \phi$ results in unpaired primary progenitors: $\Delta M > M_{\rm res}$	A □
	$2 \to 1$	21%	12%	$M_1 \in [M_0/2, \mu M_0]$ and $M_0 = M_1 + M_2$: binary pairs generated from $\phi_{\rm sym}$	A ■
	$3+ \to 1$	18%	18%	$M_1 \in [\mu M_0, M_0 - M_{\rm res}]$: $\phi_{\rm sym} > \phi$ results in excess secondary progenitors. $M_0 < M_1 + M_2 + M_3 + ...$	A ■
B	$1 \to 1$	60%	69%	$M_1 \in [M_0 - M_{\rm res}, M_0]$: binary-turned-singles due to $M_2 < M_{\rm res}$	B ■
	$2 \to 1$	20%	14%	Binary paired progenitors generated by the iterative algorithm of §6.5.3: $M_0 \geq M_1 + M_2$	B ■
	$3+ \to 1$	20%	17%	$M_1 \in [\mu M_0, M_0 - M_{\rm res}]$: identical to $3+ \to 1$ configuration in method A	B ■
C	$1 \to 1$	60%	69%	$M_1 \in [M_0 - M_{\rm res}, M_0]$: binary-turned-singles due to $M_2 < M_{\rm res}$	C ■
	$1 \to 1$	35%	29%	All secondary progenitors have already been assigned to smaller primary progenitors: these remaining primary progenitors have $\Delta M > M_{\rm res}$	C ■
	$3 \to 1$	0.1%	0.01%	Merger configurations with 3 progenitors	C ■
	$4 \to 1$	2%	0.3%	Merger configurations with 4 progenitors	C □
	$5+ \to 1$	2.9%	1.7%	Merger configurations with 5 or more progenitors	C □
KW93	$1 \to 1$	60%	69%	$M_1 \in [M_0 - M_{\rm res}, M_0]$: binary-turned-singles due to $M_2 < M_{\rm res}$	N/A
	$1 \to 1$	15%	9%	Merger configurations with a single progenitor with $M_1 < M_0 - M_{\rm res}$	N/A
	$2 \to 1$	11%	9%	Merger configurations with 2 progenitors	N/A
	$3+ \to 1$	14%	13%	Merger configurations with 3 or more progenitors	N/A

Table 6.3: A summary of our three new Monte Carlo methods discussed in §6.5 and the method of KW93. The percentages indicate the fractions of descendants with N_p progenitors in a given method, computed for $M_0 = 10^{13} M_\odot$ and $M_{\rm res} = 0.001 M_0$ for a single timestep $\Delta z = 0.02$ in both the spherical and ellipsoidal EPS models. They are representative of the merger configuration distributions for other descendant halo masses M_0.

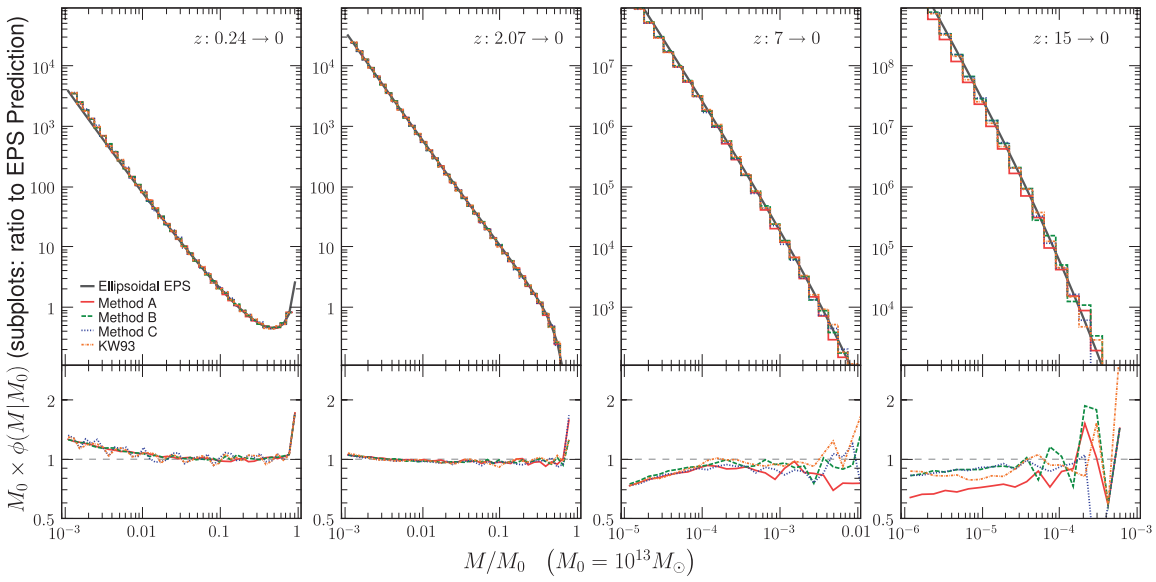

Figure 6.8: Same as Fig. 6.7 except both the Monte Carlo and analytic results are now generated from the *ellipsoidal* instead of the standard spherical EPS model. The Monte Carlo methods use eq. (6.4) as the progenitor mass function for each time step. The analytic results are calculated using the integral equation proposed by Zhang & Hui (2006). The agreement is again excellent, indicating that our new Monte Carlo algorithms work well in reproducing the EPS progenitor mass function regardless if the EPS model is based on constant (i.e. spherical collapse) or moving barrier (ellipsoidal) random walks. For completeness, we include the results from the ellipsoidal version of the KW93 method (orange dash-dotted).

This is done so long as the smaller progenitor is above the mass resolution of the Monte Carlo simulation, i.e. $M_2 \geq M_{\rm res}$ and $M_1 \leq M_0 - M_{\rm res}$. If $M_2 < M_{\rm res}$, then the second progenitor is discarded and M_1 is assumed to be a single progenitor (the darkest grey region marked $1 \to 1$ in Fig. 6.6 A). The remaining primary progenitors in the asymmetric part $\phi_{\rm asym}$ are assumed not to pair up, i.e. each descendant halo has a single progenitor (the lightest grey region marked $1 \to 1$ in Fig. 6.6 A).

In practice, we generate the merger configurations of a descendant halo of mass M_0 at each time step by repeating these two simple steps:

(i) Draw the primary progenitor mass M_1 from the mass range $[M_0/2, M_0]$ of the progenitor mass function.

(ii) If $M_1 > M_0 - M_{\rm res}$, no more progenitors are generated; if $M_1 \leq M_0 - M_{\rm res}$, the probability of having a second progenitor of mass $M_2 = M_0 - M_1$ is set to

$$r = \frac{\phi_{\rm sym}(M_1|M_0)}{\phi(M_1|M_0)} = \frac{\phi(M_0 - M_1|M_0)}{\phi(M_1|M_0)}\,. \tag{6.15}$$

Then, drawing a random number x between 0 and 1 allows us to determine whether a secondary progenitor should be generated. If $x < r$, M_2 is assigned as a secondary progenitor; otherwise M_1 is left as the sole progenitor.

We point out two subtleties with this algorithm. First, r is not always ≤ 1. It is true that r is below 1 for *most* of the relevant mass range $M_1 \in [M_0/2, \mu M_0]$ (see Fig. 6.6 A and Table 6.2) since the left side of $\phi(M|M_0)$ is slightly lower than the right side. But when $M_1 > \mu M_0$, we find that $r \gtrsim 1$, implying that on average more than one secondary progenitors should be generated to couple with the primary progenitor M_1, and we must generate merger configurations with multiple progenitors. To accommodate this feature, for each M_1 that satisfies $M_1 \in [\mu M_0, M_0 - M_{\rm res}]$, we generate[3] either $\mathrm{int}(r)$ or $\mathrm{int}(r)+1$ secondary progenitors of mass $M_0 - M_1$ according to whether a random number between 0 and 1 is larger or smaller than $r - \mathrm{int}(r)$. Note that the resulting merger configurations do not conserve mass exactly because the sum of the progenitor masses is slightly larger than the descendant mass. Typically most of these configurations only end up with 3 or 4 progenitors as $r \lesssim 2$ for $M_1 \lesssim 0.999 M_0$ and $\Delta z \lesssim 0.02$.

The second subtlety with method A is that since the total number of progenitors in the mass range of $[M_0/2, M_0]$ (which is equal to $\int_{M_0/2}^{M_0} \phi(M|M_0) dM$) is always slightly smaller than one (typically by 0.2% to 0.4% for $\Delta z = 0.02$; recall from Table 6.2 that $\alpha M_0 < M_0/2$), it is possible that we sometimes cannot assign any progenitors to a given descendant halo. When this happens, the descendant halo does not have any progenitor halos and is a "$0 \to 1$ event".

For a thorough description of our algorithm A, we list below all the possible merger configurations and their frequencies of occurrence for descendant halos at $z = 0$ over a single simulation timestep $\Delta z = 0.02$ and mass resolution $M_{\rm res} = 0.001 M_0$. This information is also summarised in Table 6.3 and Fig. 6.6. In general, the relative frequencies of different

[3]Here $\mathrm{int}(r)$ is defined to be the largest integer n that satisfies $n \leq r$.

merger configurations are insensitive to the descendant mass M_0 but do depend on the Δz and $M_{\rm res}$ used in the Monte Carlo simulation. For example, the fraction of $1 \to 1$ events increases as Δz decreases; and if $M_{\rm res}$ is chosen to be larger than $(1-\mu)M_0 \sim 0.03M_0$, then there are no $3 \to 1$ or $4 \to 1$ mergers at each timestep and mass conservation is exactly respected.

