Cosmological Halo Formation and Mergers

LILIAN JIANG

Cosmological Halo Formation and Mergers

Lilian Jiang

A Thesis presented for the degree of
Doctor of Philosophy

Insitute for Computational Cosmology
Department of Physics
University of Durham
England

September 2014

Cosmological Halo Formation and Mergers

Lilian Jiang

Submitted for the degree of Doctor of Philosophy

September 2014

Abstract

My research has centred around establishing the nature of dark matter haloes by investigating their abundance as a function of halo mass, the formation history of each halo, commonly called the merger tree, and the internal structure of the halo, in terms of their radial density profiles and angular momentum.

In the first part of this thesis, I present a new algorithm which groups the subhaloes found in cosmological N-body simulations by structure finders such as SUBFIND into dark matter haloes whose formation histories are strictly hierarchical. One advantage of these 'Dhaloes' over the commonly used friends-of-friends (FoF) haloes is that they retain their individual identity in cases when FoF haloes are artificially merged by tenuous bridges of particles or by an overlap of their outer diffuse haloes. Dhaloes are thus well suited for modelling galaxy formation and their merger trees form the basis of the Durham semi-analytic galaxy formation model, GALFORM. Applying the Dhalo construction to the ΛCDM Millennium-2 simulation we find that approximately 90% of Dhaloes have a one-to-one, bijective match with a corresponding FoF halo. The remaining 10% are typically secondary components of large FoF haloes. Although the mass functions of both types of haloes are similar, the mass of Dhaloes correlates much more tightly with the virial mass, M_{200}, than FoF masses. Approximately 80% of FoF and bijective and non-bijective Dhaloes are relaxed according to standard criteria. For these relaxed haloes all three types have similar concentration–M_{200} relations and, at fixed mass, the concentration distributions are described accurately by log-normal distributions.

In the second part of this thesis, I present distributions of orbital parameters of infalling satellite haloes at the time of crossing the virial radius of their host halo. Detailed investigation of the orbits is crucial as it represents the initial con-

ditions which determine the later evolution of the substructure within the host. I use merger trees in a high resolution cosmological N-body simulation to trace the satellite haloes and measure their orbits when they first infall into the host halo. I find that there is a trend of the orbital parameters with the ratio between the satellite halo mass and the host halo mass at infall. I find that the more massive satellites move along more eccentric orbits with lower specific angular momentum than less massive satellites. I also search for possible correlations between different orbital parameters and provide accurate fitting formulae for the two independent orbital parameters (the total velocity and the radial-to-total velocity ratio). Using combinations of these formulae, we successfully fit all the other orbital parameters.

Cosmological Halo Formation and Mergers

Lilian Jiang

A Thesis presented for the degree of
Doctor of Philosophy

Insitute for Computational Cosmology
Department of Physics
University of Durham
England

September 2014

CONTENTS

LIST OF FIGURES

LIST OF TABLES

DECLARATION

The work described in this thesis was undertaken between 2010 and 2014 while the author was a research student under the supervision of Professor Shaun Cole and Professor Carlos Frenk at the Institute for Computational Cosmology, the Department of Physics, Durham University, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text. Unless stated otherwise all the work presented in this thesis is my own.

Chapter 2 presents a merger tree algorithm that was designed by Dr John Helly. My role in its development was to carry out a series of tests of the resulting halo catalogues and merger trees. These enabled Dr John Helly to refine the algorithm. The resulting algorithm is presented in full in Chapter 2 as its details have important implications for the halo catalogues analysed in Chapters 3 and 4. For the case of the final algorithm some examples of the tests I performed are included as part of Chapter 3.

Chapter 2, Chapter 3 and Chapter 4 of this thesis have been published in the form of a paper,

- Jiang L., Helly J, Cole S., Frenk C., *"N-body Dark Matter Haloes with simple Hierarchical Histories"*, 2014, MNRAS, 440, 2115J

Chapter 5 of this thesis has been published in the form of a paper,

- Jiang L., Cole S., Sawala T., Frenk C.S., *"Orbital parameters of infalling satellite haloes in the hierarchical ΛCDM model"*, 2015, MNRAS, 448, 1674J

ACKNOWLEDGEMENTS

The thesis is dedicated firstly to my supervisors, Professor Shaun Cole and Professor Carlos Frenk. They have given me enormous help during my PhD study. I have learned a lot from not only their remarkable experiences in research and mentoring but also their perfect attitudes in life. They are always friendly and kind to all people even including those who hold different opinions. I feel indebted to have taken up so much of Shaun's extra time in helping me with every problem I have met in research. His hard work has helped me greatly in finishing the research papers and the thesis. The most important lesson Carlos has taught me is that being smart is not enough to be a top researcher; one needs to work harder and more efficiently to get to the top. All in all, what I have learned from them will make me both a better researcher and person. I thank Doctor John Helly and Doctor Till Sawala for their excellent co-supervision in my two main projects. For John, I would like to thank for his patience in helping me sort out all the technical problems. For Till, I would like to thank his valuable suggestions in my second project.

I would like to thank everyone at the ICC and the astronomy group in Durham. Especially, I would like to thank Professor Carlton Baugh, the first person in the world to ask me to carry heavy computers and boxes. It is not a complaint, as I want to thank him for believing my ability to try some impossible tasks which I did not believe I could accomplish. When I was finishing the thesis as when I was moving a heavy fridge, he gave me consistent support to

enable me to carry on. Now I am even strong enough to help moving heavy wardrobes and sofas for my friends. I would like to thank Mark, Daniel, Yetli, Michelle and other friends in the Physics Department. I also would like to thank my Chinese friends Difu, Yan, and Baojiu for their companionship.

Furthermore, I want to thank my parents and my elder sister for their support which is the source of my strength. For my father, although we rarely speak to each other and he is the one who is skeptical whenever I make progress in my life, I understand that he only wants me to be a modest person and push me to work harder. Finally, I should thank my husband, Mao, who works one hundred times harder than me to build our future and looks after me carefully although we are separated by the Atlantic Ocean.

Dedicated to

My husband Mao, my dad, Daji, my mum, Shurong, finally my dear sister, Lihui

1 INTRODUCTION

During my childhood, my dad lent part of our house to a school as the library. I spent most of my childhood there and had a dream of being an astronomer after reading the biography about the excellent astronomer, Zhang Heng (Chinese; AD 78 to 139) during the Han dynasty. With the improvement of my knowledge of fundamental physics and math, I have started to devote myself to astronomy research under the supervision of excellent astronomers. During my PhD, my research has centred around hierarchical structure and galaxy formation models.

In hierarchical dark matter dominated cosmologies, such as standard ΛCDM, galaxy formation is believed to be intimately linked to the formation and evolution of dark matter haloes (a hypothetical component of a galaxy that envelops the galactic disc and extends well beyond the edge of the visible galaxy). Baryonic gas falls into dark matter haloes, cools and settles into centrifugally supported star-forming discs. Thus the evolution of the galaxy population is driven by the evolution of the population of dark matter haloes which grow hierarchically via mergers and accretion.

To model galaxy formation one must first have an accurate model of the evolution of dark matter haloes. Hierarchical galaxy formation models require three basic pieces of information about dark matter haloes: (i) The abundance of haloes of different masses. (ii) The formation history of each halo, commonly called the merger tree. (iii) The internal structure of the halo, in terms of the radial density and their angular momentum profiles. With the improvement of N-body simulations in recent decades, the formation and evolution of dark mat-

ter haloes from cosmological initial conditions in large representative volumes can now be routinely and reliably simulated. In order to have a comprehensive understanding of galaxy formation, we can start by modelling galaxy formation by providing an accurate model of the evolution of dark matter haloes. The key starting point for this approach is halo merger trees which quantify the hierarchical growth of individual dark matter haloes. By generating N-body merger trees, we can build a halo catalogue within which we can perform detailed study of the internal structure of the haloes.

1.1 ΛCDM cosmology

Zwicky (1933) observed that galaxies in the Coma cluster seemed to be moving too rapidly to be held together by the gravitational attraction of the visible matter. This was the first evidence of dark matter. Later, Rubin et al. (1980) proposed that dark matter accounts for the missing mass suggested by the dynamics of the luminous matter in individual spiral galaxies.

The scale factor of the Universe $a(t)$ is very crucial in Cosmology, as it measures the universal expansion rate, the evolution of the $a(t)$ is described by the Friedmann equation:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \tag{1.1.1}$$

where k is the curvature and ρ is the density of the Universe (Friedmann, 1922). The Hubble parameter is defined as $H = \dot{a}/a$, and so the Friedmann equation can be written as,

$$H^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}. \tag{1.1.2}$$

For a given value of H, there is a special value of the density which would be required in order to make the Universe flat, $k = 0$, it is known as the critical density ρ_c, which is given by

$$\rho_c(t) = \frac{3H^2}{8\pi G}. \tag{1.1.3}$$

Big Bang Nucleosynthesis theory (Gamow, 1948) predicts that, in terms of the present day critical density of matter, the required density of baryons is a few percent (the exact value depends on the assumed value of the Hubble constant). This relatively low value means that not all of the dark matter can be baryonic, i.e. we are forced to consider more exotic particle candidates.

It is conventional to express the baryon density, ρ_b, the total mass density, ρ_m, and even the dark energy density, $\rho_{\rm DE}c^2$, postulated to be responsible for the accelerating expansion of the universe (Perkins, 2008) in units of the critical density.

The respective dimensionless density parameters for the baryon density, $\Omega_{\rm b,0}$, the total mass density, $\Omega_{\rm m,0}$ and dark energy density, $\Omega_{\Lambda,0}$ are given by

$$\Omega_{i,t} \equiv \frac{\rho_{i,t}}{\rho_c}. \tag{1.1.4}$$

Since the discovery of the cosmic expansion acceleration in 1998 (Riess et al., 1998; Perlmutter et al., 1999), one of the most debated questions in physics and cosmology has been the existence and nature of the dark energy. In this model (Riess et al., 1998), it is the negative pressure of the elusive dark energy which drives the accelerating expansion. In the simplest case this corresponds to a cosmological constant, perhaps associated to the zero point energy of the quantum vacuum. On the assumption of homogeneity and isotropy on large scales, the Friedmann-Robertson-Walker (FRW) metric is given by

$$ds^2 = -c^2dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + sin^2\theta d\phi^2)\right], \tag{1.1.5}$$

In equation 1.1.1, k can have positive, zero, negative curvature with respect to the spatial hypersurfaces (Kolb & Turner, 1990). The evolution of the scale factor, a, the expansion of the universe, is determined by the Friedmann equation:

$$H^2(a) = H_0^2\left[\Omega_\Lambda + \Omega_m a^{-3} + \Omega_r a^{-4} - (\Omega_{tot} - 1)^2 a^{-2}\right], \tag{1.1.6}$$

where a is conventionally taken to be equal to 1 at the day, the redshift, z, is given by $a = 1/(1+z)$ and H_0 is the present value of the Hubble parameter.

The 'Λ' in ΛCDM represents Einstein's cosmological constant, which also refers to the dark energy. By splitting the right hand side of the acceleration equation into parts for the vacuum and for all other material, we can have:

$$\left(\frac{\ddot{a}}{a}\right)^2 = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right) + \frac{\Lambda}{3}, \tag{1.1.7}$$

where $\Lambda \equiv 8\pi G\rho_{DE}$ and ρ_{DE} is the effective density of the vacuum defined such that the energy density is $\epsilon_{DE} = \rho_{DE}c^2$.

Using the same method, we can find that Λ enters the Friedmann equation as

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda}{3}. \tag{1.1.8}$$

The Λ dominates as a gets very large and results in exponential expansion $a \propto \exp(Ht)$ where $H = \sqrt{\Lambda/3}$. As $P_{DE} = -\rho_{DE}c^2$, the negative pressure exerted from the positive energy density can accelerate the expansion of the Universe, in contrast to matter, which slows down the expansion. Although the acceleration was detected through observations of supernovae (Riess et al., 1998; Schmidt et al., 1998; Perlmutter et al., 1999), the existence of dark energy still needs more evidence to be confirmed. CMB (Cosmic microwave background) radiation is produced after $t_{rec} = 400000$ years when the Universe has cooled sufficiently that there are no longer sufficiently high energy photons to ionize hydrogen. The proton-electron plasma recombines into hydrogen atoms and makes the universe transparent.

The CMB (Spergel et al., 2003, 2007; Komatsu et al., 2009, 2011) can constrain the geometry of the Universe to be close to flat which strengthens the evidence for dark energy as this requires $\Omega_{\rm tot} = 1$. The baryon acoustic peaks are particularly important and they are produced because baryons are tightly coupled to the photons before recombination. The pressure in the photon-baryon fluid sets up oscillations in the fluid by providing a restoring force to the initial fluctuations. Therefore, there are ripples of all wavelengths and any perturbation in the baryon-photon plasma thus behaves as an acoustic wave. The sound horizon at which baryons were released from the Compton drag of photons determines the location of baryon acoustic oscillations. The epoch, called the drag epoch, occurs at the redshift z_d (Eisenstein & Hu, 1998) and then the sound horizon is given by $r_s(z_d) = \int_0^{\eta_d} d1990s\eta c_s(\eta)$, where c_s is the sound speed. By measuring the angular scale of these fluctuations, astronomers (Eisenstein et al., 2005; Per-

cival et al., 2007, 2010) have continued to confirm that about 70% of the energy density of the present Universe consists of dark energy.

1.1.1 Galaxy clustering and galaxy clusters

The clustering of galaxies is still a hot topic in structure formation, two of the most famous modern surveys are 2dF galaxy redshift survey (Colless et al., 2001) and Sloan Digital Sky Survey (Strauss et al., 1999). The current understanding of galaxy clustering is in terms of being driven by the dark matter in the Universe. As the dark matter is the dominant matter and it interacts gravitationally with baryons, galaxies generally follow the dark matter distribution, and the gravitational attraction of the dark matter causes the galaxies to clump together.

The two most widely used models of the galaxy-halo connection are the Halo Occupation Distribution (HOD, e.g., Berlind & Weinberg, 2002; Zheng et al., 2005) and the Conditional Luminosity Function (CLF, e.g., Yang et al., 2003; van den Bosch et al., 2013). Both the HOD and the CLF assume that galaxy occupation statistics are governed exclusively by the masses of the dark matter halos hosting the galaxies of interest. Simple models of this form have been sucessful in accurately reproducing the observed galaxy clustering including its dependence on galaxy luminosity and colour (Norberg et al., 2002; Zehavi et al., 2011).

The most massive dark matter halos can contain thousands of galaxies. These galaxy clusters are the largest relaxed structures in the Universe. They have provided very strong and direct evidence for the nature of dark matter and played a very important role as cosmological probes. Clusters of galaxies are believed to form hierarchically, with small systems forming first which are then pulled together by gravity to form larger clusters. When these mergers take place the diffuse gas involved is processed through hydrodynamical shocks which converts its infalling kinetic energy to thermal energy. This establishes a hot pressure supported intra-cluster medium that emits at X-ray wavelengths and can be detected via the Sunyaev-Zel'dovich (SZ) effect through its imprint on the cosmic microwave background radiation (Grego et al., 2000). Cosmological pa-

rameters can be constrained by measuring the number density of galaxy clusters as a function of X-ray luminosity, temperature or SZ decrement (Molnar et al., 2002; Planck Collaboration et al., 2014a).

Measurements of gravitational lensing (the bending of light paths from distant sources caused by the gravitational field of all the intervening mass), can also provide an accurate way of measuring the dark matter distribution in galaxy clusters in order to test the current theoretical predictions (e.g. Velander et al., 2014a). The dynamical nature of mergers can also provide very direct tests of basic gravitational physics and the location of dark matter can be probed with weak gravitational lensing observations (Mandelbaum et al., 2013; Cacciato et al., 2013; Velander et al., 2014b; von der Linden et al., 2014). The observational studies of galaxy clusters are increasingly multi-wavelength (IR, X-ray etc) (Pipino & Pierpaoli, 2010; Israel et al., 2014), the astronomers are now focusing on using existing Herschel and Planck data to analyse lensing galaxy clusters, high redshift clusters, and cluster mergers.

1.1.2 Dark Matter Candidates

All of the evidence for dark matter noted above is based on its gravitational interactions, but it does little to point out what dark matter is. I will focus on dark matter candidates that are motivated not only by cosmology but also by particle physics in this section.

The standard model (SM) of particle physics is a successful theory of elementary particles (e.g. Herrero, 1998). The open questions raised by the SM have helped to encourage the hypothesis of many of the leading dark matter candidates. The neutrino is the only SM particle that could be DM and it is known to have a mass of at least meV due to neutrino oscillations (Gonzalez-Garcia & Nir, 2003). We also need to consider dark matter candidates beyond the SM. Weakly-interacting massive particles (WIMPs) are the most studied none SM dark matter candidates, as they are predicted to have the correct relic density and the WIMPs, like neutralino and Kaluza-Klein dark matter, can be detected in many ways. If WIMPs do not contribute a significant amount of the current

dark matter density, it implies that they must have annihilated to other particles because, if they exist, they initially have the observed relic density. Besides the neutralinoes and KK dark matter, there have been many other WIMP candidates like T-odd particles in little Higgs theories (Birkedal et al., 2006). All the WIMPs are produced through thermal freeze out and are cold, collisionless. Alternative new physics beyond the SM suggests a sterile neutrino is a intermediate mass, keV, dark matter candidate (Dodelson & Widrow, 1994).

There are three scenarios depending on the mass of the DM particle which determines the thermal velocities of the particles when they decouple. Neutrinos if they had masses of eV could make up all the DM and would be a candidate for hot dark matter (HDM). Here the large thermal velocities of the light particles cause them to erase, vie free streaming, all the initial structure on scales less than that of galaxy clusters. But the CMB data indicates that only a small amount of HDM could exist at the CMB time (Gariazzo et al., 2013), as a dominant fraction is not simultaneously compatible with the amplitude of fluctuations in the CMB and high redshift structure formation as evidenced by the existence of high redshift quasars. This scenario has been ruled out. WIMPS are a leading cold dark matter (CDM) candidate. They are cold and leave the primordial spectrum of fluctuations intact down to Earth masses. In this scenario structure formation begins at high redshift and is hierarchical with structure forming through repeated mergers to low redshift. Sterile neutrinos are candidates for warm dark matter (WDM), which is intermediate between hot and cold dark matter. Here modest thermal velocities only erase structure on scales comparable to those of dwarf galaxies. Their effect could be significant on the internal structure of galactic haloes and their dwarf satellites. Hence there is considerable literature on the effect of WDM on the so called cusp vs. core, too-big-to-fail and missing satellites problems (see Weinberg et al., 2013, for a review). These issues have also prompted the discussion of other classes of dark matter. For instance, if dark matter and active neutrinos interact with a new gauge boson with mass of around an meV, this can be a solution to these issues (van den Aarssen et al., 2012). All of the small cosmological scales problems can be resolved by this

proposal. With a self-scattering cross section of around 1 cm^2/g, dark matter scattering off the neutrino background leads to its kinetic decoupling at a temperature of about a few eV which is equates to the virial temperature of dwarf galaxies with halo masses of around $10^9 M_\odot$.

In this thesis, I concentrate on the ΛCDM model (although similar analysis could be applied to WDM models if provided with suitable WDM simulations) for the following reasons. All the evidence from the CMB and large scale structure is consistent with the simple ΛCDM model (Planck Collaboration et al., 2014b) and the weak deviations that do exist on small scale could be caused by baryonic effects (Sawala et al., 2014).

1.2 Semi-analytic modelling of galaxy formation

The picture of galaxy formation in ΛCDM involves gas collecting and cooling in dark matter haloes which then lead to star formation. Over the past decade, astronomers have developed a new approach to follow the formation of galaxies in cosmologies in which structure grows hierarchically. The technique is called semi-analytic modelling, allowing a wide range of properties to be predicted for the galaxy populations at any redshift. Semi-analytic modelling is complementary to gas dynamics simulations. The recipes used in semi-analytic models can be refined to mimic the simulations. Where simulations break down due to a lack of resolution or understanding of the relevant physics, semi-analytical models can be used to improve our knowledge and extend the modelling.

The first semi-analytic models were constructed in the early 1990s (Cole, 1991; White & Frenk, 1991) and have become steadily more sophisticated (Kauffmann & White, 1993; Somerville & Primack, 1999; Cole et al., 2000; Bower et al., 2006; Somerville et al., 2008; Benson & Bower, 2010). Specifying the semi-analytic model of galaxy formation typically consists of the following steps:

1. Choose the cosmological parameters that specify the geometry and matter content of the universe ($\Omega_{m,0}$, $\Omega_{b,0}$, $\Omega_{\Lambda,0}$, H_0) and the amplitude at 8 h^{-1}Mpc of the initial density fluctuation spectrum σ_8.

2. Use N-body simulations or Monte-Carlo methods to trace the merger histories for a series of dark matter haloes.

3. Use the merger histories of dark matter haloes to follow the distinct baryonic components: hot gas, cold gas disc, and stars.

4. Provide compact prescriptions to specify these components: the cooling of hot gas, star formation from the cold gas, feedback from massive stars and active galactic nucleus (AGNs) reheating the cold gas to hot gas. Also, the metallicity of each of is tracked using chemical evolution models.

1.3 Formation and structure of dark matter haloes

Density perturbations grow linearly until they reach an amplitude of order unity, δ_c, then they turn around from the expansion of the Universe and collapse to form virialized dark matter haloes. These haloes grow in mass by accreting material from their neighbourhood or by merging with other haloes. Some of these haloes survive as bound structures after merging with larger haloes and become subhaloes within the larger halo. In summary, the picture of the halo formation is that: small-scale perturbations grow and collapse to form small haloes, and these small haloes merge together to form a single virialized dark matter halo containing remnant subhaloes.

Dark matter haloes are the hosts of galaxies, the halo properties will have a direct link to the mass function, progenitor mass function, merger rate, clustering properties and internal properties of galaxies. Thus, it is essential to understand the structure and formation of dark matter haloes as their properties will help us to understand the formation and evolution of galaxies. In this section, I will discuss the following topics: The mass function of dark matter haloes, the halo merger trees, and finally the internal structure of dark matter haloes.

1.3.1 Halo mass function

Consider the overdensity field $\delta(x,t) = (\rho(x,t) - \bar{\rho}(t))/\bar{\rho}(t)$, which evolves as $\delta(x,t) = \delta_0(x)D(t)$ in the linear regime, where δ_0 is the overdensity field linearly extrapolated to the present time and $D(t)$ is the linear growth rate at time t, normalized to unity at the present (Kolb & Turner, 1990). In the spherical collapse model, the regions with $\delta_0(x) > \delta_c/D(t)$, will collapse to form virialized objects. The initial work of predicting the mass function of collapsed objects can be traced back as early as four decades ago, when Press & Schechter (1974) presented the famous Press & Schechter (PS) mass function to calculate the number density of collapsed objects with masses in the range dM given by

$$n(M,t)dM = \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M^2}\frac{\delta_c}{\sigma}\exp\left(-\frac{\delta_c^2}{2\sigma^2}\right)\left|\frac{d\ln\sigma}{d\ln M}\right|dM. \tag{1.3.9}$$

The PS mass function can be written in the form:

$$n(M,t)dM = \frac{\bar{\rho}}{M^2}f_{\rm PS}(\sigma)\left|\frac{d\ln\sigma}{d\ln M}\right|dM, \tag{1.3.10}$$

where

$$f_{\rm PS}(\sigma) = \sqrt{\frac{2}{\pi}}\frac{\delta_c}{\sigma}\exp\left(-\frac{\sigma_c^2}{2\sigma^2}\right). \tag{1.3.11}$$

In the PS formalism, all mass is contained in haloes:

$$\int_{-\infty}^{\infty} f_{\rm PS}(\sigma)\mathrm{d}\ln\sigma^{-1} = 1. \tag{1.3.12}$$

The S-T model is a modification to the PS model motivated by ellipsoidal halo collapse models and fitted to the results of N-body simulations (Sheth et al., 2001; Sheth & Tormen, 2002)

$$f(\sigma;\mathrm{S-T}) = A\sqrt{\frac{2a}{\pi}}\left[1+(\frac{\sigma^2}{a\delta_c^2})^p\right]\frac{\delta_c}{\sigma}\exp\left[-\frac{a\delta_c^2}{2\sigma^2}\right], \tag{1.3.13}$$

merger tree where A=0.3222, $a = 0.707$ and $p = 0.3$.

Besides the analytic formulas, the topic of the mass function has been investigated by groups working with N-body simulations. I list some famous formulas starting with Jenkins et al. (2001) offer an empirical fit using high resolution simulations of a range of cosmologies. Their fit is constructed in the $f - \ln(\sigma^{-1})$ plane,

$$f(\ln\sigma^{-1}) = 0.315\ \exp\left[-|\ln\sigma^{-1}+0.61|^{3.8}\right]. \tag{1.3.14}$$

Warren et al. (2006) provide a fitting formula using dark matter simulation,

$$f(\sigma) = A(\sigma^{-a} + b)e^{-c/\sigma^2} \tag{1.3.15}$$

with parameters $\mathbf{q} = (A, a, b, c)$. Reed et al. (2007) modifies this to

$$f(\sigma) = A\sqrt{\frac{2a}{\pi}}\left[1 + (\frac{\sigma^2}{a\delta_c^2})^p + 0.2G_1\right]\frac{\delta_c}{\sigma}\exp\left[-\frac{ca\delta_c^2}{2\sigma^2}\right] \tag{1.3.16}$$
$$G_1 = \exp\left[-\frac{[\ln(\sigma^{-1})-0.4]^2}{2(0.6)^2}\right],$$

by steepening the high mass slope of the S-T function with the addition of a new parameter, $c = 1.08$, in the exponential term, and simultaneously including a Gaussian in $\ln\sigma^{-1}$ centred at $\ln\sigma^{-1} = 0.4$ to fit their simulation data. Watson et al. (2013) provide the fitting formula,

$$f(\sigma) = A\left[\left(\frac{\beta}{\sigma}\right)^{\alpha} + 1\right]e^{-\gamma/\sigma^2}, \tag{1.3.17}$$

where A = 0.282, α = 2.163, β = 1.406, γ = 1.210 which is good to fit the simulation data to within 10%.

1.3.2 Halo merger trees

Halo merger trees play a very important role in hierarchical models of galaxy formation and they are the backbone of the semi-analytic models of galaxy formation. There are main two methods of constructing merger trees: Monte-Carlo methods based on the extended Press-Schechter (EPS; Bond et al. (1991)) formalism and N-body simulations. Different analytic algorithms for Monte-Carlo methods have been developed since the early 90s by Bond et al. (1991); Kauffmann & White (1993); Lacey & Cole (1993); Somerville & Kolatt (1999); Cole et al. (2000); Parkinson et al. (2008). Since the EPS formalism itself is not rigorous and it does not specify many merger tree properties, they require detailed comparison with simulations of structure formation. Trees directly extracted from the N-body outputs can be an alternative to Monte-Carlo merger trees. The main advantage over Monte-Carlo merger trees of N-body merger trees is that, they do not depend on the EPS assumptions and the outputs can provide spatial

and velocity information automatically. Although high-resolution simulations are expensive and time-consuming and there is no unambiguous definition for a halo, the approach is still very powerful due to the improvement of the simulation resolution (Boylan-Kolchin et al., 2009) and halo definitions (Helly et al., 2003; Springel et al., 2005; Harker et al., 2006). In Chapter 2 and Chapter 3, we have provided the detailed prescription of a new advanced halo definition and investigation of the Dhalo merger trees which is a more advanced N-body merger tree than Friends-of-Friends (FoF) (Davis et al., 1985) halo merger trees.

1.3.3 Internal structure of dark matter haloes

In this section, I describe three models for the internal structure of dark matter haloes, which are the singular isothermal sphere, Hernquist (Hernquist, 1990) and the NFW (Navarro et al., 1995) models. Motivated by flat rotation curves, the density profile for the singular isothermal sphere is,

$$\rho(r) \propto \frac{1}{r^2}. \tag{1.3.18}$$

Both Hernquist and NFW density profiles can be described with a double power law density distribution given by

$$\rho(r) \propto \left(\frac{r}{r_s}\right)^{-\gamma} \left[1 + (\frac{r}{r_s})^{\alpha}\right]^{(\gamma-\beta)/\alpha}. \tag{1.3.19}$$

where $\gamma = 1, \beta = 3, \alpha = 1$ for NFW profile, $\gamma = 1, \beta = 4, \alpha = 1$ for Hernquist profile, r_s is a characteristic length scale. At small radii $\rho(r) \propto r^{-\gamma}$, while at large radii $\rho(r) \propto r^{-\beta}$.

I adopt the NFW profile in the rest of this thesis. It is conventionally expressed as,

$$\rho(r) = \rho_{crit} \frac{\delta_{char}}{r/r_s(1 + r/r_s)^2}, \tag{1.3.20}$$

where δ_{char} is a characteristic overdensity and r_s is a characteristic scale . The enclosed mass of the NFW profile is

$$M(r) = 4\pi\bar{\rho}\delta_{\rm char} r_{\rm s}^3 \left[\ln(1 + cx) - \frac{cx}{1 + cx}\right], \tag{1.3.21}$$

where $x \equiv r/r_{\rm h}$, and the halo concentration is defined as $c \equiv r_{\rm h}/r_{\rm s}$. We can obtain the relation between characteristic overdensity and concentration parameters,

$$\delta_{char} = \frac{\Delta_h}{3} \frac{c^3}{\ln(1+c) - c/(1+c)}, \tag{1.3.22}$$

where Δ_h is 200 if $\rho_h = 200 \times \rho_c$. The NFW profile is completely characterized by its mass, M_{Δ_h}, and its concentration parameter, c, or equivalently by r_s and $\delta_{\rm char}$.

1.3.4 Structure of the thesis

The thesis is structured as follows. In Chapter 2 I present a detailed description of the latest N-body merger tree algorithm that has been developed for use with the semi-analytic code GALFORM. In Chapter 3 I compare and contrast the properties of the haloes (Dhaloes) resulting from the N-body merger tree described in Chapter 2 with the more commonly used FoF haloes (Davis et al., 1985). I show specific rare examples where Dhaloes and their matched FoF counterparts exhibit gross differences with either one FoF halo being decomposed into several Dhaloes or vice versa. I also examine the distribution of mass ratios for matching Dhalo and FoF pairs. Then in Chapter 4 we compare statistical properties of halo populations including halo mass functions and their concentration–mass relation. In Chapter 5, I briefly outline the methods including the identification of halo mergers in N-body simulation and the measurement of orbital parameters, then we present detailed analysis of the orbital parameters.

2

DHALO ALGORITHM

In this thesis, we present a detailed description of the latest N-body merger tree algorithm that has been developed for use with the semi-analytic code GALFORM. The algorithm is an improvement over the earlier version, described in Merson et al. (2013), which was run on the Millennium simulation (Springel, 2005a) and widely exploited in a range of applications (Bower et al., 2006; Font et al., 2008; Kim et al., 2011; Merson et al., 2013). Our improvements mainly include the following two aspects, we can identify a descendant more than one snapshot later even if something else has merged with the subhalo between the snapshots. The old algorithm would fail in this case because it required the descendant to have no progenitors at all rather than just no main progenitor. Second, we use a new definition of the main progenitor to find the ”same” object between snapshots when deciding how to group halos into subhaloes. More details are given in Section 2.2. The resulting differences between the two algorithms are very small when applied to relatively low resolution simulations such as the Millennium, but the improvements in the new algorithm do a better job of tracking halo descendants in high resolution simulations such as the Millennium II (Boylan-Kolchin et al., 2009) and Aquarius simulations (Springel et al., 2008). The starting point for our merger trees are FoF haloes that are decomposed into subhaloes, distinct self-bound structures, by the substructure finder, SUBFIND (Springel et al., 2001). Subhaloes are tracked between output times and agglomerated into a new set of haloes, dubbed Dhaloes, that have consistent membership over time in the sense that once a subhalo is accreted by a Dhalo it

never demerges. In this process we also split some FoF haloes into two or more Dhaloes when SUBFIND substructures are well separated and only linked into a single FoF halo by bridges of low density material. most bound particles

Here we describe in detail the algorithm used to produce the Dhalo merger trees developed by Dr John Helly in Section 2.2 and I summarize my post-processing algorithm in Section 2.3. These merger trees are intended to be used as input to the GALFORM semi-analytic model of galaxy formation. The need for consistency between the halo model used in the semi-analytic calculation and the N-body simulation imposes some requirements on the construction of the merger trees.

The GALFORM galaxy formation model makes the approximation that mergers between haloes are instantaneous events and assumes that haloes, once merged, do not fragment. However, in N-body simulations halo mergers take a finite amount of time and it is not uncommon for a halo falling into another, more massive halo to escape to well beyond the virial radius after its initial infall (Gill et al., 2005; Ludlow et al., 2009). We therefore need to choose when to consider N-body haloes to have merged in the semi-analytic model and define our haloes such that they remain merged at all later times. We also wish to define the haloes used to construct the trees such that, as far as possible, they resemble the spherically symmetric, virialised objects assumed in the galaxy formation model. Quantifying the extent to which we have achieved this is one of the main aims of this chapter.

2.1 Halo catalogues

Immediately below, we summarise the specification of the Millennium II simulation which we use to test and illustrate the application of our merger tree algorithm. We then give the details of the construction of the merger trees and their constituent haloes with the complete specification.