I. About 12% in the ellipsoidal model (21% for spherical) of descendant halos have two progenitors each. These are binary pairs drawn from the symmetric part of the progenitor mass function $\phi_{\rm sym}$, where $M_1 \in [M_0/2, \mu M_0]$ and $M_0 = M_1 + M_2$ (Fig. 6.6 A■).

II. About 69% (60%) of descendant halos have only one progenitor each. The majority ($\gtrsim 99\%$) of these descendants originally have binary progenitors but the smaller progenitor is discarded since $M_2 < M_{\rm res}$ (i.e. $M_1 \gtrsim M_0 - M_{\rm res}$) (Fig. 6.6 A■). The rest ($\lesssim 1\%$) of these descendant have progenitors with $M_1 \in [M_0/2, \mu M_0]$ and originate from the small asymmetric part $\phi_{\rm asym}$ of the progenitor mass function where $r < 1$ (Fig. 6.6 A□).

III. About 18% of descendant halos have three or four progenitors each, typically consisting of one massive progenitor and two or three very small secondary progenitors ($\lesssim (1-\mu)M_0 \sim 0.03M_0$). These have $M_1 \in [\mu M_0, M_0 - M_{\rm res}]$ (Fig. 6.6 A■).

IV. About 0.4% (0.3%) of the descendants have no progenitors due to the sharp cutoff of the primary progenitor mass at $M_0/2$ discussed above.

The red solid curves in Fig. 6.7 compare the progenitor mass functions from this Monte Carlo algorithm with the analytic eq. (6.3) of the spherical EPS model. Fig. 6.8 shows the same thing except everything is for the ellipsoidal EPS model, where we use eq. (6.4) to compute the progenitor mass function for each small simulation timestep. Both figures show excellent agreement ($< 10\%$ deviation) at $z_1 - z_0 = 0.24$, 2.07, 7, and 15 for a descendant halo of mass $10^{13} M_\odot$ at $z_0 = 0$. We have tested other descendant masses ($10^{12} \lesssim M_0 \lesssim 10^{14} M_\odot$) and found equally good agreement. This agreement also provides numerical verification of the criterion introduced in §6.3.1.

6.5.3 Method B

Two features in method A may seem unnatural. First, as shown in Table 6.3 and discussed in the previous section, a small fraction ($\sim 0.3\%$ to 0.4%) of the descendant halos in method A are not assigned any progenitors in one timestep because $\int_{M_0/2}^{M_0} \phi(M|M_0) dM \approx 0.997$ (for $\Delta z = 0.02$ and a large range of M_0) and is not exactly unity. It is important to note that though these descendants are rare, one cannot remove them from method A by modifying the normalisation of $\phi(M|M_0)$ in the mass range of $[M_0/2, M_0]$, as such a modification is amplified with iterations and leads to a large error in $\phi(M|M_0)$ after many timesteps.

Second, due to the asymmetry in the EPS $\phi(M|M_0)$, we have assigned a small fraction (0.4% to 0.8% for parameters used in Table 6.3) of the descendant halos to $1 \to 1$ events. There is therefore a small chance that a progenitor of mass comparable to half of the descendant mass does not have any companions, corresponding to a large deficit between the mass of the descendant halo and the total mass of its progenitors.

The first feature can be avoided by decreasing the lower limit of the mass range from which the primary progenitor is drawn from $M_0/2$ to αM_0, where α is defined in eq. (6.14) and ranges from $\alpha \approx 0.36$ to 0.45 in Table 6.2. The second feature can be altered by distributing the secondary progenitors in a slightly different way. These options motivate us to invent Method B with the following set up:

1. We assume the primary progenitor mass lies in the mass range $[\alpha M_0, M_0]$. This condition guarantees that every descendant halo has a primary progenitor of mass $> \alpha M_0$ due to the definition of α.

2. We then assign secondary progenitors to primary progenitors from the left side of αM_0. For simplicity, whenever possible, we make only binary configurations, each of which contains one primary and one secondary progenitor. We start with the primary and secondary progenitor bins that share the αM_0 boundary (i.e. nearly equal-mass pairs) and work our way outwards to the $M_1 \gg M_2$ pairs. This is a natural decision as this way of pairing the primary and secondary masses minimises the difference between M_0 and $M_1 + M_2$. Specifically, for a given M_1 bin, we determine its binary companion's mass M_2 from

$$\int_{M_2}^{\alpha M_0} \phi(M|M_0)\, dM = \int_{\alpha M_0}^{M_1} \phi(M|M_0)\, dM\,, \tag{6.16}$$

which guarantees that we always have an equal number of secondary progenitors to pair up with the primary halos. Note that since $\alpha < 0.5$ it is generally true that $M_0 > M_1 + M_2$.

3. One caveat with step 2 above is that this simple binary pairing scheme works for a large range of masses but needs to be modified near the end points when M_1 is close to M_0 and $M_2 \ll M_1$. This is because the scheme starts out with nearly equal-mass pairs at $M_1 \sim M_2 \sim \alpha M_0$ and $M_1 + M_2 < M_0$, and the asymmetric shape of the progenitor mass function is such that the method produces pairs with increasing $M_1 + M_2$ as we move outward from αM_0. The equality $M_1 + M_2 = M_0$ is reached when M_1 is slightly larger than μM_0 (typically at $0.99 M_0$), beyond which there are more secondary progenitors left to be paired than the primary ones. We therefore stop the binary pairing when $M_1 + M_2 = M_0$ is reached. From this point on, we instead use the same multiple merger configurations as in method A. For simplicity in the following few paragraphs, we denote this transitional M_1 as $\mu' M_0$.

In summary, methods A and B are closely related and are compared side-by-side in Table 6.3 and Fig. 6.6. They have identical merger configurations in the following regions:

I. The high-M_1 region $M_1 \in [M_0 - M_{\rm res}, M_0]$, where 60% to 70% of descendant halos belong. The secondary progenitor is below $M_{\rm res}$, so M_1 is effectively the sole progenitor (i.e. $N_p = 1$) for these descendants

II. The region $M_1 \in [\mu' M_0, M_0 - M_{\rm res}]$ (μ' replaced by μ in method A), where 17% to 20% of descendant halos belong. These descendants all each have 3 or more progenitors ($N_p = 3+$).

Methods A and B differ in the following regions:

III. The binary pairing algorithm used in method B removes the sliver of $1 \to 1$ configurations in the $M_1 \in [M_0/2, \mu M_0]$ region in method A (**A** ▫) and redistributes the binary merger

configurations in this region (A) to yield a robust set of binary configurations between $\alpha M_0 \leq M_1 \leq \mu' M_0$ (B). This affects $\sim 20\%$ of the descendant halos.

IV. Since the primary progenitor mass range extends down to αM_0 instead of $M_0/2$, method B does not have any of the $0 \rightarrow 1$ configurations that are present in method A.

The green dashed curves in Figs. 6.7 and 6.8 compare the progenitor mass functions from method B with the analytic predictions of the spherical and ellipsoidal EPS models, respectively. The agreement is again excellent ($< 10\%$ deviation) at $z_1 - z_0 = 0.24$, 2.07, 7, and 15 for a descendant halo of mass $10^{13} M_\odot$ at $z_0 = 0$.

Finally, we note that mass is not strictly conserved for the multiple merger configurations generated in the $M_1 \in [\mu M_0, M_0 - M_{\rm res}]$ region of method A and the $M_1 \in [\mu' M_0, M_0 - M_{\rm res}]$ region of method B (Fig. 6.6 A, B). These configurations have more than one companion of mass $M_0 - M_1$, making the total mass of the progenitors slightly above the descendant halo mass. This issue is due to the rapid rise of the progenitor number as the secondary progenitor mass approaches zero. In principle, the small progenitors ($\lesssim (1-\mu) M_0$) that are causing this problem can be re-distributed and combined, *e.g.* , with progenitors in some of the $1 \rightarrow 1$ and $0 \rightarrow 1$ merger configurations in method A, or with some binary configurations of total masses smaller than the descendant mass in method B, to form multiple merger configurations that obey mass conservation (this, in fact, is what happens in method C below, where mass conservation is strictly respected). We have checked that this can be done successfully without violating mass conservation down to very small $M_{\rm res}$ and find that in practice, these modifications do not introduce significant changes to the statistical properties of the halo merger histories. We have therefore chosen to present the simpler version of each model. It is also worth noting that in the EPS theory, mass conservation only has to be obeyed statistically and is *not* required for individual merger configurations.

6.5.4 Method C (Multiple Mergers)

As shown in Table 6.3, methods A and B both produce comparable number of descendants with binary ($N_p = 2$) and multiple ($N_p = 3+$) progenitors in a single timestep. The importance of multiple merger configurations have been emphasised by a number of authors (e.g., Kauffmann and White 1993, Somerville and Kolatt 1999, Neistein and Dekel 2008a). It is therefore interesting to explore the relative importance of binary vs multiple mergers by relaxing the binary assumption. Our method C is designed for this purpose. More specifically, this method does not have any restrictions on the number of progenitors in each merger configuration. We only require that the total progenitor mass of every merger configuration be smaller than (or equal to) the descendant halo mass.

We now describe method C:

1. To prevent the formation of $0 \rightarrow 1$ merger configurations we mimic the setup of method B and choose to draw primary progenitors from the mass range $M_1 \in [\alpha M_0, M_0]$. Thus methods B and C share the same distribution of primary and secondary progenitor mass bins.

2. As with method B, we form merger configurations by assigning secondary progenitors

to progenitors in primary bins. Every primary bin starts with one merger configuration: that which contains only the primary progenitor itself, and has a probability $N_{\rm conf}$ equal to the number of primary progenitors in the bin. The assignment of secondary progenitors to primary bins is done in order of decreasing secondary progenitor mass. For each secondary bin, we scan the primary bins in order of increasing primary progenitor mass to find configurations with room to hold *at least* one secondary progenitor from the bin in question (recall that we require the sum of progenitor masses to never exceed the descendant mass).

3. When a valid configuration is found, we always assign the *maximal* number of secondary progenitors to that configuration. For example, suppose we start to assign secondary progenitors from a bin with central mass M_2 (say there are N_2 such progenitors in this bin), and find a valid configuration of probability $N_{\rm conf}$ and total progenitor mass M_{tot}. The maximum number n_{max} of secondary progenitors that can be added into each realisation of this configuration is equal to int$[(M_0 - M_{tot})/M_2]$. Therefore, we can maximally assign $N_{max} = n_{max} \times N_{\rm conf}$ secondary progenitors to this configuration.

I. If $N_{max} > N_2$, we break the configuration into two: one contains the original set of progenitors, with a probability equal to $(1 - N_2/N_{max}) \times N_{\rm conf}$; the other contains the original set of progenitors plus n_{max} secondary progenitors of mass M_2, with a probability equal to $(N_2/N_{max}) \times N_{\rm conf}$. In this case all the secondary progenitors of the current secondary bin are assigned.

II. If $N_{max} \leq N_2$ we simply add the n_{max} secondary progenitors of mass M_2 to the configuration, and update the list of progenitors in the configuration. $N_{\rm conf}$, the configuration's probability does not change. The number of remaining secondary progenitors to be matched is now $N_2 - N_{max}$, and we continue our search across merger configurations (in order of increasing primary progenitor mass) until all of them have been assigned.

Once a secondary bin is fully assigned, we move on to the next secondary bin (of a slightly smaller mass) and repeat the same assignment procedure. As this process goes on all configurations are gradually filled with secondary progenitors of smaller and smaller mass. For technical convenience, the number of configurations in each primary bin and the number of unique progenitor masses in each configuration are both limited to be fewer than 6. In practice, we find that this setup allows us to successfully assign all secondary progenitors in the mass range $[M_{\rm res}, \alpha M_0]$, even when the mass resolution of each time step is as low as $M_{\rm res} = 0.001 M_0$.

In fact this dense packing of secondary progenitors into primary bin configurations manages to distribute efficiently *all* secondary progenitors in $[M_{\rm res}, \alpha M_0]$ in only a fraction of the available primary progenitors. As seen in Fig. 6.6 C, only 2% (5% for spheircal) of the primary progenitors (at the low mass end) are grouped with secondary progenitors and the remaining 98% (95%) are $1 \rightarrow 1$ events. We note that even though there are far more secondary progenitors than primary progenitors, this is possible because many secondary progenitors have exceedingly small masses and can be efficiently distributed into the mass reservoirs of relatively few primary progenitors.

The execution of method C is as follows:

(i) Generate a primary progenitor M_1 from the mass range $[\alpha M_0, M_0]$ of the EPS pro-

genitor mass function. Determine which primary bin contains M_1.

(ii) If $M_1 > M_0 - M_{\rm res}$, no more progenitors are generated; if $M_1 \leq M_0 - M_{\rm res}$, a random number determines which merger configuration to choose according to the probability distribution of all possible configurations associated with the given primary bin. The progenitors of the chosen configuration are then generated.

For a better understanding of method C, we show in Table 6.3 and discuss below all the possible merger configurations and their frequencies of occurrence for descendant halos (regardless of their masses) at $z = 0$, assuming timestep $\Delta z = 0.02$ and mass resolution $M_{\rm res} = 0.001 M_0$:

I. About 98% (95% for spherical) of the descendant halos have only one progenitor each.

A) About 2/3 of these descendants' progenitors are within the resolution limit of the descendant mass (i.e. $M_1 \gtrsim M_0 - M_{\rm res}$, see figure 6.6 c■).

B) The remaining 1/3 of these descendant halos' progenitors have masses below $M_0 - M_{\rm res}$. As discussed above, these massive primary progenitors are not assigned any secondary companions because all the available secondary progenitors are maximally assigned to the less massive primary bins. Note that this region extends to masses below μM_0 (c■).

II. For the remaining primary progenitor bins, there are no configurations having only two progenitors. All in all, 0.01% (0.1% for spherical) of all descendants have three progenitors (c■); 0.3% (2%) have four progenitors (c□); 1.7% (2.9%) have five or more progenitors (c□). The progenitor count for a given configuration can be rather large reaching values of more than 100.

As in methods A and B, the values quoted above depend on Δz and $M_{\rm res}$. They also depend on the maximal number of configurations allowed in each primary bin and the maximal number of unique progenitor masses allowed in each configuration.

The blue dotted curves in Figs. 6.7 and 6.8 compare the progenitor mass functions from this Monte Carlo algorithm with the analytic predictions of the spherical and ellipsoidal EPS models, respectively. They again show excellent agreement ($< 10\%$ deviation) at $z_1 - z_0 =$ 0.24, 2.07, 7, and 15 for a descendant halo of mass $10^{13} M_\odot$ at $z_0 = 0$.

6.6 Comparison of Higher-Moment Statistics in Algorithms A, B, C, and KW93

We have designed Monte Carlo algorithms A, B, and C for constructing merger trees that can accurately reproduce the EPS prediction for the progenitor mass function $\phi(M|M_0)$ across each individual timestep. According to the discussion in §6.3.1, these methods should then accurately generate the progenitor mass function at *any* look-back time in any number of timesteps. Figs. 6.7 and 6.8 show that this is indeed the case for both the spherical and ellipsoidal EPS models. Including KW93, there are now four methods that are completely consistent with the EPS $\phi(M|M_0)$. The results of the ellipsoidal version of KW93 have been shown in Fig. 6.8 as well.

Despite this agreement, we recall that the progenitor mass function is only one of many statistical properties of a halo merger tree. Even though all four algorithms are degenerate in $\phi(M|M_0)$, they are likely to (and should) differ in their predictions for other statistical quantities. Here we investigate two such quantities as an illustration: (i) $\phi^{(N_p)}(M|M_0)$, the progenitor mass function for the subset of descendant halos that have N_p progenitors. The sum of $\phi^{(N_p)}(M|M_0)$ over all N_p is equal to $\phi(M|M_0)$. (ii) $\phi^{(i_{th})}(M|M_0)$, the distribution of the i_{th} most massive progenitor of each descendant halo. Again, the sum of $\phi^{(i_{th})}(M|M_0)$ over all i is equal to $\phi(M|M_0)$. These two statistics are two obvious ways of decomposing the total $\phi(M|M_0)$ into individual moments: $\phi^{(N_p)}$ separates flourishing trees from quiescent trees, while $\phi^{(i_{th})}$ compares the individual distributions of the primary, secondary and more minor progenitors, which are relevant for modelling galaxy formation through mergers. Other statistics such as the distributions of halo formation time and last major merger time (e.g., Cole et al. 2008, Moreno et al. 2008) and the factorial moments of the partition function [Sheth and Pitman, 1997] are also useful. Some of these will be examined in our next paper.

To compute these moments, we set the descendant halo at redshift zero to be $10^{13}M_{\odot}$, and the mass resolution to be $4 \times 10^{10}M_{\odot}$. The results are plotted at two look-back times ($z_1 - z_0 = 0.51, 2.07$) in Figs. 6.9-6.12, where Figs. 6.9 and 6.10 show $\phi^{(N_p)}(M|M_0)$ for the spherical and ellipsoidal EPS models, respectively, while Figs. 6.11 and 6.12 show $\phi^{(i_{th})}(M|M_0)$. In each figure, results from our three methods (red solid for A, green dashed for B, blue dotted for C) and from our implementation of KW93 (orange dash-dotted) are shown for comparison. These figures clearly indicate that methods A, B, C, and KW93 generate distinct predictions for these specific moments of the progenitor mass distribution. Some of the notable differences are:

1. Method C produces a much lower amplitude for the $N_p = 2$ and 3 moments than methods A and B. This is because C is designed to be a multiple-merger algorithm that effectively does not generate *any* binary configuration in one individual timestep (note the absence of the $N_p = 2$ entry for method C in Table 6.3). This feature can been seen by the absence of blue short-dashed curves in the $N_p = 2$ and 3 panels in Figs. 6.9 and 6.10, i.e., there are almost no descendant halos having only two or three progenitors in method C at $z = 0.51$. By contrast, methods A and B have a wealth of descendants with binary progenitors at these redshifts.

2. The removal of the binary assumption in method C leads to many features in the moments of the progenitor distributions. By contrast, the predictions from A and B are mostly power-laws, or at least smooth functions, in the progenitor mass. This difference is due to the fact that the merger configurations in the binary methods are much more regulated than those in the non-binary method: a binary configuration contains only two progenitors, the total mass of which is always quite close (if not equal) to the descendant mass, whereas the distribution of progenitor masses in a multiple configuration can have various forms, which can easily affect, *e.g.* , the ranking of the progenitor masses and the number of progenitors. It is interesting to note that the predictions of KW93 are fairly smooth functions in spite of the fact that it does not assume binary. This is likely because the way progenitors are assigned in KW93 effectively suppresses the probability of mergers

involving multiple progenitors.