The Millennium-II Simulation

The Millennium-II (MSII) simulation[1] (Boylan-Kolchin et al., 2009) was carried out with the GADGET3 N-body code, which uses a "TreePM" method to calculate gravitational forces. The MSII is a cosmological simulation of the standard ΛCDM cosmology in a periodic box of side $L_{\rm box}$= $100h^{-1}$ Mpc containing $N = 2160^3$ particles of mass $6.95 \times 10^6\ h^{-1}{\rm M}_\odot$. The cosmological parameters for the MSII are: $\Omega_{\rm m} = 0.25$, $\Omega_{\rm b} = 0.045$, $h = 0.73$, $\Omega_\Lambda = 0.75$, $n = 1$ and $\sigma_8 = 0.9$. Here $\Omega_{\rm m}$ denotes the total matter density in units of the critical density, $\rho_{\rm crit} = 3H_0^2/(8\pi G)$. $\Omega_{\rm b}$ and Ω_Λ denote the densities of baryons and dark energy at the present day in units of the critical density. The Hubble constant is $H_0 = 100h\,{\rm km\,s^{-1}}$ Mpc^{-1}, n is the primordial spectral index and σ_8 is the rms density fluctuation within a sphere of radius $8h^{-1}$Mpc extrapolated to z = 0 using linear theory. These cosmological parameters are consistent with a combined analysis of the 2dFGRS (Colless et al., 2001; Percival et al., 2001) and first year WMAP data (Spergel et al., 2003; Sánchez et al., 2006).

2.2 Building the subhalo merger trees

Before we can construct the Dhalo merger trees, it is necessary to define subhalo merger trees by identifying the descendant of each subhalo. The code we use to do this was included in the merger trees comparison project carried out by Srisawat et al. (2013) under the name D-Trees. The project concluded that it was desirable feature for a merger tree code to use particle IDs to match haloes between snapshots and have the ability to search multiple snapshots for descendants. The latter requirement was due to the tendency of the AHF group finder (Knollmann & Knebe, 2009) used in the project to temporarily fail to detect substructures during mergers.

Since SUBFIND suffers from a similar problem, we allow for the possibility that

[1]The Millennium-II simulation data will be available from an SQL relational database that can be accessed at http://galaxy-catalogue.dur.ac.uk:8080/Millennium .

the descendant of a subhalo may be found more than one snapshot later. Our approach is to devise an algorithm which can identify the descendant of a halo at any single, later snapshot, apply it to the next $N_{\rm step}$ snapshots (where $N_{\rm step} = 5$), and pick one of these $N_{\rm step}$ possible descendants to use as the descendant of the subhalo in the merger trees.

Alternative solutions to this problem include allowing the merger tree code to modify the subhalo catalogue to ensure consistency of subhalo properties between snapshots (ConsistentTrees, Behroozi et al. 2013) and using information from previous snapshots to define the subhalo cataloguemuch longer (HBT, Han et al. 2012).

In common with all but one of the merger tree codes in the comparison (Jmerge, which relies entirely on aggregate properties of the haloes), we identify descendants by finding subhaloes at differentmuch longer snapshots which have particles in common.

2.2.1 Identifying a descendant at a single, later snapshot

To find the descendant at snapshot j, of a halo which exists at an earlier snapshot, i, the following method is used. For each halo containing N_p particles the $N_{\rm link}$ most bound particles are identified, where $N_{\rm link}$ is given by

$$N_{\rm link} = \min(N_{\rm linkmax}, \max(f_{\rm trace} N_p, N_{\rm linkmin})) \tag{2.2.1}$$

with $N_{\rm linkmin} = 10$, $N_{\rm linkmax} = 100$ and $f_{\rm trace} = 0.1$.

For each of the haloes at snapshot i, descendant candidates are found by locating all haloes at snapshot j which received at least one particle from the earlier halo. Then, a single descendant is chosen from these candidates as follows. If any of the descendant candidates received a larger fraction of their $N_{\rm link}$ most bound particles from the progenitor halo than from any other halo at the earlier snapshot, then the descendant is chosen from these candidates only and the halo at snapshot i will be designated the main progenitor of the chosen descendant; otherwise, all candidates are considered and the halo will not be the main progenitor of its descendant. The descendant of the halo at snapshot i is taken to

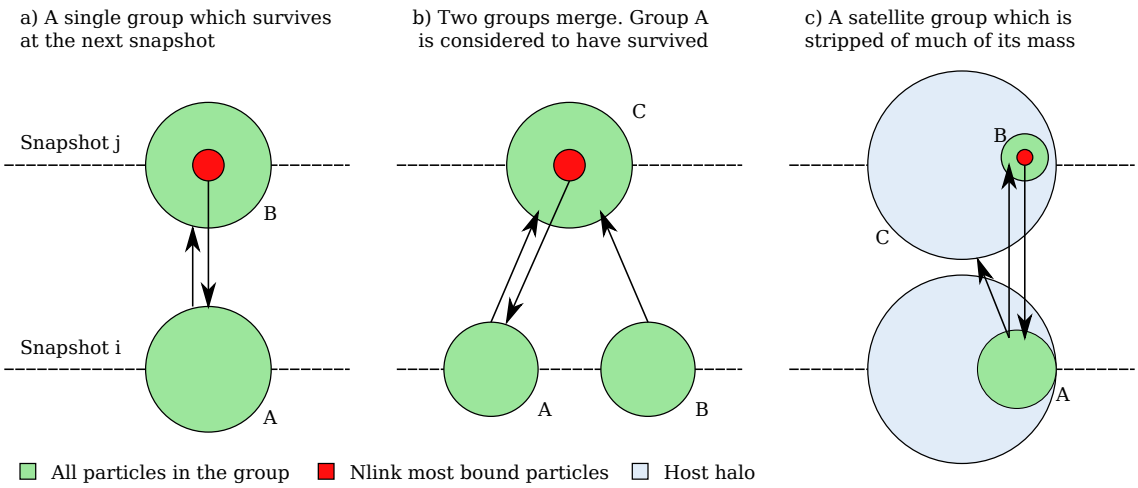

Figure 2.1: Schematic examples illustrating the method used to link SUBFIND subhaloes between pairs of snapshots i and j, where $i < j$. The green circles represent SUBFIND subhaloes. The most bound particles $N_{\rm link}$ particles in each subhalo at the later time are shown in red. From left to right are a) a single, isolated subhalo which still exists at the next snapshot, b) a merger between subhaloes A and B where more of the most bound partciles of the merged halo C come from halo A than from any other halo and therefore halo A is considered to be the main progenitor of halo C, and c) a satellite subhalo orbiting within a background halo which loses a large fraction of its particles to its host halo at the next snapshot but is still identified by one or more snapshots SUBFIND. Arrows between green circles show the location of the majority of the particles in the subhalo at the later snapshot. Arrows starting from red circles show the location of the majority of the most bound particles at the earlier snapshot.

be the remaining candidate which received the largest fraction of the $N_{\rm link}$ most bound of the progenitor halo. For each halo at snapshot j, this method identifies zero or more progenitors of which at most one may be a main progenitor. Note that it is not guaranteed that a main progenitor will be found for every halo.

By following the most bound part of the subhalo, we ensure that if the core of a subhalo survives at the later snapshot it is identified as the descendant irrespective of how much mass has been lost. It also means that in cases where an object at the later snapshot has multiple progenitors we can determine which one of the progenitors contributed the largest fraction of the most bound core of the descendant object. We consider this main progenitor to have survived the merger while the other progenitors have merged onto it and ceased to exist as independent objects.

Fig. 2.1 shows three examples of this linking procedure. In the simplest case (left) a single, isolated subhalo B at snapshot j is identified as the descendant of subhalo A which exists at the earlier snapshot i. Since more of the most bound particles of subhalo B come from subhalo A than from any other subhalo, we conclude that A is the main progenitor of B. In the second case (centre) two subhaloes A and B merge to form subhalo C at the later snapshot. Subhalo A is determined to be the main progenitor because it contributed the largest fraction of the most bound particles of the descendant, C. In the third example (right) a satellite subhalo A exists within a more massive host halo. In this case, particles from the subhalo A are split between subhalo B and the host halo C at the later snapshot. While a large fraction (or even the vast majority) of the particles from subhalo A may belong to the host halo at the later snapshot, we choose subhalo B as the descendant because its most bound part came from subhalo A.

Searching multiple snapshots for descendants

If a subhalo is not found to be the main progenitor of its descendant, this may indicate that the subhalo has merged with another subhalo and no longer exists as an independent object. However, it is also possible that the substructure finder has simply failed to identify the object at the later snapshot because it is

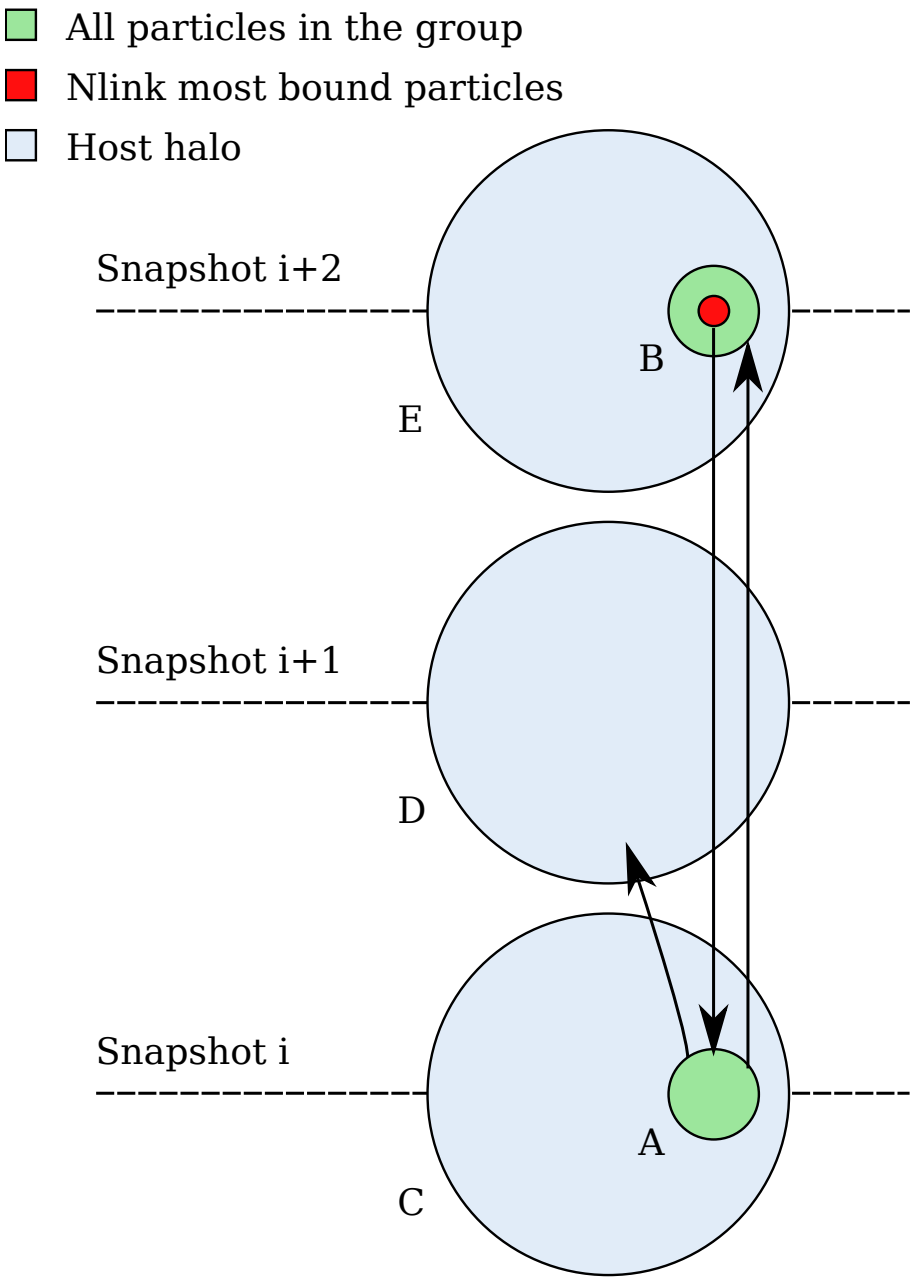

Figure 2.2: A schematic example of a case where the descendant of a subhalo is found to be more than one snapshot later. The green circles represent a satellite SUBFIND subhalo within a larger host halo which is represented by the blue circles. Three consecutive snapshots are shown.

superimposed on the dense central parts of a larger subhalo. Typically this phase lasts for a small fraction of the host halo dynamical time (Behroozi et al., 2013) which in turn is much longer than the usual interval between the snapshots of cosmological N-body simulations. Hence by looking one snapshot ahead we will normally find the missed subhalo, but one can be unlucky and catch it half an orbit later when again it is hidden by the dense core of the more massive subhalo in which it is orbiting. Hence looking several snapshots ahead exponentially suppreses this possibility. Thus in order to distinguish between subhalo mergers and subhaloes which are just temporarily lost it is necessary to search multiple snapshots for descendants.

In our algorithm for each snapshot i in the simulation descendants are identified at later snapshots in the range $i+1$ to $i+N_{\rm step}$ using the method described in section 2.2.1. For each subhalo at snapshot i this gives up to $N_{\rm step}$ possible descendants. One of these descendants is picked for use in the merger trees as follows: if the subhalo at snapshot i is the main progenitor of one or more of the descendants, the earliest of these descendants which does not have a main progenitor at a snapshot later than i is chosen. If no such descendant exists, the earliest descendant found is chosen irrespective of main progenitor status.

Descendants more than one snapshot later are only chosen in cases where the earlier subhalo is the main progenitor — i.e. where the group still survives as an independent object. If the subhalo does not survive we have no way to determine whether it merged immediately or if SUBFIND failed to detect it for one or more snapshots prior to the merger, so we simply assume that the merger happened between snapshots i and $i+1$.

Fig. 2.2 shows a case where a descendant more than one snapshot later is chosen. Subhalo A exists at snapshot i. Its descendant at snapshot $i+1$ is found to be the subhalo D. However, the most bound particles of D were not contributed by subhalo A, but by another progenitor, subhalo C. This means that A is not the main progenitor of its descendant at snapshot $i+1$ and so it is necessary to consider possible descendants at later snapshots. Two subhaloes at snapshot $i+2$ (B and E) receive particles from subhalo A. Since the most bound

particles of subhalo B came from subhalo A, A is the main progenitor of B and subhalo B is taken to be the descendant of A.

2.2.2 Constructing a halo catalogue

At this point we have a descendant for each subhalo. This is sufficient to define merger trees for the subhaloes. These SUBFIND trees can be split into "branches" as follows. A new branch begins whenever a new subhalo forms (i.e. the subhalo has no progenitors). The remaining subhaloes that make up the branch are found by following the descendant pointers until either a subhalo is reached that is not the main progenitor of its descendant, a subhalo is reached that has no descendant, or the final snapshot of the simulation is reached. Each of these branches represents the life-time of an independent halo or sub-halo in the simulation. We construct haloes and halo merger trees by grouping together these branches of the subhalo merger trees using methods which will be described below. We refer to the resulting collections of subhaloes as "Dhaloes". Fig. 2.3 is an example of a Dhalo merger tree with the subhalo merger tree branches marked. In this case there are three branches. Branch A is a single, massive halo which exists as an independent halo at all four snapshots. Branch B is a smaller halo which becomes a satellite subhalo within halo A, but continues to exist. Branch C is another small halo which briefly becomes a satellite before merging with A.

For each subhalo in a FoF halo we identify the hierarchy of subhaloes in which it is embedded. To achieve this we identify for each subhalo the least massive of all the more massive subhaloes that it is enclosed within. Subhalo A is said to enclose subhalo B if B's centre lies within twice the half mass radius of A. A pointer to the enclosing subhalo is stored for each subhalo that is enclosed. This produces a tree structure which is intended to represent the hierarchy of haloes, sub-haloes, sub-sub-haloes etc. in the FoF halo. Any subhalo which is not enclosed by any other becomes a new Dhalo. Any subhaloes enclosed by this subhalo are assigned to the new Dhalo.

We then iterate through the snapshots from high redshift to low redshift. For each subhalo we find the maximum number of particles it ever contained while

it was the most massive subhalo in its parent FoF halo. If a satellite subhalo in a Dhalo retains a fraction $f_{\rm split}$ of its maximum isolated mass then it is split from its parent Dhalo and becomes a new Dhalo. Any subhaloes enclosed by this subhalo are assigned to the new Dhalo too. We usually set $f_{\rm split} = 0.75$, so that when a halo falls into another, more massive halo the two haloes will only be considered to have merged into one once the smaller halo has been stripped of some of its mass. This is to ensure that haloes artificially linked by the FoF algorithm are still treated as separate objects.