3. The differences between method A and B are more subtle because they are both mostly binary methods. The main feature that distinguishes A from B is in the distribution of the most massive progenitor (i.e. $i_{th} = 1$) shown in the first columns of Figs. 6.11 and 6.12. At the high mass end, method B has a slightly broader shape for the primary progenitor mass than method A. This is expected, because it is the case across every time step by construction (the primary bins in method B extend down to αM_0 as opposed to $M_0/2$ for method A). At the low mass end, however, there is a long tail in the distribution of primary progenitor masses in method A, which is not present in other methods. This tail is caused by the fact that in method A, there is a small chance ($\sim 0.3\%$) at every timestep that a primary progenitor completely disappears, transferring the rank of "primary" to one of the much smaller secondary progenitors. Over several timesteps this rare occurrence affects more and more branches of the merger tree and can significantly modify the primary progenitor statistics.

In summary, we have constructed three Monte Carlo algorithms that can all reproduce closely the progenitor mass function of the EPS model (both spherical and ellipsoidal). The methods, however, produce significantly different higher moments of the progenitor distributions. They are also very different from KW93. Either a theoretical model more complete than the EPS or direct N-body results will be needed to determine which, if any, of the thus-far successful algorithms is the winner. We will turn to this subject in the next paper (Zhang, Fakhouri & Ma, in preparation).

6.7 Conclusions and Discussion

Monte Carlo algorithms based on the spherical EPS model have been an essential tool for many studies of galaxy and structure formation. These algorithms allow one to generate realisations of actual halo merger histories starting from a limited set of statistical information about dark matter halo properties provided by the EPS model. Since the EPS model does not uniquely determine many statistical quantities of halo mergers beyond the progenitor mass function, there is considerable freedom in how to combine progenitors to form descendant halos in each time step in a Monte Carlo algorithm.

The emphasis of this paper is on elucidating and quantifying the ability of a Monte Carlo algorithm to construct merger trees that match the analytic progenitor mass function of the EPS model (both the spherical and ellipsoidal versions). Four main conclusions can be drawn:

1. We have shown rigourously that to match the EPS progenitor mass function accurately at any look-back time, it is necessary and *sufficient* for a Monte Carlo algorithm to reproduce the exact progenitor mass function at each time step.

2. We have reviewed and compared the four most frequently used Monte Carlo algorithms based on the spherical EPS model in the literature: Lacey and Cole 1993, Kauffmann and White 1993, Somerville and Kolatt 1999, and Cole et al. 2000. As seen in Figs. 6.2-6.4,

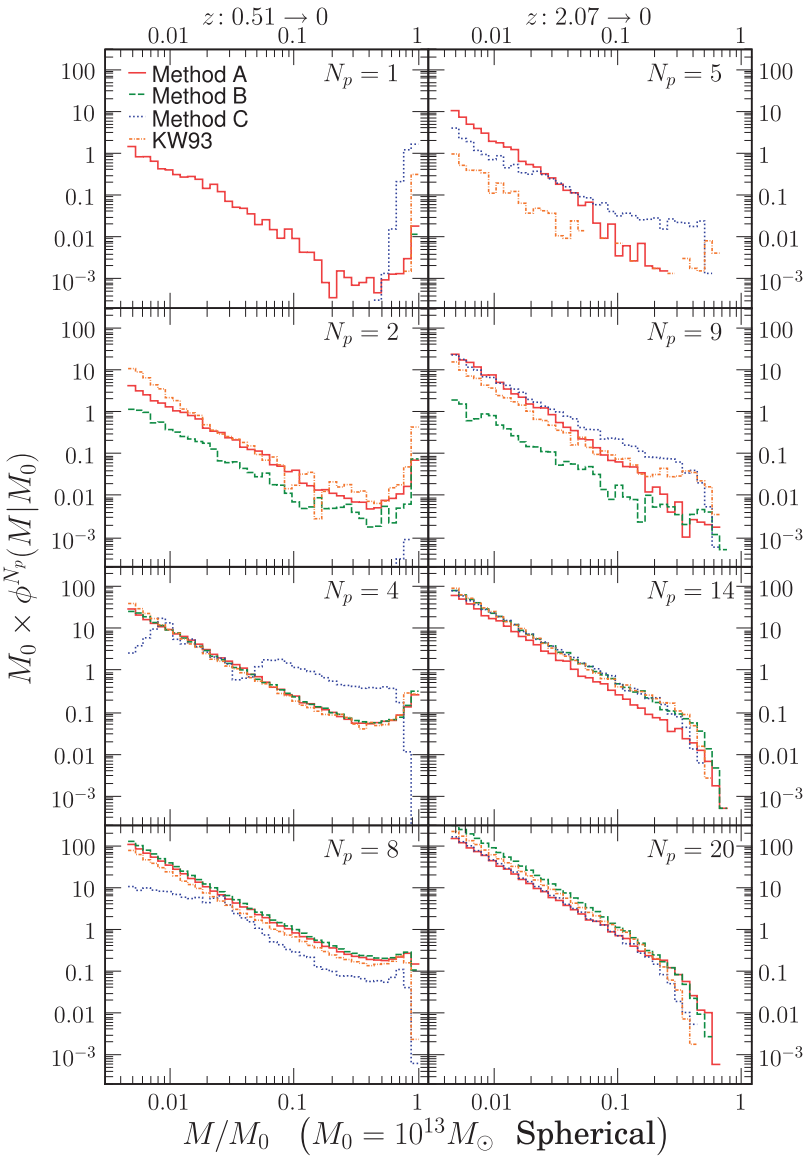

Figure 6.9: Predictions of algorithms A (red solid), B (green dashed), C (blue dotted), and KW93 (orange dash-dotted) for $\phi^{(N_p)}(M, z|M_0, z_0)$, the mass function of progenitors for descendant halos that have a total of N_p progenitors. Two look-back redshifts are shown: $z - z_0 = 0.51$ (left) and 2.07 (right). For each redshift, four representative values of N_p are shown (from top down). The simulations are for the spherical EPS model and assume a descendant halo mass of $10^{13} M_\odot$ at $z_0 = 0$ and mass resolution of $M_{\rm res} = 4 \times 10^{10} M_\odot$.

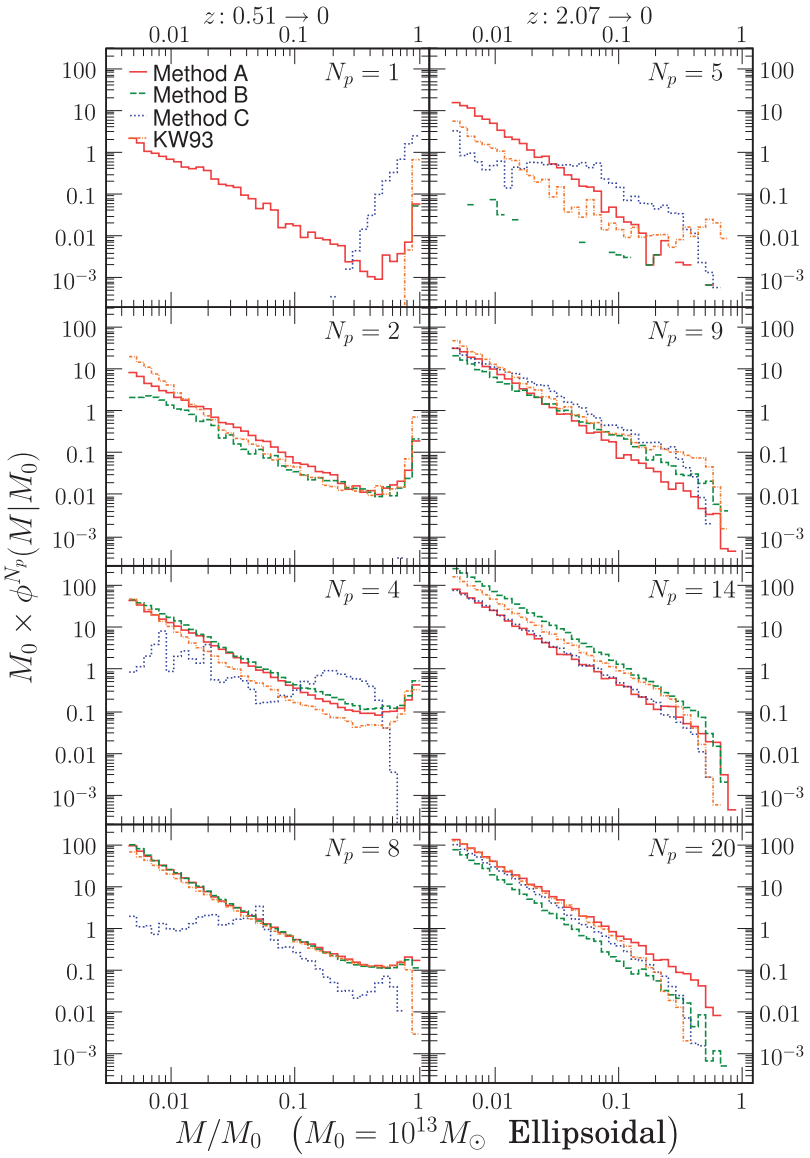

Figure 6.10: Same as Fig. 6.9 except the Monte Carlo results are generated from the ellipsoidal instead of the standard spherical EPS model.

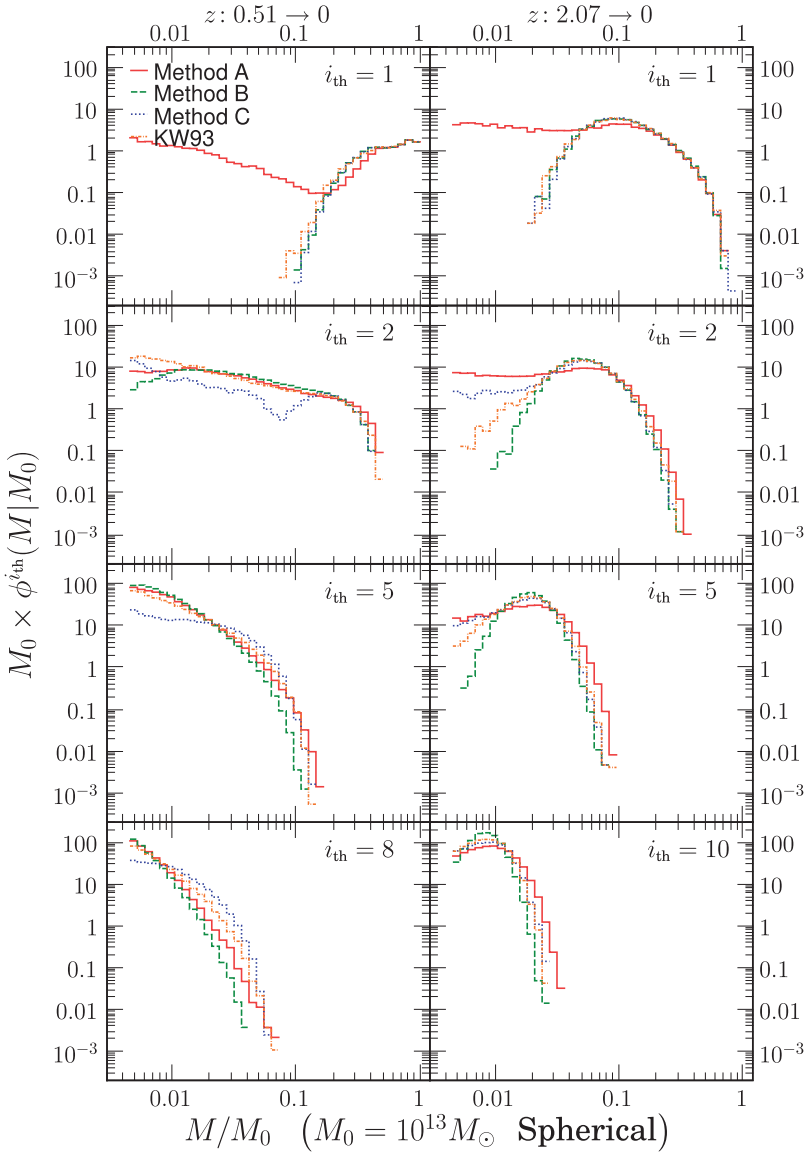

Figure 6.11: Same as Fig. 6.9 except for a different progenitor statistic: $\phi^{(i_{\rm th})}(M, z|M_0, z_0)$, the mass function of the i_{th} most massive progenitor.

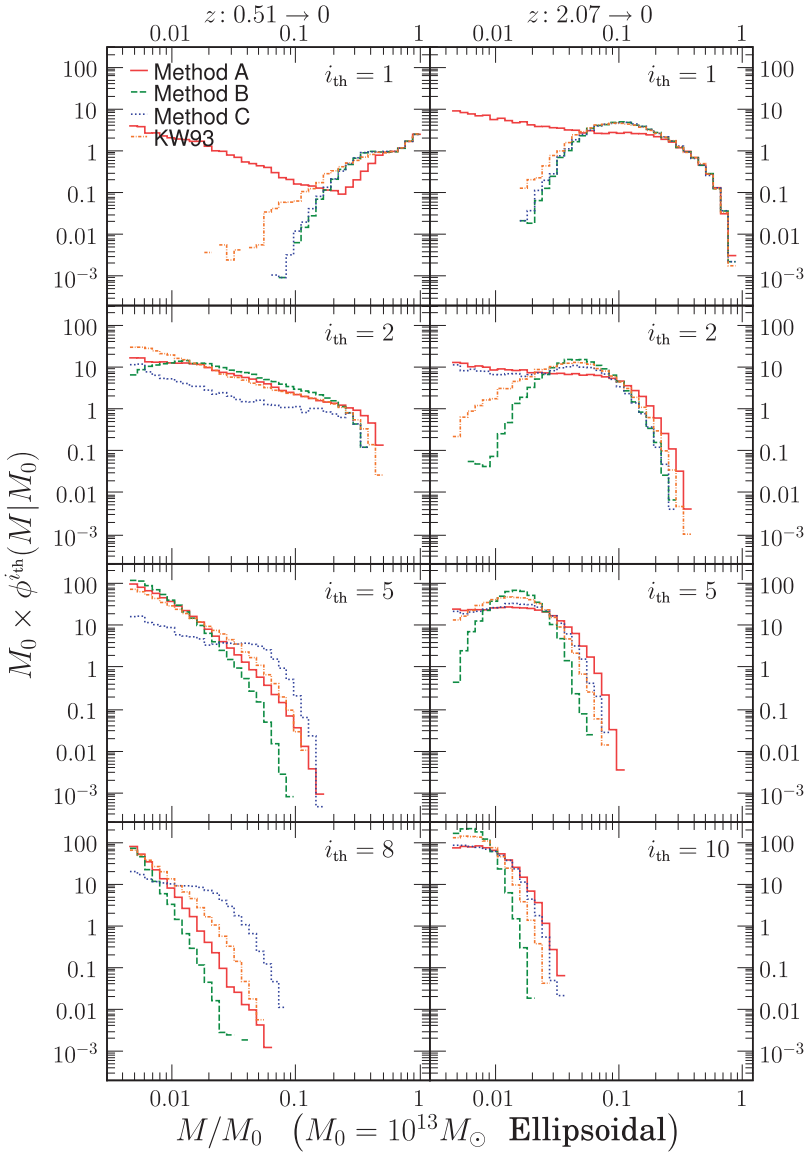

Figure 6.12: Same as Fig. 6.11 except the Monte Carlo results are generated from the ellipsoidal instead of the spherical EPS model.

all but KW93 only approximately reproduce the spherical EPS progenitor mass function at each time step, resulting in large deviations from the spherical EPS predictions after the accumulation of small errors over many time steps.

Their problems (see Table 6.1 for details) can be summarised as: (i) SK99 generally over-estimates the abundances of small progenitors by about a factor of two; (ii) LC93 over-produces progenitors by a factor of a few when the look-back time is large ($\Delta z \gg 1$); (iii) C00 under-predicts the progenitor abundance at the high mass end when the look-back time is large. The origin of these discrepancies frequently comes from the incompatibility between the binary merger assumption used in the Monte Carlo algorithm (e.g. LC93, C00) and the asymmetric progenitor mass function of the EPS model.

3. We have designed three new Monte Carlo algorithms that all reproduce closely the EPS progenitor mass function over a broad range of redshift ($z_1 - z_0$ up to at least 15) and halo mass. In addition, the algorithms are tested for both the spherical and ellipsoidal EPS models and the results are shown in Figs.6.7 and 6.8. We see that all three methods perform equally well at reproducing the respective progenitor mass function at higher redshifts, regardless of whether the spherical progenitor mass function eq. (6.3) or ellipsoidal progenitor mass function eq. (6.4) is used as input.

4. As emphasised throughout the paper, the EPS model only provides a partial statistical description of dark matter halo properties; it does not tell us explicitly how to group progenitors into descendants in a Monte Carlo realisation. Therefore, there are different ways to combine progenitors into descendant halos in consistent Monte Carlo algorithms.

We have used our three new algorithms to illustrate this exact point. Despite their success in generating merger trees that accurately reproduce the EPS progenitor mass function, Figs. 6.9-6.12 show that the three algorithms make significantly different predictions for quantities such as the distribution of the most (or the 2_{nd} or 3_{rd} most) massive progenitor masses, and the mass function of progenitors in descendant halos with N_p ($= 1, 2, 3...$) progenitors. A theory more complete than EPS would be needed to predict these higher-order merger statistics and break the degeneracies in the progenitor mass function. Alternatively, comparisons with N-body simulations should determine which, if any, of the three new algorithms is viable. We view the EPS models (spherical or ellipsoidal), Monto Carlo algorithms, and N-body simulations as three major components in the general study of the formation, growth, and clustering of dark matter halos. In this paper we have focused on the first two areas, comparing various Monte Carlo algorithms for generating halo merger trees and quantifying their abilities to consistently match the analytical EPS progenitor mass functions over a broad range of mass and redshift. In our next paper (Zhang, Fakhouri, Ma 2008b), we will turn to comparisons with the Millennium simulation.

Several recent papers have investigated other Monte Carlo methods (see, e.g., Parkinson et al. 2008, Neistein and Dekel 2008b, Moreno et al. 2008, Neistein and Dekel 2008a. Although a complete review of these methods is beyond the scope of this paper, it is worth pointing out some of their features. The method of Moreno et al. [2008] is essentially equivalent to LC93 but is based on the ellipsoidal collapse model and is discretized in mass instead of redshift. The two progenitor masses for each time step are assigned using computer-

generated random walks and moving barriers. Since the asymmetry problem of the progenitor mass function is also present in the ellipsoidal model, this method does not accurately reproduce the theory-predicted progenitor mass function at each time step. Such a discrepancy is amplified with increasing redshift and is indeed shown in Fig. 5, 6, and 7 of Moreno et al. [2008].