In some cases a subhalo may escape from its parent halo. This happens to halo *B* in Fig. 2.3. For the purposes of semi-analytic galaxy formation modelling, we would like to continue to treat such subhaloes as satellites in the parent halo so that each in-falling halo contributes a single branch to the halo merger tree. This is done by merging such objects back on to the Dhalo they escaped from; the subhalo is recorded as a satellite within the original Dhalo at all later times regardless of its spatial position. Any subhaloes it encloses will also be considered to be part of this Dhalo.

In practice the re-merging is carried out in the following way. For each Dhalo *A* we identify a descendant Dhalo *B* by determining which later Dhalo contains the descendant of the most massive subhalo in *A* which survives at the next snapshot. In every case where a subhalo in *A* survives, we assign the descendant of the subhalo to Dhalo *B*. We repeat this process for all Dhaloes at each snapshot in decreasing order of redshift. This ensures that if any two subhaloes are in the same Dhalo at one snapshot, and both survive at the next snapshot, they will both be in the same Dhalo at the next snapshot.

This process produces a Dhalo catalogue for each snapshot. Each Dhalo contains one or more subhaloes and each subhalo may have a pointer to a descendant at some later snapshot. Any subhaloes in a Dhalo which survive at the next snapshot are guaranteed to belong to the same Dhalo at the next snapshot. This provides a simple way to identify a descendant for each Dhalo and defines the Dhalo merger trees. Fig. 2.3 shows an example of a Dhalo merger tree. The two smaller haloes *B* and *C* merge with a larger halo *A*. Halo *C* survives as a

satellite for one snapshot before merging with the descendant of A. Halo B also becomes a satellite sub-halo and then temporarily escapes from the parent halo before falling back in. At all times after the initial infall it is considered to be part of the parent Dhalo.

2.3 The post-processing algorithm

In the FoF halo catalogue, we have the index Dhaloid for each subhalo in the FoF halo. I start with this index and build a catalogue of Dhaloes by combining the subhaloes with the same Dhaloid. At this stage, I have built a new Dhalo group catalogue which contains quantities I have computed for each group including, "Dhalo masses", $M_{\rm halo}$, which are the sum of all the masses of all subhaloes belonging to the same Dhalo, "Length", which refers to the number of particles the halo, "Dhalo centre", which is the potential minimum of the most massive subhalo contained in the halo and all the information from each subhalo, now re-ordered according to their host Dhalo. Then, I run my codes using the centres of Dhalo through all the snapshots in MSII to calculate virial mass and virial radius for each Dhalo. This results in a complete Dhalo group catalogue.

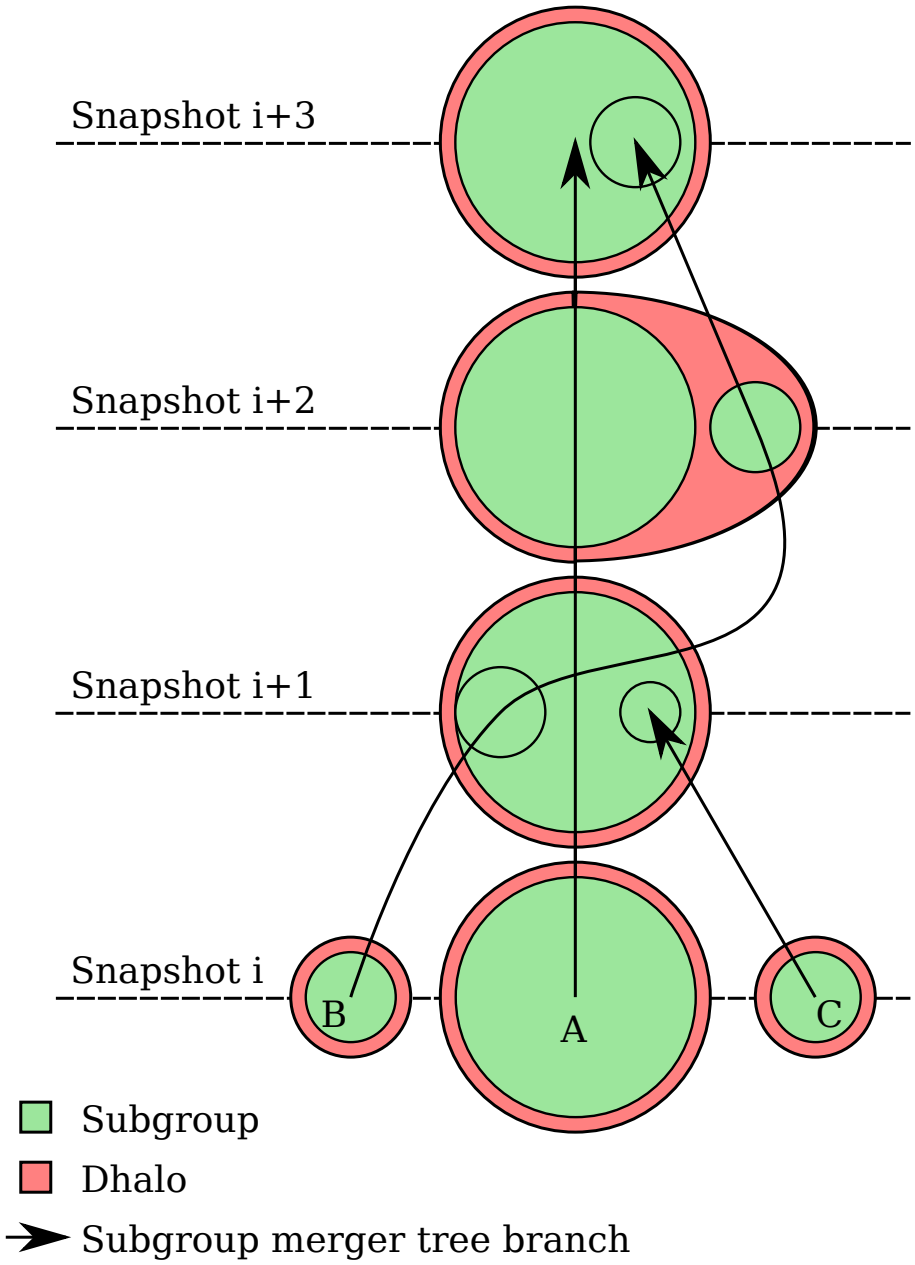

Figure 2.3: An example of a Dhalo merger tree showing two less massive haloes falling into another, more massive halo. Subhaloes are shown in green. Red areas indicate subhaloes which belong to the same Dhalo. The black arrows show branches of the subhalo merger tree.

3 COMPARISON OF FOF AND DHALOES

3.1 Introduction

The properties of FoF haloes, especially those defined by the conventional linking length parameter of $b = 0.2$ (the linking length is defined as b times the mean inter-particle separation), are well documented in the literature (e.g. Frenk et al., 1988; Lacey & Cole, 1994; Summers et al., 1995; Audit et al., 1998; Huchra & Geller, 1982; Press & Davis, 1982; Einasto et al., 1984; Davis et al., 1985; Klypin et al., 1999; Jenkins et al., 2001; Warren et al., 2006; Eke et al., 2004; Gottlöber & Yepes, 2007) and such haloes are widely used as the starting point for relating the dark matter and galaxy distributions (Peacock & Smith, 2000; Seljak, 2000; Berlind & Weinberg, 2002). Thus as the semi-analytic model GALFORM (Bower et al., 2006; Font et al., 2008, 2011; Lagos et al., 2011) instead uses Dhaloes as its starting point, it is interesting to contrast the properties of haloes defined by these two algorithms.

As described in Section 2.1, FoF haloes are decomposed by SUBFIND into subhaloes and those are then regrouped into Dhaloes. Hence for every FoF halo, we can find its matching Dhalo by finding which Dhalo contains the most massive subhalo from the FoF group. We can perform this matching the other way round by finding the FoF halo containing the most massive subhalo from the Dhalo. In cases where the most massive subhalo of a FoF halo is also the most massive subhalo of a Dhalo, these two matching procedures produce identical associations. We refer to such cases as bijective matches.

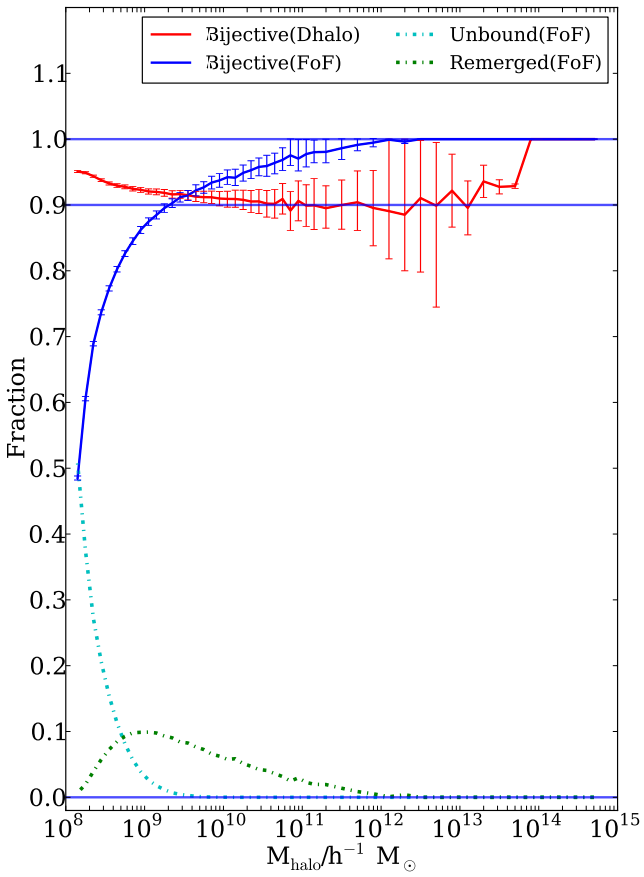

Figure 3.1: The upper two curves, with bootstrap error bars, show the fraction of Dhalo (red) and FoF haloes (blue) in the MSII catalogues that have a bijective (a unique one-to-one) match as a function of their respective Dhalo or FoF halo mass. The lower two curves show the fraction of FoF haloes that do not contain a self-bound substructure (cyan) and the fraction whose main subhaloes are remerged by the Dhalo algorithm to form part of a more massive Dhalo (green).

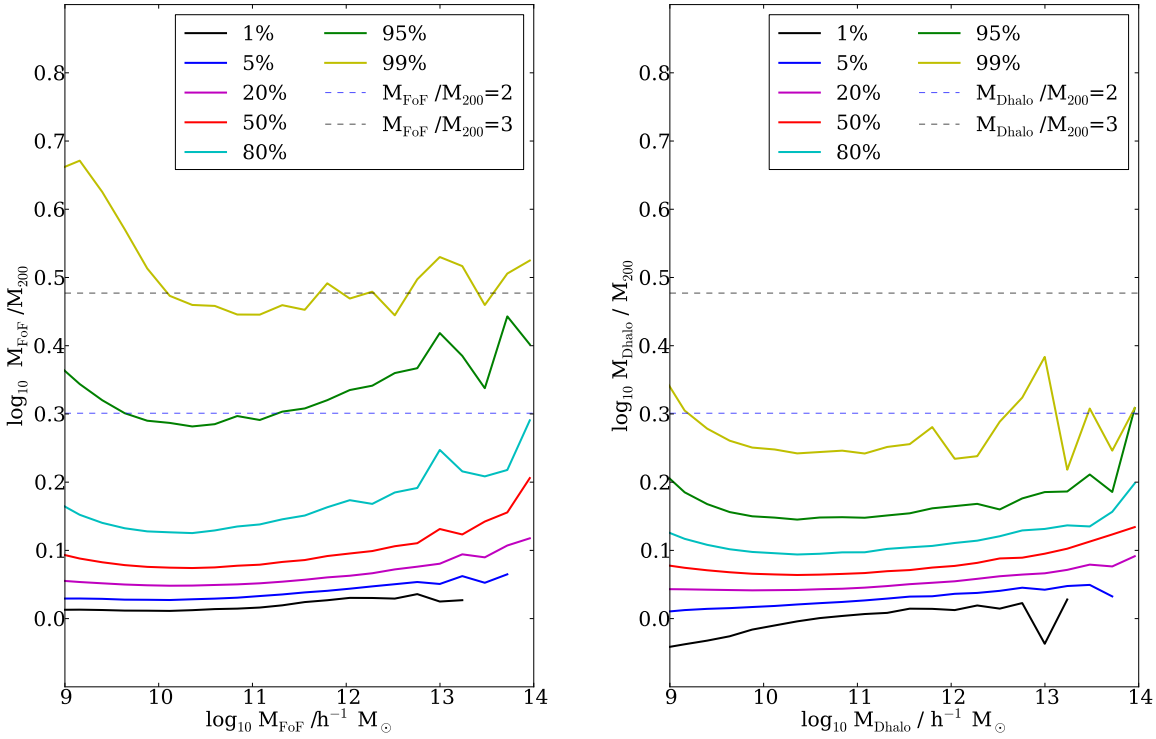

Figure 3.2: The left panel shows the median, 1, 5, 20, 80, 95, 99 percentiles of the distribution of the mass ratios between FoF halo mass, $M_{\rm FoF}$ and virial mass, M_{200} as a function of FoF halo mass for haloes identified using the FoF group finder. The right panel shows the same percentiles for the distribution of the mass ratio between Dhalo mass, $M_{\rm Dhalo}$ and virial mass, M_{200}, as a function of Dhalo mass for haloes identified using the Dhalo group finder. The blue dashed line in both panels shows where $M_{\rm Halo}/M_{200}$ =2.0 and the black one $M_{\rm Halo}/M_{200}$=3.0.

3.2 Bijectively matched FoF and Dhaloes

Before comparing the properties of this subset of bijectively matched Dhaloes and FoF haloes we first quantify how representative they are by looking at the fraction of each set of haloes that have these bijective matches. The two upper curves in Fig. 3.1 show the dependence of the bijective fraction of Dhaloes on Dhalo mass and FoF haloes on FoF mass. The first thing to note is that the fraction of bijectively matched Dhaloes is large, being 90% or greater over the full range from 10^8 to 10^{14} $h^{-1}\mathrm{M}_\odot$ and so to a first approximation there is a good correspondence between FoF and Dhaloes. Above 3×10^{10} $h^{-1}\mathrm{M}_\odot$ about 10% of the Dhaloes do not have a bijective match which means they instead represent secondary fragments of more massive FoF haloes that the Dhalo algorithm has split into two or more subhaloes. Below 3×10^{10} $h^{-1}\mathrm{M}_\odot$ this non-bijective fraction drops indicating that lower mass FoF haloes are less likely to be split into two or more comparable mass Dhaloes. This behaviour is consistent with the results of Lukić et al. (2009) who found that 15-20% of FoF haloes are irregular structures that have two or more major components linked together by low density bridges and that this fraction is an increasing function of halo mass. This is also to be expected in the hierarchical merging picture as the most massive haloes formed most recently and so are the least dynamically relaxed.

For the FoF haloes with mass above 10^{12} $h^{-1}\mathrm{M}_\odot$ the bijectively matched fraction is unity, indicating that the most massive subhalo of such FoF haloes together with the subhaloes embedded within it always gives rise to a Dhalo. Below 10^{12} $h^{-1}\mathrm{M}_\odot$ the the bijective fraction begins to decrease steadily with decreasing mass. This happens because as the FoF mass decreases there is an increasing probability that the progenitor of this FoF halo has previously passed through a more massive neigbouring halo and this results in the Dhalo algorithm remerging the FoF halo with its more massive neighbour. This fraction of FoF haloes that are remerged to form part of a more massive Dhalo is shown by the green curve in Fig. 3.1. As one approaches 10^8 $h^{-1}\mathrm{M}_\odot$ ($\sim$15 particles) the bijective fraction plummets as at very low masses many of the FoF haloes are not

self-bound and so do not contain any subhalo from which to build a Dhalo. The fraction of FoF haloes which do not contain a self-bound substructure is shown by the cyan curve in Fig. 3.1 and can be seen to reach 50% at a FoF mass of 20 particles.

3.2.1 Virial masses

It is conventional to define the virial mass, $M_{\rm vir}$, and associated virial radius, $r_{\rm vir}$, of a dark matter halo using a simple spherical overdensity criterion centred on the potential minimum.

$$M_{\rm vir} = \frac{4}{3}\pi\Delta\,\rho_{\rm crit}\,r_{\rm vir}^3 \tag{3.2.1}$$

where $\rho_{\rm crit}$ is the cosmological critical density and Δ is the specified overdensity. In applying this definition we adopt $\Delta = 200$ and include all the particles inside this spherical volume, not only the particles grouped by the FoF or Dhalo algorithm, to define the enclosed mass, M_{200} , and associated radius r_{200}. This choice is largely a matter of convention but has been shown to roughly correspond to boundary at which the haloes are in approximate dynamical equilibrium (e.g. Cole & Lacey, 1996).

If the halo finding algorithm has succeeded in partitioning the dark matter distribution into virialized haloes we would expect to see a good correspondence between the grouped mass of the halo and M_{200}. For instance, as FoF haloes are essentially bounded by an isodensity contour, whose value is set by the linking parameter (Davis et al., 1985), then if they have relaxed quasi-spherical configuration a tight relation between $M_{\rm halo}$ and M_{200} is inevitable. The only way $M_{\rm halo} \gg M_{200}$ is if the halo has multiple components which have been spuriously linked together as illustrated in the typical example shown in the lower panels of Fig. 3.4.[1] $M_{\rm halo} \ll M_{200}$ could indicate cases where the group

[1]These grossly non-virialized multi-component systems are not always detected by more often used relaxation criteria (Neto et al., 2007; Power et al., 2012, and see Section 4.3), as such criteria focus on the mass within r_{200} which can be in equillibrium even if diffusely linked to secondary mass concentrations.

finder has split a virialized object into small fragments. Hence it is interesting to look at the distribution of $M_{\rm halo}/M_{200}$ for both the FoF and Dhalo algorithms to simply see how $M_{\rm halo}$ compares to the conventional M_{200} definition of halo mass and to give an indication of the frequency of over linking and fragmentation.

The two panels of Fig. 3.2 quantify the distribution of $M_{\rm halo}/M_{200}$ for both the standard FoF haloes and for haloes defined by the Dhalo algorithm. We immediately see that the distribution is much tighter for the Dhalo definition than for FoF haloes. For FoF haloes 5% of the haloes have $M_{\rm FOF}/M_{200} \gtrsim 2$ and 1% $M_{\rm FOF}/M_{200} \gtrsim 3$. In contrast for Dhaloes only 5% have $M_{\rm Dhalo}/M_{200} \gtrsim 1.5$ and less than 1% have $M_{\rm Dhalo}/M_{200} > 2$. In the Dhalo panel only Dhaloes that are bijectively matched with FoF haloes are included. Since such pairs of haloes contain the same most massive subhalo, the centres used for calculating M_{200} are identical and result in the same M_{200}. Furthermore, since Fig. 3.1 indicates that all FoF haloes more massive than $10^{12}\ h^{-1}\ {\rm M}_\odot$ have a bijectively matching Dhalo, then above $10^{12}\ h^{-1}\ {\rm M}_\odot$ we are comparing the same population of haloes and using the same values of M_{200}. Consequently the wider distribution of $M_{\rm halo}/M_{200}$ for FoF is directly caused by the wider spread in $M_{\rm FoF}$ masses. For the cases where $M_{\rm FoF} \gg M_{200}$ there is one or more substantial components of the FoF halo that lies outside r_{200}. We will see in Fig. 3.4 that these are generally secondary mass concentrations that are linked by tenuous bridges of quite diffuse material. The Dhaloes have a tighter distribution of $M_{\rm halo}/M_{200}$ as in this algorithm these secondary concentrations are successfully split off and result in separate distinct Dhaloes.

Our results for FoF haloes are consistent with earlier investigations. Harker et al. (2006); Evrard et al. (2008); Lukić et al. (2009) found that approximately 80-85% of FoF haloes are isolated haloes while 15-20% of FoF haloes have irregular morphologies, most of which are described in Lukić et al. (2009) as "bridged haloes". The distribution of $M_{\rm FoF}/M_{200}$ for "bridged haloes" given in figure 7 of Lukić et al. (2009) is very similar to the 20% tail of our distribution above $M_{\rm FoF}/M_{200} = 1.5$, while the isolated haloes in Lukić et al. (2009) have a distribution similar to the remaining 80% of our distribution.

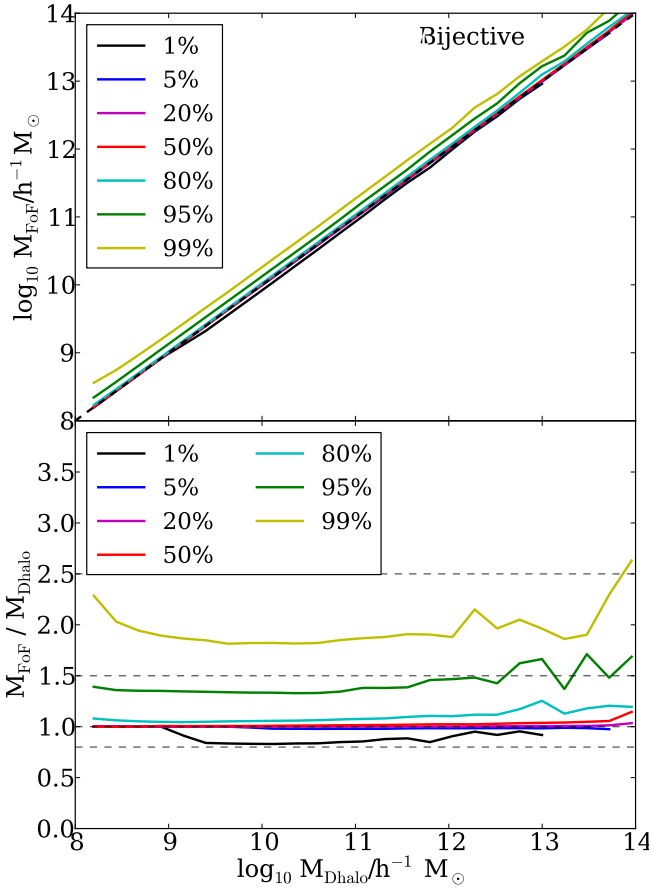

Figure 3.3: In the top panel, the 1, 5, 20, 50, 80, 95 and 99 percentiles of the distribution of FoF halo mass, $M_{\rm FoF}$, is plotted against $M_{\rm Dhalo}$ for the bijectively matched pairs of haloes. In the bottom panel, the same percentiles of the distribution of the mass ratio $M_{\rm FoF}/\ M_{\rm Dhalo}$ is plotted as a function of Dhalo mass. The black dashed lines are where $M_{\rm FoF}/M_{\rm Dhalo}$ =0.8, 1, 1.5 and 2.5.

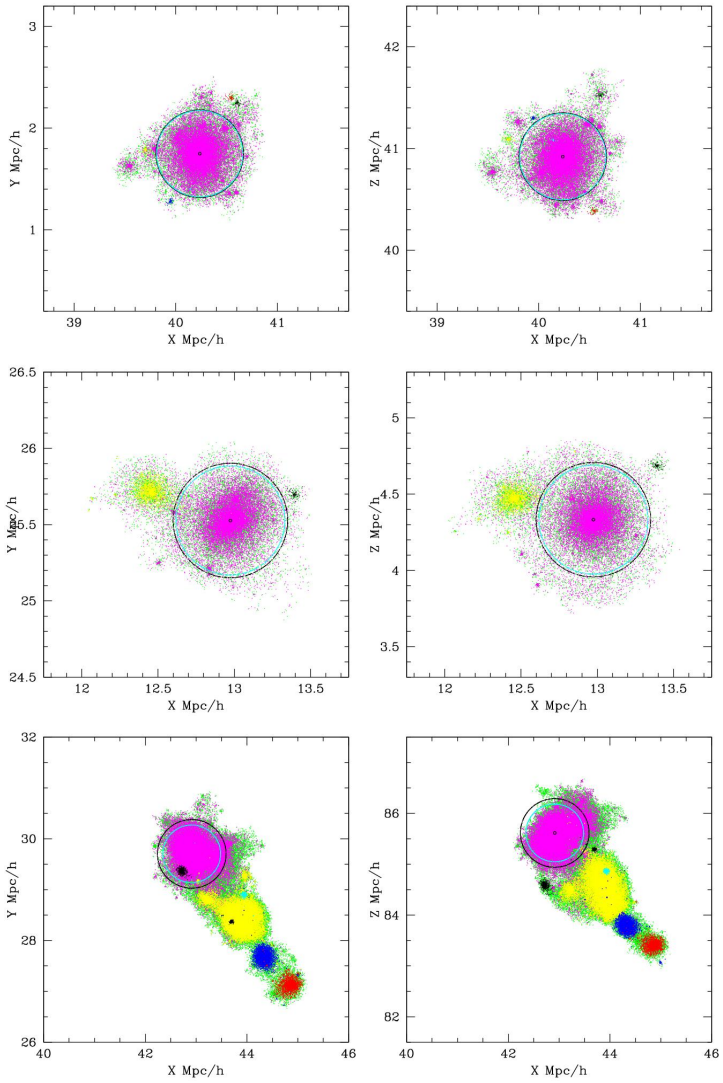

Figure 3.4: Three examples of the relationship between FoF haloes and Dhaloes. In each panel all the points plotted are from a single FoF halo. First all the FoF particles were plotted in green and then subsets belonging to specific Dhaloes were over-plotted. The magenta points are those belonging to the bijectively matched Dhaloes. Other colours are used to indicate particles belonging to other non-bijective Dhaloes with a unique colour used for each separate Dhalo. Two projections of each halo are shown. The left panels show the X-Y and right the X-Z plane. The black circle marks r_{200} of the FoF halo and the cyan circle marks twice the half mass radius of the main subhalo of the FoF halo. The top row shows a typical case where $M_{\mathrm{FoF}} \approx M_{\mathrm{Dhalo}}$. Here $M_{\mathrm{FoF}} = 2.6 \times 10^{13} h^{-1}$ $\mathrm{M}_{\odot}$, $M_{200} = 1.9 \times 10^{13} h^{-1}$ $\mathrm{M}_{\odot}$, and $r_{200} = 0.43 h^{-1}$ Mpc. The middle panel shows **an** example where the mass ratio $M_{\mathrm{FoF}}/M_{\mathrm{Dhalo}} = 1.5$ with $M_{\mathrm{FoF}} = 1.7 \times 10^{13} h^{-1}$ $\mathrm{M}_{\odot}$, $M_{200} = 1.2 \times 10^{13} h^{-1}$ $\mathrm{M}_{\odot}$ and $r_{200} = 0.375 h^{-1}$ Mpc. The bottom row shows an extreme example where $M_{\mathrm{FoF}} \gg M_{\mathrm{Dhalo}}$ and the FoF halo is split into many Dhaloes. Here $M_{\mathrm{FoF}} = 1.4 \times 10^{14} h^{-1}$ $\mathrm{M}_{\odot}$, $M_{200} = 7.1 \times 10^{13} h^{-1}$ $\mathrm{M}_{\odot}$ and $r_{200} = 0.67 h^{-1}$ Mpc

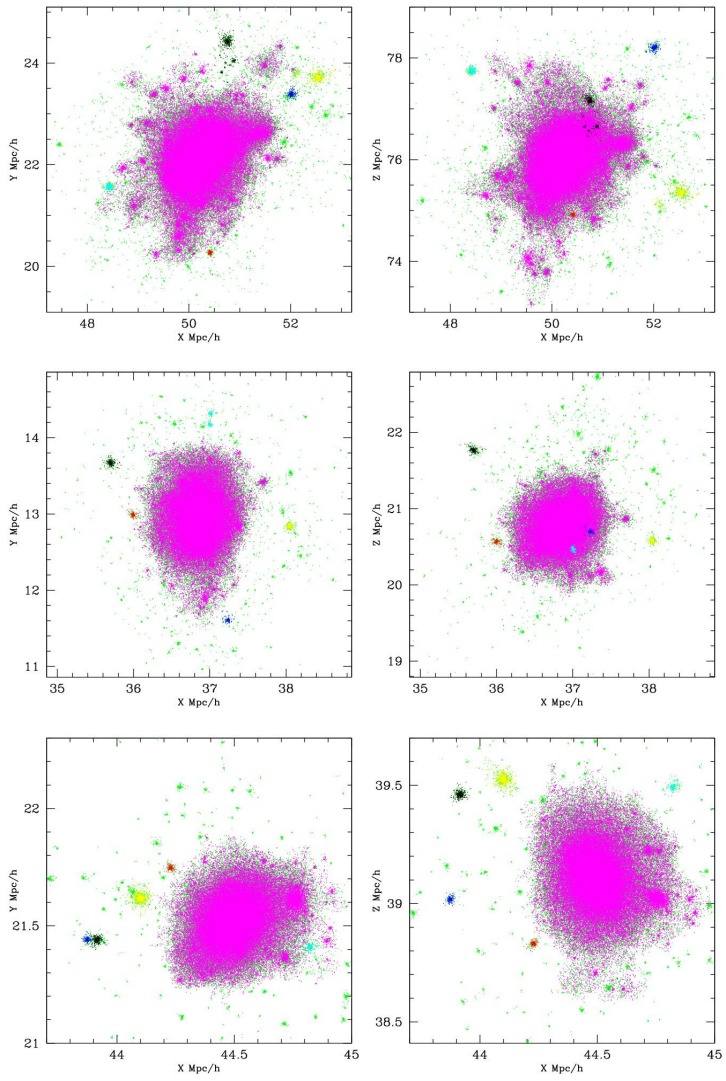

Figure 3.5: Examples of three typical Dhaloes showing how a single Dhalo can be composed of more than one FoF halo. In each panel all the points plotted are from a single Dhalo. First all the Dhalo particles were plotted in green and then subsets belonging to specific FoF haloes were over plotted. The magenta points are those belonging to the bijectively matched FoF halo. Other colours are used to indicate particles belonging to other FoF haloes with a unique colour used for each separate FoF halo. Two projections of each halo are shown. The left panels show the X-Y and right the X-Z plane. From top to bottom the Dhalo masses of these examples are $M_{\rm Dhalo} = 4.2 \times 10^{14} h^{-1}$ M$_\odot$, $M_{\rm Dhalo} = 6.8 \times 10^{13} h^{-1}$ M$_\odot$ and $M_{\rm Dhalo} = 5.4 \times 10^{12} h^{-1}$ M$_\odot$. In all cases the majority of the Dhalo mass is contained in the single bijectively matched FoF halo and the secondary FoF haloes are typically 100 times less massive.

3.2.2 Mass scatter plots

We now turn to directly comparing the mass assigned to FoF haloes and their corresponding Dhaloes. Fig. 3.3 compares the distributions of these two masses and their ratio for bijectively matched FoF and Dhaloes, i.e. haloes which contain the same most massive subhalo. First we see that the median of the distribution is very close to the one-to-one line. Furthermore on one side the distribution cuts off very sharply with far fewer than 1% of haloes having FoF masses significantly lower than their corresponding Dhalo mass. In principal $M_{\rm Dhalo} > M_{\rm FoF}$ can occur as one aspect of the Dhalo algorithm as it includes satellite subhaloes that previously passed through the main halo even if they are now sufficiently distant so as not to be linked into the corresponding FoF halo. However, such subhaloes are typically much less massive than the main subhalo and the mass gained in this way is out weighed by other sources of mass loss. On the other side of the distribution there is a significant tail of haloes for which $M_{\rm FoF} > M_{\rm Dhalo}$. We see that approximately 5% have $M_{\rm FoF} > 1.5M_{\rm Dhalo}$ and 1% have $M_{\rm FoF} > 2M_{\rm Dhalo}$. These fractions are largely independent of Dhalo mass. The main reason for this tail is the presence of FoF haloes that have a significant secondary mass concentration, often linked by a low density bridge, that the Dhalo algorithm succeeds in splitting off. For these bijectively matched haloes $M_{\rm FoF}$ is unlikely to significantly exceed $2M_{\rm Dhalo}$ as if a single secondary mass concentration had a subhalo of mass greater than that of the most massive subhalo in the Dhalo we would not have a bijective match. However, in rare instances $M_{\rm FoF} > 2M_{\rm Dhalo}$ can occur when the FoF halo contains several massive secondary mass concentrations.

To illustrate the relationship between FoF and Dhaloes we show three examples in Fig. 3.4 that have been chosen to be representative of different points in the $M_{\rm FoF}$–$M_{\rm Dhalo}$ distribution. The halo shown in the top row is representative of the majority of cases, namely those with $M_{\rm FoF} \approx M_{\rm Dhalo}$. Here the only particles from the FoF halo that are not included in the Dhalo are a diffuse cloud of unbound particles and the particles in a couple of subhaloes whose centres lie outside twice the half mass radius of the main subhalo. We stress that these small differences are what is typical for corresponding FoF and Dhaloes.

The middle row of Fig. 3.4 shows an example where $M_{\rm FoF}/M_{\rm Dhalo} = 1.5$, which corresponds to the 95th percentile of the distribution shown in Fig. 3.3. Here the FoF halo is split into three well separated Dhaloes. The main Dhalo is dominant, but there two secondary Dhaloes, one a lot more massive than the other, lying outside the r_{200} of the main Dhalo. For the purposes of semi-analytic galaxy formation models such as GALFORM the three separate haloes given by the Dhalo definition is clearly a better description than the single FoF halo as one would not expect the gas reservoirs associated with these distinct haloes to have merged at this stage and so each should be able to provide cooling gas to their respective central galaxies.