Neistein and Dekel [2008a] have proposed a method that exactly reproduces the progenitor mass function of the spherical EPS model at each time step. This feature alone guarantees it to be consistent with EPS at any look-back time according to our discussion in §3.1. However, since the method requires solving several differential equations with non-trivial boundary conditions for the progenitor masses, it is technically harder to implement it.

Finally, the methods described in Parkinson et al. [2008] and Neistein and Dekel [2008b] are proposed to mimic N-body results. They are based on fitting to N-body data rather than the EPS theory. It will be interesting to compare the predictions for the various merger statistics discussed in this paper from these methods with those from our ellipsoidal EPS-based methods and from N-body simulations. This will be done in the next paper.

Bibliography

U. Abbas and R. K. Sheth. The environmental dependence of clustering in hierarchical models. *MNRAS*, 364:1327–1336, December 2005. doi: 10.1111/j.1365-2966.2005.09654.x.

R. E. Angulo, C. G. Lacey, C. M. Baugh, and C. S. Frenk. The fate of substructures in cold dark matter haloes. *MNRAS*, 399:983–995, October 2009. doi: 10.1111/j.1365-2966.2009.15333.x.

V. Avila-Reese, P. Colín, S. Gottlöber, C. Firmani, and C. Maulbetsch. The Dependence on Environment of Cold Dark Matter Halo Properties. *ApJ*, 634:51–69, November 2005. doi: 10.1086/491726.

J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay. The statistics of peaks of Gaussian random fields. *ApJ*, 304:15–61, May 1986. doi: 10.1086/164143.

C. M. Baugh. A primer on hierarchical galaxy formation: the semi-analytical approach. *Reports of Progress in Physics*, 69:3101–3156, December 2006. doi: 10.1088/0034-4885/69/12/R02.

A. J. Benson, S. Cole, C. S. Frenk, C. M. Baugh, and C. G. Lacey. The nature of galaxy bias and clustering. *MNRAS*, 311:793–808, February 2000.

A. J. Benson, M. Kamionkowski, and S. H. Hassani. Self-consistent theory of halo mergers. *MNRAS*, 357:847–858, March 2005. doi: 10.1111/j.1365-2966.2005.08679.x.

J. C. Berrier, J. S. Bullock, E. J. Barton, H. D. Guenther, A. R. Zentner, and R. H. Wechsler. Close Galaxy Counts as a Probe of Hierarchical Structure Formation. *ApJ*, 652:56–70, November 2006. doi: 10.1086/507573.

P. Bett, V. Eke, C. S. Frenk, A. Jenkins, J. Helly, and J. Navarro. The spin and shape of dark matter haloes in the Millennium simulation of a Λ cold dark matter universe. *MNRAS*, 376:215–232, March 2007. doi: 10.1111/j.1365-2966.2007.11432.x.

J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion set mass functions for hierarchical Gaussian fluctuations. *ApJ*, 379:440–460, October 1991. doi: 10.1086/170520.

R. G. Bower, A. J. Benson, R. Malbon, J. C. Helly, C. S. Frenk, C. M. Baugh, S. Cole, and C. G. Lacey. Breaking the hierarchy of galaxy formation. *MNRAS*, 370:645–655, August 2006. doi: 10.1111/j.1365-2966.2006.10519.x.

M. Boylan-Kolchin, C.-P. Ma, and E. Quataert. Dynamical friction and galaxy merging time-scales. *MNRAS*, 383:93–101, January 2008. doi: 10.1111/j.1365-2966.2007.12530.x.

M. Boylan-Kolchin, V. Springel, S. D. M. White, and A. Jenkins. There's no place like home? Statistics of Milky Way-mass dark matter halos. *arXiv:0911.4484 [astro-ph]*, November 2009a.

M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, and G. Lemson. Resolving cosmic structure formation with the Millennium-II Simulation. *MNRAS*, 398:1150–1164, September 2009b. doi: 10.1111/j.1365-2966.2009.15191.x.

K. Bundy, M. Fukugita, R. S. Ellis, T. Kodama, and C. J. Conselice. A Slow Merger History of Field Galaxies since z ˜ 1. *ApJL*, 601:L123–L126, February 2004. doi: 10.1086/381891.

J. M. Burkey, W. C. Keel, R. A. Windhorst, and B. E. Franklin. Galaxy pairs in deep HST images: Evidence for evolution in the galaxy merger rate. *ApJ*, 429:L13–L17, July 1994. doi: 10.1086/187402.

R. G. Carlberg, C. J. Pritchet, and L. Infante. A survey of faint galaxy pairs. *ApJ*, 435: 540–547, November 1994. doi: 10.1086/174835.

J. D. Cohn and M. White. The formation histories of galaxy clusters. *Astroparticle Physics*, 24:316–333, December 2005. doi: 10.1016/j.astropartphys.2005.07.006.

S. Cole, C. G. Lacey, C. M. Baugh, and C. S. Frenk. Hierarchical galaxy formation. *MNRAS*, 319:168–204, November 2000.

S. Cole, J. Helly, C. S. Frenk, and H. Parkinson. The statistical properties of Λ cold dark matter halo formation. *MNRAS*, 383:546–556, January 2008. doi: 10.1111/j.1365-2966.2007.12516.x.

C. J. Conselice, M. A. Bershady, M. Dickinson, and C. Papovich. A Direct Measurement of Major Galaxy Mergers at z¡ 3. *AJ*, 126:1183–1207, September 2003. doi: 10.1086/377318.

Scott M Croom, B. J Boyle, T Shanks, R. J Smith, L Miller, P. J Outram, N. S Loaring, F Hoyle, and J da Ângela. The 2df qso redshift survey - xiv. structure and evolution from the two-point correlation function. *Monthly Notices of the Royal Astronomical Society*, 356:415, Jan 2005. doi: 10.1111/j.1365-2966.2004.08379.x. (c) 2005 RAS.

M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *ApJ*, 292:371–394, May 1985. doi: 10.1086/163168.

G. De Lucia, V. Springel, S. D. M. White, D. Croton, and G. Kauffmann. The formation history of elliptical galaxies. *MNRAS*, 366:499–509, February 2006. doi: 10.1111/j.1365-2966.2005.09879.x.

V. Desjacques. Environmental dependence in the ellipsoidal collapse model. *MNRAS*, 388: 638–658, August 2008. doi: 10.1111/j.1365-2966.2008.13420.x.

R. Diehl, H. Halloin, K. Kretschmer, G. G. Lichti, V. Schönfelder, A. W. Strong, A. von Kienlin, W. Wang, P. Jean, J. Knödlseder, J.-P. Roques, G. Weidenspointner, S. Schanne, D. H. Hartmann, C. Winkler, and C. Wunderer. Radioactive 26Al from massive stars in the Galaxy. *Nature*, 439:45–47, January 2006. doi: 10.1038/nature04364.

J. Diemand, B. Moore, and J. Stadel. Earth-mass dark-matter haloes as the first structures in the early Universe. *Nature*, 433:389–391, January 2005. doi: 10.1038/nature03270.

D. J. Eisenstein and W. Hu. Baryonic Features in the Matter Transfer Function. *ApJ*, 496: 605–+, March 1998. doi: 10.1086/305424.

D. J. Eisenstein and W. Hu. Power Spectra for Cold Dark Matter and Its Variants. *ApJ*, 511:5–15, January 1999. doi: 10.1086/306640.

V. R. Eke, S. Cole, and C. S. Frenk. Cluster evolution as a diagnostic for Omega. *MNRAS*, 282:263–280, September 1996.

O. Fakhouri and C.-P. Ma. The nearly universal merger rate of dark matter haloes in ΛCDM cosmology. *MNRAS*, 386:577–592, May 2008. doi: 10.1111/j.1365-2966.2008.13075.x.

O. Fakhouri and C.-P. Ma. Environmental dependence of dark matter halo growth - I. Halo merger rates. *MNRAS*, 394:1825–1840, April 2009. doi: 10.1111/j.1365-2966.2009.14480.x.

O. Fakhouri and C.-P. Ma. Dark matter halo growth - II. Diffuse accretion and its environmental dependence. *MNRAS*, 401:2245–2256, February 2010. doi: 10.1111/j.1365-2966.2009.15844.x.

L. Gao and S. D. M. White. Assembly bias in the clustering of dark matter haloes. *MNRAS*, 377:L5–L9, April 2007. doi: 10.1111/j.1745-3933.2007.00292.x.

L. Gao, V. Springel, and S. D. M. White. The age dependence of halo clustering. *MNRAS*, 363:L66–L70, October 2005. doi: 10.1111/j.1745-3933.2005.00084.x.

J. M. Gelb and E. Bertschinger. Cold dark matter. 1: The formation of dark halos. *ApJ*, 436:467–490, December 1994. doi: 10.1086/174922.

S. Genel, R. Genzel, N. Bouché, A. Sternberg, T. Naab, N. M. F. Schreiber, K. L. Shapiro, L. J. Tacconi, D. Lutz, G. Cresci, P. Buschkamp, R. I. Davies, and E. K. S. Hicks. Mergers and Mass Accretion Rates in Galaxy Assembly: The Millennium Simulation Compared to Observations of z ˜ 2 Galaxies. *ApJ*, 688:789–793, December 2008. doi: 10.1086/592241.

S. Genel, R. Genzel, N. Bouché, T. Naab, and A. Sternberg. The Halo Merger Rate in the Millennium Simulation and Implications for Observed Galaxy Merger Fractions. *ApJ*, 701: 2002–2018, August 2009. doi: 10.1088/0004-637X/701/2/2002.