The bottom row of Fig. 3.4 shows a rare example with $M_{\rm FoF}/M_{\rm Dhalo} \approx 2$, the 99th percentile of the distribution, in which a single FoF halo is split into several substantial Dhaloes. In this and the previous example the FoF halo is clearly far from spherical and a large proportion of the FoF halo mass lies outside the virial radius that is defined by centring on the potential minimum of the most massive substructure. Clearly characterising such haloes by a NFW profile fit just to the mass within the virial radius would be an inadequate description of the halo. In fact, in most studies of halo concentrations, including our analysis presented in Section 4.3, these haloes would be deemed to be unrelaxed and excluded from subsequent analysis. In contrast, the Dhaloes in each of the examples presented are much closer to being spherical with only a small amount of mass outside their respective virial radii. Each of the primary Dhaloes in Fig. 3.4, including the one in the bottom panel, are sufficiently symmetrical and virialized to pass the relaxation criteria that we employ in Section 4.3 even though the corresponding FoF haloes in the bottom two panels are not.

In the example shown in the bottom row of Fig. 3.4 we also see a case of a Dhalo that has two distinct components. Here the two clumps of black points are a single Dhalo due to the fact that they passed directly through each other at a redshift $z = 0.89$. This extreme example must have been a high speed encounter and so any galaxies they contained would have been unlikely to merge, but their extended hot gas distributions would have interacted and possibly merged. It is

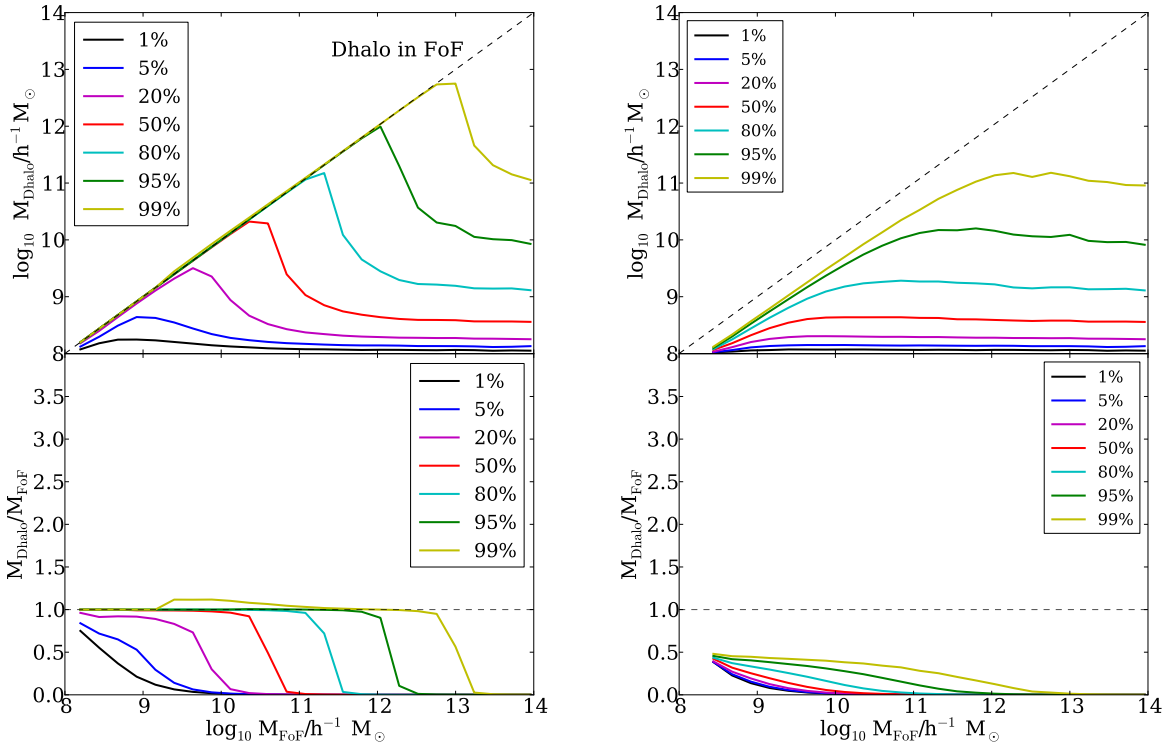

Figure 3.6: In the left hand panels, we plot the median, 1, 5, 20, 80, 95 and 99 percentiles of the distribution of Dhalo mass, $M_{\rm Dhalo}$ (upper), and mass ratio $M_{\rm Dhalo}/M_{\rm FoF}$ (lower) against $M_{\rm FoF}$ for all the Dhalo matches to each FoF halo. The black dashed lines in each panel mark where $M_{\rm Dhalo}/M_{\rm FoF}$ =1. In the right hand panel, we plot the same quantities but only for secondary Dhaloes in each FoF halo.

for this reason that it is useful in the semi-analytic models to associate them as a single halo.

The Dhalo algorithm quite frequently merges several FoF haloes together into a single Dhalo as a consequence of the way it avoids splitting up subhaloes which at an earlier timestep were in a single Dhalo. However unlike the extreme example we have just seen the typical masses of subhaloes which pass through a Dhalo and then emerge to once again become a distinct FoF halo are much lower than the mass of the main FoF halo. This is illustrated in Fig. 3.5, where we show the particles of three typical Dhaloes of a range of masses colour coded

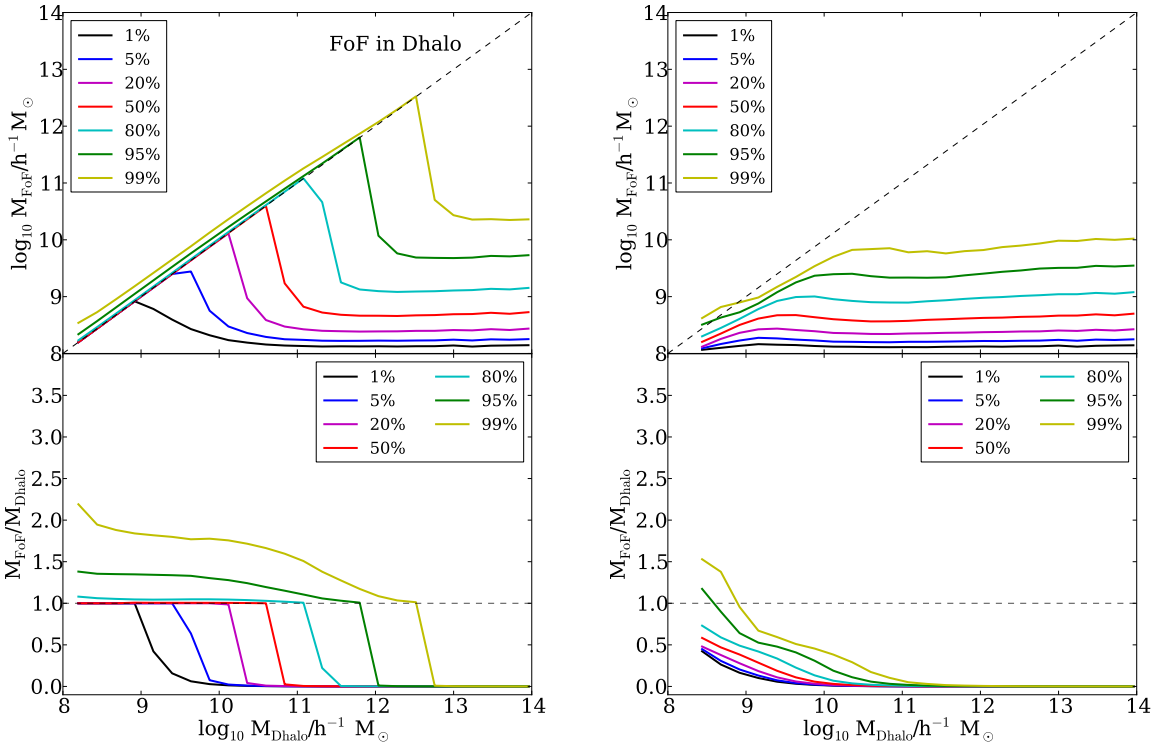

Figure 3.7: As Fig. 3.6 but with the role of FoF and Dhalo reversed. In the left hand panels, we plot the median, 1, 5, 20, 80, 95 and 99 percentiles of the distribution of FoF halo mass, $M_{\rm FoF}$ (upper), and mass ratio $M_{\rm FoF}/M_{\rm Dhalo}$ (lower) against $M_{\rm Dhalo}$ for all the FoF halo matches to each Dhalo. The black dashed lines in each panel mark where $M_{\rm FoF}/M_{\rm Dhalo}$ =1. In the right hand panel, we plot the same quantities but only for secondary FoF in each Dhalo.

by their FoF halo membership. In each case we immediately see that the vast majority of the Dhalo particles also belong to the (bijectively) matched FoF halo. However in addition there are isolated clumps of particles in the outskirts of each Dhalo which belong to much smaller distinct FoF haloes. There are also similar nearby clumps of particles which due to surrounding diffuse material are linked into the main FoF halo. In all cases each of these clumps are typically less than one percent of the mass of the main halo. From the perspective of semi-analytic galaxy formation models it makes sense to treat each of these clumps equally. For instance, they have all been within twice the half mass radius of the main Dhalo and could therefore have been ram pressure stripped of their diffuse gaseous haloes. In GALFORM satellite galaxies move with the subhalo within which they formed (or if the descendant of the subhalo drops below the 20 particle threshold with the particle that was previously the potential minimum of its subhalo) and so the satellite galaxy positions reflect the spatial distribution of these subhaloes even if they move far from the halo to which they are associated.

3.3 Non-bijective FoF and Dhalo matches

So far we have just compared FoF–Dhalo pairs which form a bijective match, that is their most massive subhaloes are identical. However there are other cases such as the examples of secondary Dhaloes in Fig. 3.4 in which the main subhalo of the Dhalo is not the most massive subhalo in the corresponding FoF halo and conversely examples such as the secondary FoF haloes in Fig. 3.5 in which the main subhalo of the FoF halo is not the most massive subhalo in the corresponding Dhalo. We will refer to this former set of matches as Dhalo in FoF halo and the latter as FoF in Dhalo matches. Note that the bijective matches are a subset of both of these sets, i.e. they are the intersection of the two sets of matches. To have a complete census of the correspondence between FoF and Dhaloes it is important that we include non-bijectively matched haloes in our comparison. We compare the Dhalo to FoF halo masses for these two sets of pairings in Fig. 3.6 and 3.7.

The left hand panels of Fig. 3.6 show for all the Dhalo in FoF matches the dependence of the mass, $M_{\rm Dhalo}$, and the mass ratio, $M_{\rm Dhalo}/M_{\rm FoF}$ on the FoF halo mass. The right hand panel shows the same quantities but only for secondary Dhalo in FoF halo haloes, i.e. excluding the bijective matches. Focusing first on the right hand panels, we see that the percentiles of the distribution of secondary $M_{\rm Dhalo}$ values are all horizontal lines at high $M_{\rm FoF}$, indicating that in this regime the distribution of $M_{\rm Dhalo}$ is independent of $M_{\rm FoF}$. This suggests that the secondary Dhaloes that are linked into high mass FoF haloes by bridges of diffuse material are essentially drawn at random from the Dhalo population. We note that in this way the FoF halo can be hundreds or more times more massive than many of the Dhaloes it contains. In these same panels, we see that at lower masses the distribution of Dhalo masses is sharply truncated at $M_{\rm Dhalo} = M_{\rm FoF}/2$. This is essentially by construction as if a Dhalo with mass greater than $M_{\rm FoF}/2$ were linked into the FoF halo then its most massive subhalo would very likely to be the most massive subhalo of the whole FoF halo and hence there would be a bijective match and this pairing would be excluded from this plot. The left hand panels of Fig. 3.6, which includes the bijective matches show a more complex distribution. However it can be easily understood as resulting from the superposition of the distribution from the right hand panel with the distribution of bijective matches shown in Fig. 3.3. At very low masses most FoF haloes contain only a single resolved subhalo and so the FoF halo cannot be split into multiple Dhaloes and so the overall distribution is dominated by the bijective matches resulting in a tight correlation between $M_{\rm Dhalo}$ and $M_{\rm FoF}$. With increasing FoF mass there are more and more secondary Dhaloes per FoF halo. They increasingly dominate over the bijective matches and so the contours tend to their values in the right hand panel.

Fig. 3.7 shows the distribution of FoF halo mass for the FoF in Dhalo matches. Again the right hand panes show the distribution for just the secondary matches while the left hand panels also include the primary or bijective matches. Comparing the right hand panels of Fig. 3.7 and Fig. 3.6 we see that the corresponding contours are shifted to lower masses. Thus it is rarer for a Dhalo to contain

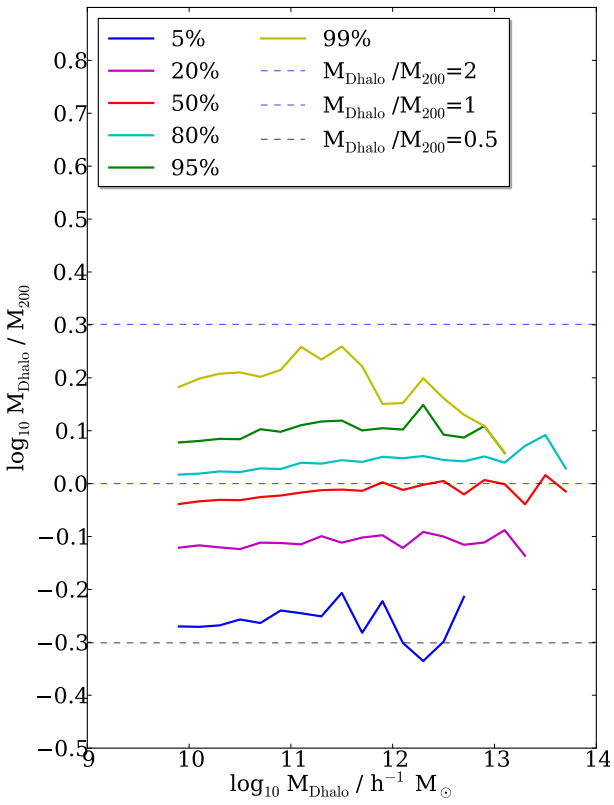

Figure 3.8: Like the right hand panel of Fig. 3.2, but for non-bijective Dhaloes. The curves show the median, 5, 20, 80, 95, 99 percentiles of the ratio between the Dhalo mass, $M_{\rm Dhalo}$, and the virial mass, M_{200}. The horizontal dashed lines indicate $M_{\rm Dhalo}/M_{200}$ = 0.5, 1.0, 2.0.

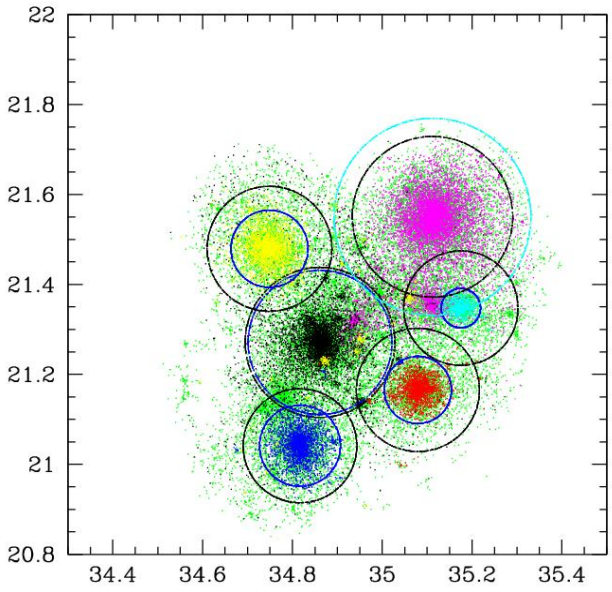

Figure 3.9: An example of one FoF halo split by the Dhalo algorithm into several Dhaloes. All the points plotted are from a single FoF halo. First all the FoF particles are plotted in green and then subsets belonging to specific Dhaloes are over plotted. The magenta points are those belonging to the bijectively matched Dhalo. Other colours are used to indicate particles belonging to other Dhaloes with a unique colour used for each separate Dhalo. The black circle around the magenta points marks r_{200} of the FoF halo and is also the r_{200} of the bijective Dhalo. The concentric cyan circle marks twice the half mass radius of this main subhalo. The other black circles show r_{200} locations for the non-bijective Dhaloes, while the concentric blue circles indicate twice the half mass radius of the corresponding subhalo.

massive secondary FoF halo than it is for FoF halo to contain massive secondary Dhalo. The secondary Dhaloes arise from the remerging step in the Dhalo algorithm whereby two subhaloes that have passed through each other (the smaller has come within twice the half mass radius of the larger) are deemed thereafter always to be part (or satellite components) of the same Dhalo even if they subsequently separate sufficiently to become distinct FoF haloes. This occurs reasonably frequently, but as in the examples shown in Fig 3.4 the secondary FoF haloes are typically much less massive than the primary and contribute little to the total mass of the halo. Interestingly the near horizontal contours in the upper right hand panel Fig. 3.7 indicate that the mass distribution of this population of secondary FoF haloes is approximately independent of $M_{\rm Dhalo}$ for high Dhalo masses. As these FoF haloes are often heavily stripped by their passage through the main Dhalo this is not a trivial result. The contours begin to dip at lower masses reflecting the fact it is unlikely for a matched FoF halo to have a mass greater than about one half of $M_{\rm Dhalo}$ without it being the primary or bijective match. This expectation is violated for $M_{\rm Dhalo} < 10^9 h^{-1}$ M$_\odot$, but this is a resolution effect because at such low masses secondaries with $M_{\rm FOF} \ll M_{\rm Dhalo}$ fall below the 20 particle limit of the catalogue and so their absence biases the distribution towards higher ratios.