D. M. Goldberg and M. S. Vogeley. Simulating Voids. *ApJ*, 605:1–6, April 2004. doi: 10.1086/382143.

S. Gottlöber, A. Klypin, and A. V. Kravtsov. Merging History as a Function of Halo Environment. *ApJ*, 546:223–233, January 2001. doi: 10.1086/318248.

S. Gottlöber, M. Kerscher, A. V. Kravtsov, A. Faltenbacher, A. Klypin, and V. Müller. Spatial distribution of galactic halos and their merger histories. *A&A*, 387:778–787, June 2002. doi: 10.1051/0004-6361:20020339.

F. Governato, J. P. Gardner, J. Stadel, T. Quinn, and G. Lake. On the Origin of Early-Type Galaxies and the Evolution of the Interaction Rate in the Field. *AJ*, 117:1651–1656, April 1999. doi: 10.1086/300805.

Q. Guo and S. D. M. White. Galaxy growth in the concordance ΛCDM cosmology. *MNRAS*, 384:2–10, February 2008. doi: 10.1111/j.1365-2966.2007.12619.x.

O. Hahn, C. M. Carollo, C. Porciani, and A. Dekel. The evolution of dark matter halo properties in clusters, filaments, sheets and voids. *MNRAS*, 381:41–51, October 2007. doi: 10.1111/j.1365-2966.2007.12249.x.

O. Hahn, C. Porciani, A. Dekel, and C. M. Carollo. The Tidal Origin of the Environment Dependence of Halo Assembly. *ArXiv e-prints*, 803, March 2008.

Zoltán Haiman, Eliot Quataert, and Geoffrey C Bower. Modeling the counts of faint radio-loud quasars: Constraints on the supermassive black hole population and predictions for high redshift. *The Astrophysical Journal*, 612:698, Sep 2004. doi: 10.1086/422834. (c) 2004: The American Astronomical Society.

G. Harker, S. Cole, J. Helly, C. Frenk, and A. Jenkins. A marked correlation function analysis of halo formation times in the Millennium Simulation. *MNRAS*, 367:1039–1049, April 2006. doi: 10.1111/j.1365-2966.2006.10022.x.

J. C. Helly, S. Cole, C. S. Frenk, C. M. Baugh, A. Benson, and C. Lacey. Galaxy formation using halo merger histories taken from N-body simulations. *MNRAS*, 338:903–912, February 2003. doi: 10.1046/j.1365-8711.2003.06151.x.

Joseph F Hennawi, Michael A Strauss, Masamune Oguri, Naohisa Inada, Gordon T Richards, Bartosz Pindor, Donald P Schneider, Robert H Becker, Michael D Gregg, Patrick B Hall, David E Johnston, Xiaohui Fan, Scott Burles, David J Schlegel, James E Gunn, Robert H Lupton, Neta A Bahcall, Robert J Brunner, and Jon Brinkmann. Binary quasars in the

sloan digital sky survey: Evidence for excess clustering on small scales. *The Astronomical Journal*, 131:1, Jan 2006. doi: 10.1086/498235. (c) 2006: The American Astronomical Society.

Philip F Hopkins, Lars Hernquist, Thomas J Cox, Tiziana Di Matteo, Brant Robertson, and Volker Springel. A unified, merger-driven model of the origin of starbursts, quasars, the cosmic x-ray background, supermassive black holes, and galaxy spheroids. *The Astrophysical Journal Supplement Series*, 163:1, Mar 2006. doi: 10.1086/499298. (c) 2006: The American Astronomical Society.

A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida. The mass function of dark matter haloes. *MNRAS*, 321: 372–384, February 2001.

Y. P. Jing, Y. Suto, and H. J. Mo. The Dependence of Dark Halo Clustering on Formation Epoch and Concentration Parameter. *ApJ*, 657:664–668, March 2007. doi: 10.1086/511130.

X. Kang, Y. P. Jing, H. J. Mo, and G. Börner. Semianalytical Model of Galaxy Formation with High-Resolution N-Body Simulations. *ApJ*, 631:21–40, September 2005. doi: 10.1086/432493.

G. Kauffmann and S. D. M. White. The merging history of dark matter haloes in a hierarchical universe. *MNRAS*, 261:921–928, April 1993.

G. Kauffmann, S. D. M. White, and B. Guiderdoni. The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes. *MNRAS*, 264:201–+, September 1993.

G. Kauffmann, J. M. Colberg, A. Diaferio, and S. D. M. White. Clustering of galaxies in a hierarchical universe - I. Methods and results at z=0. *MNRAS*, 303:188–206, February 1999.

Guinevere Kauffmann and Martin Haehnelt. A unified model for the evolution of galaxies and quasars. *Monthly Notices of the Royal Astronomical Society*, 311:576, Jan 2000. (c) 2000 The Royal Astronomical Society.

C. Lacey and S. Cole. Merger Rates in Hierarchical Models of Galaxy Formation. *MNRAS*, 262:627–649, June 1993.

C. Lacey and S. Cole. Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations. *MNRAS*, 271:676, December 1994.

R. J. Lavery, A. Remijan, V. Charmandaris, R. D. Hayes, and A. A. Ring. Probing the Evolution of the Galaxy Interaction/Merger Rate Using Collisional Ring Galaxies. *ApJ*, 612:679–689, September 2004. doi: 10.1086/422420.

O. Le Fèvre, R. Abraham, S. J. Lilly, R. S. Ellis, J. Brinchmann, D. Schade, L. Tresse, M. Colless, D. Crampton, K. Glazebrook, F. Hammer, and T. Broadhurst. Hubble Space Telescope imaging of the CFRS and LDSS redshift surveys - IV. Influence of mergers in the evolution of faint field galaxies from z˜1. *MNRAS*, 311:565–575, January 2000.

G. Lemson and G. Kauffmann. Environmental influences on dark matter haloes and consequences for the galaxies within them. *MNRAS*, 302:111–117, January 1999.

Y. Li, H. J. Mo, F. C. van den Bosch, and W. P. Lin. On the assembly history of dark matter haloes. *MNRAS*, 379:689–701, August 2007. doi: 10.1111/j.1365-2966.2007.11942.x.

Y. Li, H. J. Mo, and L. Gao. On halo formation times and assembly bias. *MNRAS*, 389: 1419–1426, September 2008. doi: 10.1111/j.1365-2966.2008.13667.x.

Adam Lidz, Philip F Hopkins, Thomas J Cox, Lars Hernquist, and Brant Robertson. The luminosity dependence of quasar clustering. *The Astrophysical Journal*, 641:41, Apr 2006. doi: 10.1086/500444. (c) 2006: The American Astronomical Society.

L. Lin, D. C. Koo, C. N. A. Willmer, D. R. Patton, C. J. Conselice, R. Yan, A. L. Coil, M. C. Cooper, M. Davis, S. M. Faber, B. F. Gerke, P. Guhathakurta, and J. A. Newman. The DEEP2 Galaxy Redshift Survey: Evolution of Close Galaxy Pairs and Major-Merger Rates up to z ˜ 1.2. *ApJl*, 617:L9–L12, December 2004. doi: 10.1086/427183.

A. D. Ludlow, J. F. Navarro, V. Springel, A. Jenkins, C. S. Frenk, and A. Helmi. The Unorthodox Orbits of Substructure Halos. *ArXiv e-prints*, January 2008.

Z. Lukic, K. Heitmann, S. Habib, S. Bashinsky, and P. M. Ricker. The Halo Mass Function: High Redshift Evolution and Universality. *ArXiv e-prints 0702360*, February 2007.

C.-P. Ma and E. Bertschinger. Do galactic systems form too late in cold + hot dark matter models? *ApJL*, 434:L5–L9, October 1994. doi: 10.1086/187559.

A. H. Maller, N. Katz, D. Kereš, R. Davé, and D. H. Weinberg. Galaxy Merger Statistics and Inferred Bulge-to-Disk Ratios in Cosmological SPH Simulations. *ApJ*, 647:763–772, August 2006. doi: 10.1086/503319.

Tiziana Di Matteo, Volker Springel, and Lars Hernquist. Energy input from quasars regulates the growth and activity of black holes and their host galaxies. *Nature*, 433:604, Feb 2005. doi: 10.1038/nature03335. (c) 2005: Nature.

C. Maulbetsch, V. Avila-Reese, P. Colín, S. Gottlöber, A. Khalatyan, and M. Steinmetz. The Dependence of the Mass Assembly History of Cold Dark Matter Halos on Environment. *ApJ*, 654:53–65, January 2007. doi: 10.1086/509706.

J. McBride, O. Fakhouri, and C.-P. Ma. Mass accretion rates and histories of dark matter haloes. *MNRAS*, 398:1858–1868, October 2009. doi: 10.1111/j.1365-2966.2009.15329.x.

H. J. Mo and S. D. M. White. An analytic model for the spatial clustering of dark matter haloes. *MNRAS*, 282:347–361, September 1996.

J. Moreno, C. Giocoli, and R. K. Sheth. Merger history trees of dark matter haloes in moving barrier models. *MNRAS*, 391:1729–1740, December 2008. doi: 10.1111/j.1365-2966.2008.13766.x.

C. Murali, N. Katz, L. Hernquist, D. H. Weinberg, and R. Davé. The Growth of Galaxies in Cosmological Simulations of Structure Formation. *ApJ*, 571:1–14, May 2002. doi: 10.1086/339876.

E. Neistein and A. Dekel. Merger rates of dark matter haloes. *MNRAS*, 388:1792–1802, August 2008a. doi: 10.1111/j.1365-2966.2008.13525.x.