The left hand panels of Fig. 3.7 are for all the matches of FoF in Dhalo, including the bijective matches. These distributions can be understood as a superposition of the distributions in the right hand panels with the distribution for bijective matches shown in Fig. 3.3. At low masses the bijective halo matches dominate whereas at large $M_{\rm Dhalo}$ there are many FoF haloes matched to each Dhalo. Thus, for example, at $M_{\rm Dhalo} \approx 10^{10.5} h^{-1}$ M$_\odot$ we transition from 50% of the matched FoF haloes being primary to 50% of them being much lower mass ($M_{\rm FoF} \approx 10^{8.7} h^{-1}$ M$_\odot$) secondary FoF haloes.

In section 3.1.1, we examined the distribution of the $M_{\rm Dhalo}/M_{200}$ ratio for the bijectively matched haloes. We are also interested in this distribution for the non-bijective Dhaloes shown in Fig. 3.8. We immediately notice the distribution is shifted towards lower values than the corresponding distribution for

the bijective haloes shown in Fig. 3.2. The origin of this shift can be understood by reference to Fig. 3.9 which shows an example of a FoF halo which is split into several Dhaloes. The Dhalo whose particles are plotted in magenta is the bijective match of the FoF halo and the Dhaloes plotted in other colours are non-bijective matches. The black circles in Fig. 3.9 show the location of r_{200} for each of the Dhaloes, while the other circles show the location of twice the half-mass radius of each Dhalo. For bijectively matched Dhaloes, the majority of which are isolated, r_{200} is typically slightly smaller than twice the half-mass radius. In contrast we see in Fig. 3.9 that for many of the non-bijectively matched Dhaloes twice the half mass radius is much smaller than r_{200}. This is a consequence of the SUBFIND algorithm which determines the extent of a subhalo by finding saddle points in the density distribution (Springel et al., 2001). Hence as a subhalo enters a dense environment the mass assigned to it by SUBFIND is decreased. This environmentally dependent effect both lowers $M_{\rm Dhalo}$ relative to M_{200} and increases the scatter in this relation.

3.4 Conclusion

We have shown that unlike the FoF algorithm the Dhalo algorithm is successful in avoiding distinct mass concentrations being prematurely linked together into a single halo when their diffuse outer haloes touch. We have also illustrated how some Dhaloes can be composed of more than one FoF halo. This occurs as structure formation in CDM models is not strictly hierarchical and occasionally a halo, after falling into a more massive halo, may escape to beyond the virial radius of the more massive halo. For the purposes of the GALFORM it is convenient to consider such haloes as remaining as satellites of the main halo. We find that such remerged FoF haloes are not uncommon, but contribute very little mass to the larger haloes to which they are (re)attached.

Approximately 90% of the Dhaloes have a unique one-to-one, bijective, match with a corresponding FoF halo. For this subset of haloes the mass of the Dhalo, $M_{\rm Dhalo}$, correlates much more closely with the standard virial mass, M_{200}, than

does the FoF mass. The median $M_{\rm FoF}/M_{200} = 1.2$ and 90% of the distribution of this mass ratio spans a factor 1.9, while for the same Dhaloes the median $M_{\rm Dhalo}/M_{200} = 1.15$ and corresponding width of the distribution spans only a factor 1.3. The larger scatter in the FoF case is often caused by secondary mass concentrations that lie outside the r_{200} radius of the main substructure and are linked into the FoF halo by particle bridges in overlapping diffuse haloes. The non-bijective Dhaloes have a wider distribution, with 90% of the distribution spanning a factor 2.2 and with the median ratio reduced to $M_{\rm Dhalo}/M_{200} = 0.95$. This is due to the SUBFIND substructure finder, which is part of the Dhalo algorithm, assigning less mass to subhaloes when they move into overdense environments. When utilised in GALFORM this systematic loss of mass is not an issue as the merger trees are preprocessed and mass is added back in to ensure the branches of the GALFORM merger trees always have monotonically increasing masses.

4 STATISTICAL PROPERTIES OF DHALOES

4.1 Introduction

It is now quite common for semi-analytic models to use halo merger trees extracted directly from N-body simulations (Springel et al., 2001; Helly et al., 2003; Hatton et al., 2003; Bower et al., 2006; Muñoz et al., 2009; Koposov et al., 2009; Busha et al., 2010; Macciò et al., 2010; Guo et al., 2011). There are many choices to be made both in defining the halo catalogues and in constructing the links between haloes at different times. Knebe et al. (2011) and Knebe et al. (2013) have found significant differences in even the most basic properties (e.g the halo mass function) of halo catalogues constructed with different group finding codes. Additionally, these halo catalogues can often be modified by the procedure of constructing the merger trees as some of the algorithms break up or merge haloes together in order to achieve a more consistent membership over time (Helly et al., 2003; Behroozi et al., 2013). So, for example, even if one starts with standard Friends-of-Friends (FoF) groups (Davis et al., 1985) the process of building the merger trees can alter the abundance and properties of the haloes.

Semi-analytic models such as GALFORM have the option of using information extracted directly from an N-body simulation or using Monte Carlo methods (see Jiang & van den Bosch, 2014, for a comparison of different algorithms) which make use of statistical descriptions of N-body results such as analytic halo mass functions (e.g. Sheth & Tormen, 1999; Jenkins et al., 2001; Evrard et al., 2002; White, 2002; Reed et al., 2003; Linder & Jenkins, 2003; Łokas et al.,

2004; Warren et al., 2006; Heitmann et al., 2006; Reed et al., 2007; Lukić et al., 2009; Tinker et al., 2008; Boylan-Kolchin et al., 2009; Crocce et al., 2010; Courtin et al., 2010; Bhattacharya et al., 2011; Watson et al., 2013) and models for the distribution of the concentrations of halo mass profiles (e.g. Navarro et al., 1995, 1996a; Bullock et al., 2001; Eke et al., 2001; Macciò et al., 2008). These statistical descriptions are often based on the abundance and properties of FoF haloes and so may not be directly applicable to the catalogues of haloes that result from the application of a specific merger tree algorithm. The internal structure of the dark matter haloes strongly influences galaxy formation models. Often the gas density profiles within dark matter haloes are assumed to be related to the dark matter profile, e.g. through hydrostatic equilibrium and these influence the rate at which gas cools onto the central galaxy. In addition the central potential of the dark matter halo effects the size and circular velocity of the central galaxy which in turn can have a strong effect on the expulsion of gas from the galaxy via SN feedback. Hence for semi-analytic galaxy formation modelling it is important to adopt models of the individual haloes that are consistent with the haloes that appear in the merger trees used by semi-analytic models.

Having thoroughly compared individual Dhaloes with their corresponding FoF haloes, we now turn to the statistical properties of the Dhaloes. We first look at the Dhalo mass function and then the statistics of their density profiles as characterised by fitting NFW profiles (Navarro et al., 1995, 1996a, 1997).

4.2 The Dhalo mass function

For many applications it is extremely useful to have an analytic description of the number density of haloes as a function of halo mass. A relevant example for us is when semi-analytic galaxy formation models are constructed using Monte-Carlo methods (Parkinson et al., 2008; Cole et al., 2000) for generating dark matter merger trees. In this case, in order to construct predictions of galaxy luminosity functions or any other volume averaged quantity (Cole et al., 2000; Berlind et al., 2003; Baugh et al., 2005; Neistein & Dekel, 2008; Bundy et al.,

2005; Giocoli et al., 2008; Moreno et al., 2008; van den Bosch et al., 2005), one needs knowledge of the halo mass function in order to know how many of each type of tree one has per unit volume. It has become common practice to assume the halo mass function is given by analytic fitting functions which have been fitted to the abundance of haloes found by the FoF or other group finding algorithms (Davis et al., 1985; Lacey & Cole, 1994; Knollmann & Knebe, 2009) in suites of cosmological N-body simulations. Murray et al. (2013) compare all the currently proposed fitting functions. In our semi-analytic modelling we would like to achieve consistent results when using Monte-Carlo merger trees or when using merger trees extracted directly from N-body simulations using the Dhalo algorithm. Hence it is important to directly determine the Dhalo mass function and to compare it to such fitting formulae.

We do this in Fig. 4.1 which compares the Dhalo and FoF mass functions that we measure in the MSII simulations with various analytic prescriptions (Jenkins et al., 2001; Sheth & Tormen, 2002; Warren et al., 2006; Reed et al., 2007; Tinker et al., 2008; Watson et al., 2013). The left hand panel shows the number density of haloes per unit logarithmic interval of mass from the nominal 20 particle mass resolution of the simulation up to $10^{14}h^{-1}$ $M_{\odot}$ which is the mass of the most massive haloes in the simulation. In constructing these mass functions the halo mass we use is simply the aggregated mass of all the particles assigned to each halo. Thus in the FoF case this is all particles linked to the halo by the FoF algorithm while in the Dhalo case it is the sum of the masses of the subhaloes that compose an individual Dhalo. Also shown on this panel are the predictions of various analytic prescriptions. To evaluate these we use $\sigma^2(M)$, the variance of the density fluctuations as a function of mass (using a top hat filter), corresponding to the input power spectrum of the MSII propagated to the output time of the simulation using linear theory. They are all clearly very similar and so in the left hand panel we expand the dynamic range of the comparison by plotting each mass function divided by the prediction of the Sheth & Tormen (2002) model.

The first thing that we note is that despite the sometimes quite large dif-

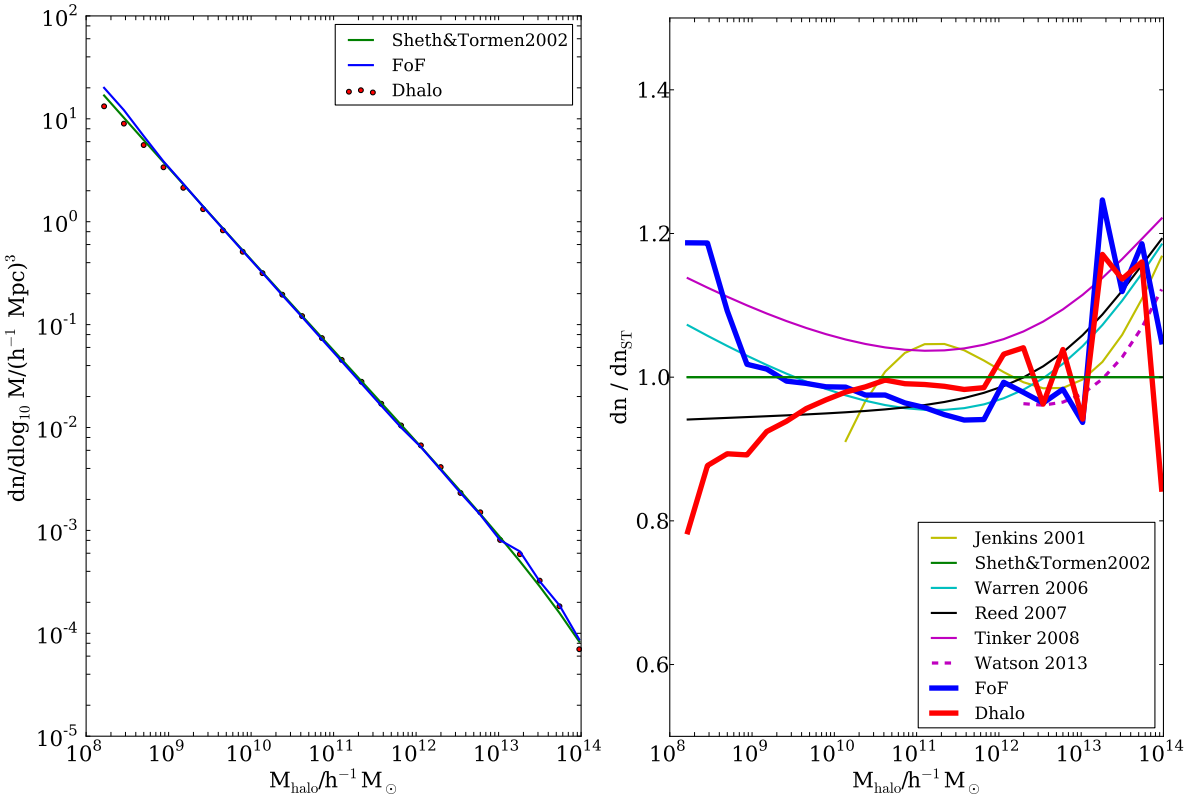

Figure 4.1: The left hand panel shows the differential mass functions for both FoF (linking length $b = 0.2$) haloes (blue line) and Dhaloes (red points) in the MSII simulation. We plot this down to $\sim 10^8 h^{-1}$ $M_\odot$, the mass corresponding to 20 particles in the MSII simulation and we also plot the Sheth and Tormen (2002) mass function as a comparison. To expand the dynamic range, the right hand panel shows the corresponding prediction of various analytic mass functions(Jenkins et al 2001, Warren et al 2006, Reed et al 2007, Tinker et al 2008, Watson et al 2013) as indicated in the legend but now relative to the Sheth and Tormen(2002) prediction. The FoF Dhalo data are now shown as the heavy blue and red lines.

ferences (see §3) in the masses of individual FoF and Dhaloes their two mass functions agree to within 5% for all masses greater than $10^{10}h^{-1}$ $M_\odot$. In the range $10^{10} \lesssim M_{\rm halo} \lesssim 10^{12.5}h^{-1}$ $M_\odot$ the Dhalo abundance is approximately 5% higher than FoF haloes as roughly 5% of Dhaloes are secondary members of FoF haloes. In other words, the FoF halo abundance has been suppressed relative to the Dhalo abundance by a fraction of them being composed of two or more Dhaloes that have been linked into one more massive FoF halo by diffuse material or bridges. There is also a competing effect, FoF haloes being remerged into single Dhaloes, which suppressed the Dhalo abundance, but this is a much smaller effect.

Below $10^{10}h^{-1}$ $M_\odot$ the abundance of FoF halo rises systematically above that of Dhaloes. Between $10^{10}h^{-1}$ $M_\odot$ and $8 \times 10^{8}h^{-1}$ $M_\odot$ this excess increases to about 10% and is caused by FoF haloes that are remerged to become secondary components of a larger Dhaloes (see Fig. 3.1). At lower masses ($\lesssim$ 100 particles) the sharp up turn in the FoF mass function relative to that of Dhaloes is due to an increasing fraction of the FoF haloes not containing a self-bound subhalo and so having no corresponding Dhalo (see Fig. 3.1). Thus this portion of the mass function is strongly affected by the resolution of the simulation.

The Jenkins et al (2001) fitting formula is within 10% of both the FoF and Dhalo mass functions for masses above $2 \times 10^{10}h^{-1}$ $M_\odot$. However below this mass it strongly under predicts the number density of low mass haloes. Note that we only plot this fit and that of Watson et al (2013) over the mass ranges used to constrain them in the original papers. The Watson et al (2013) mass function is only defined at very high masses where we have poor statistics. It lies somewhat below but is still compatible with our noisy estimates. The Warren (2006) model has the best agreement with our FoF mass function, fitting it well all the way down to 40 particles, beyond which we expect our limited resolution means that our FoF mass function is contaminated by spurious unbound chance groupings of particles. However the Reed (2007) mass function does a better job of matching the low mass end of our Dhalo mass function. The Sheth & Tormen mass function is intermediate at low masses between that of Warren (2006) and

Reed (2007), but systematically below the other models and our FoF and Dhalo mass function at high masses, though still only at the 15% level. The Tinker (2008) mass function predicts halo abundances that are about 5 to 10% higher than Warren (2006) and our estimated FoF abundances.