E. Neistein and A. Dekel. Constructing merger trees that mimic N-body simulations. *MNRAS*, 383:615–626, January 2008b. doi: 10.1111/j.1365-2966.2007.12570.x.

E. Neistein, F. C. van den Bosch, and A. Dekel. Natural downsizing in hierarchical galaxy formation. *MNRAS*, 372:933–948, October 2006. doi: 10.1111/j.1365-2966.2006.10918.x.

H. Parkinson, S. Cole, and J. Helly. Generating dark matter halo merger trees. *MNRAS*, 383:557–564, January 2008. doi: 10.1111/j.1365-2966.2007.12517.x.

D. R. Patton, C. J. Pritchet, H. K. C. Yee, E. Ellingson, and R. G. Carlberg. Close Pairs of Field Galaxies in the CNOC1 Redshift Survey. *ApJ*, 475:29–+, January 1997. doi: 10.1086/303535.

D. R. Patton, C. J. Pritchet, R. G. Carlberg, R. O. Marzke, H. K. C. Yee, P. B. Hall, H. Lin, S. L. Morris, M. Sawicki, C. W. Shepherd, and G. D. Wirth. Dynamically Close Galaxy Pairs and Merger Rate Evolution in the CNOC2 Redshift Survey. *ApJ*, 565:208–222, January 2002. doi: 10.1086/324543.

P. J. E. Peebles. Tests of cosmological models constrained by inflation. *ApJ*, 284:439–444, September 1984. doi: 10.1086/162425.

W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438, February 1974.

H. B. Sandvik, O. Möller, J. Lee, and S. D. M. White. Why does the clustering of haloes depend on their formation history? *MNRAS*, 377:234–244, May 2007. doi: 10.1111/j.1365-2966.2007.11595.x.

R. K. Sheth and G. Lemson. The forest of merger history trees associated with the formation of dark matter haloes. *MNRAS*, 305:946–956, May 1999. doi: 10.1046/j.1365-8711.1999.02477.x.

R. K. Sheth and J. Pitman. Coagulation and branching process models of gravitational clustering. *MNRAS*, 289:66–82, July 1997.

R. K. Sheth and G. Tormen. Large-scale bias and the peak background split. *MNRAS*, 308: 119–126, September 1999. doi: 10.1046/j.1365-8711.1999.02692.x.

R. K. Sheth and G. Tormen. An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier. *MNRAS*, 329:61–75, January 2002. doi: 10.1046/j.1365-8711.2002.04950.x.

R. K. Sheth and G. Tormen. On the environmental dependence of halo formation. *MNRAS*, 350:1385–1390, June 2004. doi: 10.1111/j.1365-2966.2004.07733.x.

R. K. Sheth, H. J. Mo, and G. Tormen. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *MNRAS*, 323:1–12, May 2001. doi: 10.1046/j.1365-8711.2001.04006.x.

Todd A Small and Roger D Blandford. Quasar evolution and the growth of black holes. *Royal Astronomical Society*, 259:725, Dec 1992.

M. Smoluchowski. *Phys. Zeit.*, 17:557, 1916.

R. S. Somerville and T. S. Kolatt. How to plant a merger tree. *MNRAS*, 305:1–14, May 1999.

R. S. Somerville and J. R. Primack. Semi-analytic modelling of galaxy formation: the local Universe. *MNRAS*, 310:1087–1110, December 1999.

V. Springel. The cosmological simulation code GADGET-2. *MNRAS*, 364:1105–1134, December 2005. doi: 10.1111/j.1365-2966.2005.09655.x.

V. Springel, S. D. M. White, G. Tormen, and G. Kauffmann. Populating a cluster of galaxies - I. Results at [formmu2]z=0. *MNRAS*, 328:726–750, December 2001a. doi: 10.1046/j.1365-8711.2001.04912.x.

V. Springel, N. Yoshida, and S. D. M. White. GADGET: a code for collisionless and gasdynamical cosmological simulations. *New Astronomy*, 6:79–117, April 2001b. doi: 10.1016/S1384-1076(01)00042-2.

V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nat*, 435:629–636, June 2005. doi: 10.1038/nature03597.

Volker Springel, Tiziana Di Matteo, and Lars Hernquist. Modelling feedback from stars and black holes in galaxy mergers. *Monthly Notices of the Royal Astronomical Society*, 361: 776, Aug 2005. doi: 10.1111/j.1365-2966.2005.09238.x. (c) 2005 RAS.

K. R. Stewart, J. S. Bullock, R. H. Wechsler, A. H. Maller, and A. R. Zentner. Merger Histories of Galaxy Halos and Implications for Disk Survival. *ApJ*, 683:597–610, August 2008. doi: 10.1086/588579.

K. R. Stewart, J. S. Bullock, E. J. Barton, and R. H. Wechsler. Galaxy Mergers and Dark Matter Halo Mergers in ΛCDM: Mass, Redshift, and Mass-Ratio Dependence. *ApJ*, 702: 1005–1015, September 2009. doi: 10.1088/0004-637X/702/2/1005.

A. Tasitsiomi, A. V. Kravtsov, S. Gottlöber, and A. A. Klypin. Density Profiles of ΛCDM Clusters. *ApJ*, 607:125–139, May 2004. doi: 10.1086/383219.

J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz. Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. *ApJ*, 688:709–728, December 2008. doi: 10.1086/591439.

G. Tormen. The assembly of matter in galaxy clusters. *MNRAS*, 297:648–656, June 1998. doi: 10.1046/j.1365-8711.1998.01545.x.

G. Tormen, F. R. Bouchet, and S. D. M. White. The structure and dynamical evolution of dark matter haloes. *MNRAS*, 286:865–884, April 1997.

F. C. van den Bosch. The universal mass accretion history of cold dark matter haloes. *MNRAS*, 331:98–110, March 2002. doi: 10.1046/j.1365-8711.2002.05171.x.

H. Y. Wang, H. J. Mo, and Y. P. Jing. Environmental dependence of cold dark matter halo formation. *MNRAS*, 375:633–639, February 2007. doi: 10.1111/j.1365-2966.2006.11316.x.

R. H. Wechsler, J. S. Bullock, J. R. Primack, A. V. Kravtsov, and A. Dekel. Concentrations of Dark Halos from Their Assembly Histories. *ApJ*, 568:52–70, March 2002. doi: 10.1086/338765.

R. H. Wechsler, A. R. Zentner, J. S. Bullock, A. V. Kravtsov, and B. Allgood. The Dependence of Halo Clustering on Halo Formation History, Concentration, and Occupation. *ApJ*, 652:71–84, November 2006. doi: 10.1086/507120.

A. R. Wetzel, J. D. Cohn, M. White, D. E. Holz, and M. S. Warren. The Clustering of Massive Halos. *ApJ*, 656:139–147, February 2007. doi: 10.1086/510444.

A. R. Wetzel, J. D. Cohn, and M. White. Simulating subhaloes at high redshift: merger rates, counts and types. *MNRAS*, 395:1376–1390, May 2009. doi: 10.1111/j.1365-2966.2009.14424.x.

M. White. The mass of a halo. *A&A*, 367:27–32, February 2001. doi: 10.1051/0004-6361:20000357.

S. D. M. White. Formation and Evolution of Galaxies: Les Houches Lectures. *ArXiv Astrophysics e-prints*, October 1994.

D. Woods, G. G. Fahlman, and H. B. Richer. Counting Pairs of Faint Galaxies. *ApJ*, 454: 32–+, November 1995. doi: 10.1086/176461.

J. Stuart B Wyithe and Abraham Loeb. Self-regulated growth of supermassive black holes in galaxies as the origin of the optical and x-ray luminosity functions of quasars. *The Astrophysical Journal*, 595:614, Oct 2003. doi: 10.1086/377475. (c) 2003: The American Astronomical Society.

H. K. C. Yee and E. Ellingson. Statistics of close galaxy pairs from a faint-galaxy redshift survey. *ApJ*, 445:37–45, May 1995. doi: 10.1086/175670.

A. R. Zentner. The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth. *International Journal of Modern Physics D*, 16:763–815, 2007. doi: 10.1142/S0218271807010511.

J. Zhang and L. Hui. On Random Walks with a General Moving Barrier. *ApJ*, 641:641–646, April 2006. doi: 10.1086/499802.

J. Zhang, O. Fakhouri, and C.-P. Ma. How to grow a healthy merger tree. *MNRAS*, 389: 1521–1538, October 2008a. doi: 10.1111/j.1365-2966.2008.13671.x.

J. Zhang, C.-P. Ma, and O. Fakhouri. Conditional mass functions and merger rates of dark matter haloes in the ellipsoidal collapse model. *MNRAS*, 387:L13–L17, June 2008b. doi: 10.1111/j.1745-3933.2008.00472.x.

D. H. Zhao, Y. P. Jing, H. J. Mo, and G. Börner. Mass and Redshift Dependence of Dark Halo Structure. *ApJL*, 597:L9–L12, November 2003a. doi: 10.1086/379734.

D. H. Zhao, H. J. Mo, Y. P. Jing, and G. Börner. The growth and structure of dark matter haloes. *MNRAS*, 339:12–24, February 2003b. doi: 10.1046/j.1365-8711.2003.06135.x.

D. H. Zhao, Y. P. Jing, H. J. Mo, and G. Börner. Accurate Universal Models for the Mass Accretion Histories and Concentrations of Dark Matter Halos. *ApJ*, 707:354–369, December 2009. doi: 10.1088/0004-637X/707/1/354.