In summary, the Dhalo and FoF mass functions are very similar and only differ by more than 5% below $10^{10}h^{-1}$ $\mathrm{M}_{\odot}$. As a result the established analytic mass function models fit the Dhalo mass function almost as well as they do the standard FoF mass function. The differences between the different analytic fitting formulae are greater than the difference between the FoF and Dhalo mass functions. The Reed (2007) model is a slightly better description of the Dhalo mass function due to it predicting a slightly lower abundance at low masses.

4.3 Density profile fits

We now turn to the density profiles of the haloes as these are an important ingredient in semi-analytic models such as GALFORM where they influence the rate at which gas cools and set the gravitational potential well in which galaxies form. We choose to fit the halo density profiles using NFW (Navarro et al., 1996a, 1997) profiles

$$\frac{\rho_{\mathrm{NFW}}(r)}{\rho_{\mathrm{crit}}} = \frac{\delta_c}{r/r_{\mathrm{s}}(1+r/r_{\mathrm{s}})^2} \quad (r \leq r_{200}), \tag{4.3.1}$$

where δ_c is the characteristic density contrast, and r_{s} is the scale radius. We define the virial radius, r_{200}, as the radius at which the mean interior density equals 200 times the critical density, $\rho_{\mathrm{crit}} = 3H_0^2/(8\pi G)$. The concentration is defined as $c \equiv r_{200}/r_{\mathrm{s}}$. The definition of r_{200} implies that δ_c and c must satisfy

$$\delta_c = \frac{200}{3}\frac{c^3}{\ln(1+c) - c/(c+1)}. \tag{4.3.2}$$

Our choice of NFW profiles is motivated by their accuracy as a model of CDM haloes (Navarro et al., 1996a, 1997), their widespread use and so that our results can be compared to those in Neto et al. (2007) who studied the statistics of NFW concentrations for FoF haloes identified in the Millennium Simulation

(Springel, 2005a). To allow us to compare directly with Neto et al. (2007) we have followed their fitting procedure.

For each halo, we have computed a spherically-averaged density profile by binning the halo mass into 32 equally spaced bins in $\log_{10}(r)$ between the virial radius and $\log_{10}(r/r_{200}) = -2.5$, centred on the potential minimum. We fit the two free parameters, δ_c and $r_{\rm s}$ by minimising the mean square deviation

$$\sigma^2_{\rm fit} = \frac{1}{N_{\rm bin}-1}\sum_{i}^{N_{\rm bin}}[\log_{10}\rho(r_i) - \log_{10}\rho_{\rm NFW}(r_i|\delta_c, r_{\rm s})]^2 \qquad (4.3.3)$$

between the binned $\rho(r)$ and the NFW profile. As in Neto et al. (2007), we perform the fit over the radial range $0.05 < r/r_{200} < 1$. In order to be consistent with the original NFW work, we express the results in terms of fitted virial mass, M_{200}, and a concentration, $c_{200} \equiv r_{200}/r_{\rm s}$. We note that while the fitted value of M_{200} used here and the directly measured M_{200} used earlier (e.g. in Fig. 3.2) are not identical they in general agree very accurately with an rms scatter of less than 3%.

Neto et al. (2007) distinguished relaxed haloes from haloes that were not in dynamical equilibrium due to recent or ongoing mergers. They found that relaxed haloes were well fit by NFW profiles while the profiles of unrelaxed haloes were lumpier and yielded poorer fits with systematically lower concentrations. Hence to compare to Neto et al. (2007) we use the following three objective criteria to assess whether a halo has reached equilibrium (Neto et al., 2007; Gao et al., 2008; Power et al., 2012).

1. The fraction of mass in resolved substructures whose centres lie inside r_{200}: $f_{\rm sub} = \sum_{i\neq 0}^{N_{\rm sub}} M_{{\rm sub},i}/M_{200}$. We require $f_{\rm sub} < 0.1$ for relaxed haloes.

2. The centre of mass displacement, i.e. the difference between the position of the potential minimum and the centre of mass, $s = |r_{\rm c} - r_{\rm cm}|/r_{200}$ (Thomas et al., 2001). Note that, the centre of mass is calculated using all the particles within r_{200}, not only those belonging to the FoF or Dhalo. We require $s < 0.07$ for relaxed haloes.

3. The virial ratio, $2T/|U|$, where T is the total kinetic energy of halo particles

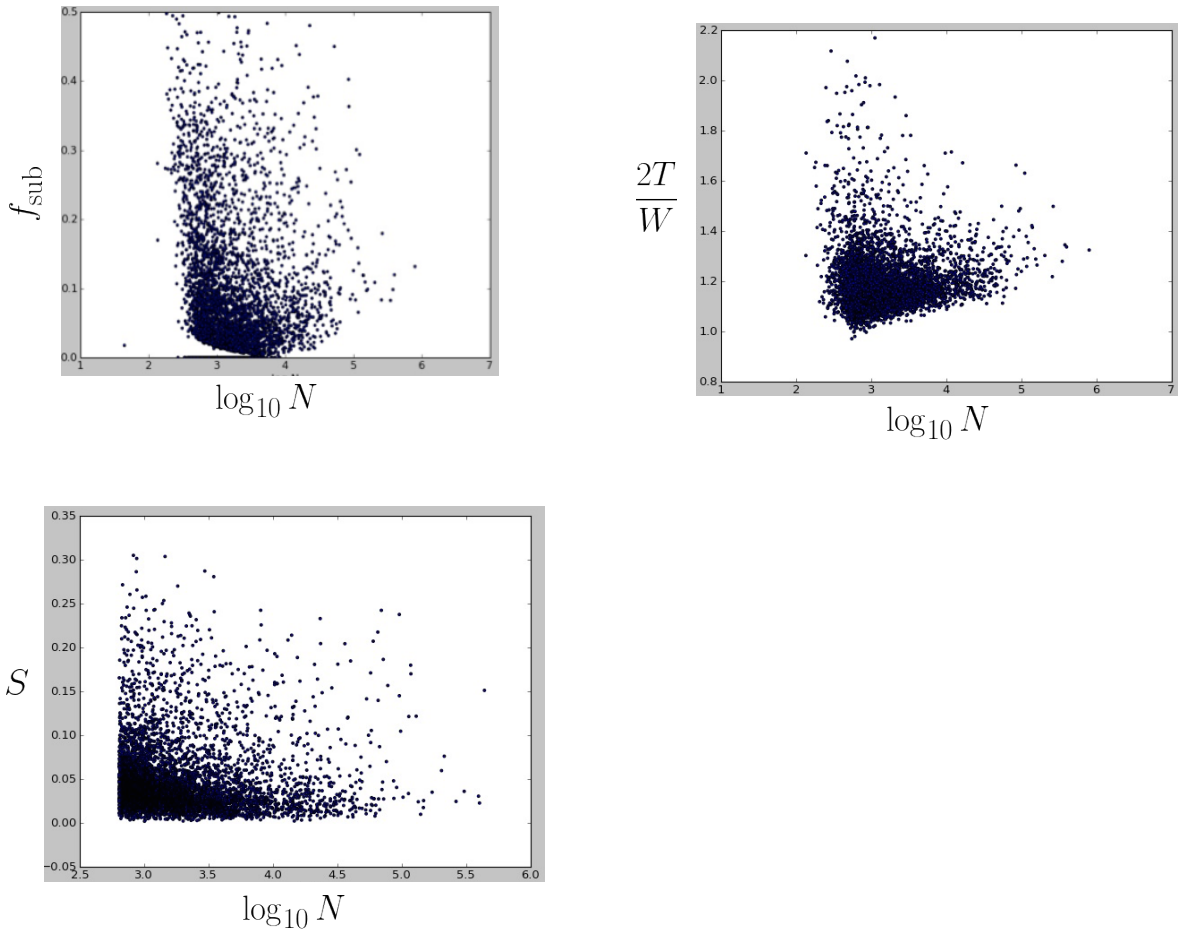

Figure 4.2: The distributions of the parameters $f_{\rm sub}$, $2T/|W|$, s, used to define our relaxed sample of haloes are shown as a function of the number of particles within the FoF halo.

within r_{200} and U is their gravitational potential self energy. We require $2T/|U| < 1.35$ for our relaxed haloes. (For haloes with more than 5000 particles we use a random subset of 5000 particles to estimate U.)

As a test of the Neto's relaxation criteria, we measured the parameters used in these criteria for all the FoF haloes in the milli-MillenniumII simulation. This simulation has the same volume, initial conditions and data format as MillenniumII (Boylan-Kolchin et al., 2009), but lower mass resolution. The distribution of these parameters for this sample is shown in Fig. 4.2. Like Neto we find the majority of haloes pass all three selection criteria. We also find that the $f_{\rm sub}$ pa-

rameter is the most sensitive to the number of particles. The distribution shown in the top left panel of Fig. 4.2 includes some haloes with no resolved substructures $f_{\rm sub} = 0$. In such cases, we can use the kinematic information based on $2T/|W|$ to reject haloes that, despite passing the criterion using $f_{\rm sub}$, are far from dynamical equilibrium. The criterion $s < 0.07$, which rejects haloes in which the centre of mass and potential are offset is the most constraining. We found that a big shift between mass centre and potential minimum often indicates an ongoing merger. This criterion is quite closely linked to the particle-bridge problem in FoF haloes. We find we can successfully remove the unrelaxed haloes in each halo mass range by combining all the three objective criteria.

Fig. 3.9 shows a single FoF halo and its component Dhaloes which we use to illustrate the application of these selection criteria and ability of NFW profiles to fit secondary/non-bijective Dhaloes. The spherically averaged density profiles and our NFW fits to each of these Dhaloes are shown in Fig. 4.3 along with the values of the three selection parameters $f_{\rm sub}$, s and $2T/|U|$. The top left panel of Fig. 4.3 shows the density profile and NFW fit for the main component of the FoF halo, which can be identified by the cyan circle in Fig. 3.9 which marks twice the half mass radius the most massive substructure in the FoF halo. In previous analyses of FoF haloes, such as Neto et al. (2007), this would be the only density profile fitted to the mass distribution shown in Fig. 3.9. The bijectively matched Dhalo has the same centre as the FoF halo and the NFW fit is performed on all the mass within r_{200}, (indicated by the concentric black circle) consequently the density profile and NFW fit of the bijectively matched Dhalo is necessarily identical to that or the corresponding FoF halo. Examining this region in Fig. 3.9, we can clearly see that the mass distribution is asymmetric and has several distinct substructures indicative of a recent merger. This halo is not relaxed according to the above selection criteria as it fails to satisfy the cut on $2T/|U|$. Also its value of the centre offset, s, comes close to the threshold. The NFW fit to its density profile can be seen to have significant deviations at both large and small radii.

We are also interested in whether NFW profiles provide acceptable fits to

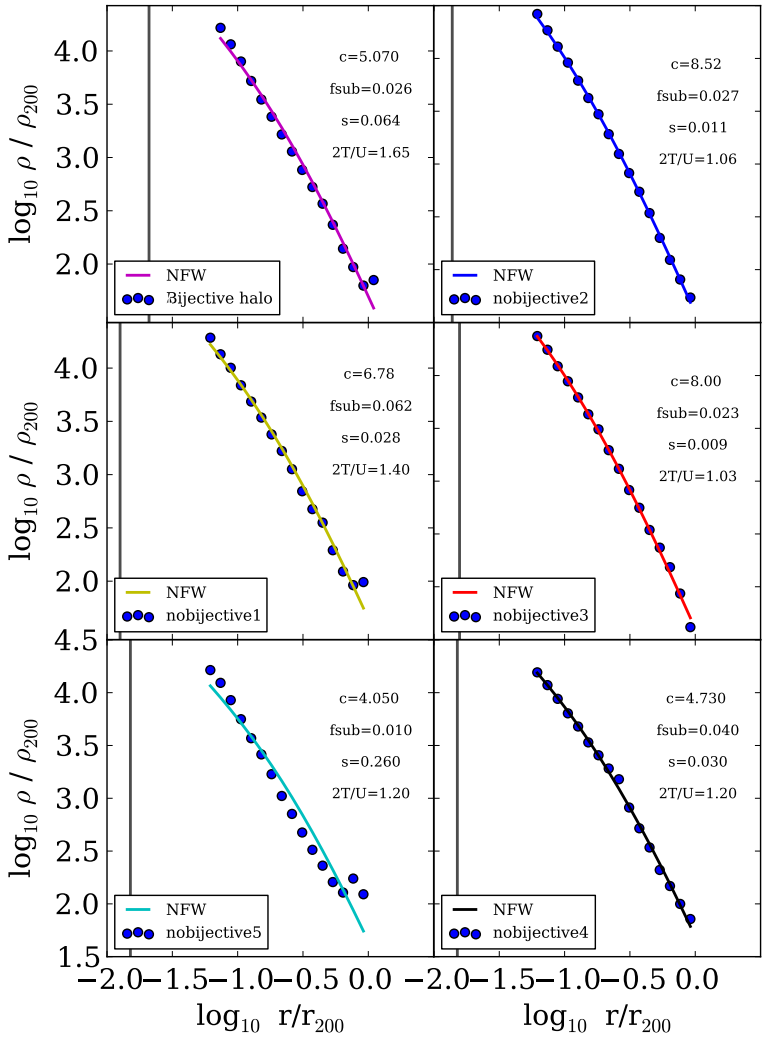

Figure 4.3: Density profiles, $\rho(r)$, for each of the Dhaloes shown in Fig. 3.9. The colour of the fitted NFW curve matches the colour coding of the individual Dhaloes in Fig. 3.9. The two-parameter, δ_c and r_s, NFW least-square fits were performed over the radial range $0.05 < r/r_{200} < 1$, shown by the black circles in Fig. 3.9. The minimum fit radius $r/r_{200} = 0.05$ is always larger than the convergence radius derived by Power et al (2003), which we indicate by the solid vertical line in each panel.

the other Dhaloes found within this single FoF halo. These are shown in the remaining panels of Fig. 4.3. According to the selection criteria three of these Dhaloes (those in the right-hand column) are relaxed. These are the blue, red and black Dhaloes in Fig. 3.9 and their density profiles are shown, respectively, in the top, middle and bottom right-hand panels of Fig. 4.3. In all cases we see that the NFW fits provide a good description of the mass profile of these relaxed Dhaloes. The remaining two Dhaloes fail one or other of the selection criteria. The yellow Dhalo of Fig. 3.9, whose density profile is shown in the middle-left panel of Fig. 4.3, marginally fails the cut on $2T/|U|$. The cyan Dhalo of Fig. 3.9, whose density profile is shown in the bottom-left panel of Fig. 4.3, which strongly exceeds the threshold on s, can be seen to be very poorly fit by the NFW profile and have a particularly low concentration. This Dhalo is very close to being within twice the half mass radius of the most massive substructure of the FoF halo, marked by the cyan circle in Fig. 3.9. This being the radius used by the Dhalo algorithm as part of its criteria to determine whether two subhaloes should be considered as two distinct haloes or components of the same halo. It is this proximity to a merger that both creates the large offset, s, between the potential minimum and the centre of mass within r_{200} and distorts the object's density profile. We also note that this Dhalo has the most extreme ratio of r_{200} to twice its half mass radius. In Fig. 3.4, we saw that for isolated haloes r_{200} and twice the half mass radius were very comparable, but in contrast we see in Fig. 3.9 that the r_{200} of secondary Dhaloes can be significantly boosted by the density of the surrounding environment.

This systematic difference in the ratio of Dhalo mass to M_{200} for bijective and non-bijective Dhaloes is illustrated in Fig. 3.8 which should be contrasted with the right-hand panel of Fig. 3.2. We see that the scatter in the ratio of $M_{\rm Dhalo}/M_{200}$ is considerably larger for the non-bijective Dhaloes than it is for bijective Dhaloes. For bijective Dhaloes the 5 to 95% range of the distribution spans only a 30% range in the ratio of $M_{\rm Dhalo}/M_{200}$, while this is increased to approximately a factor of two for the non-bijective Dhaloes. In addition the median $M_{\rm Dhalo}/M_{200}$ ratio is reduced from 1.2 for bijective Dhaloes to ≈ 0.95 for

non-bijective Dhaloes. These differences are principally caused by the way the SUBFIND algorithm (Springel et al., 2001) is affected by the local environment. SUBFIND locates the edge of a substructure by searching for a saddle point in the density distribution. Hence if the same sub-structure is placed in a denser environment this will move the saddle point in and reduce the mass that SUBFIND associates with the sub-structure (see Muldrew et al., 2011, for a detailed discussion). As a Dhalo mass is simply the sum of the masses of the subhaloes from which it is composed this in turn reduces the mass assigned to the Dhalo. This systematic dependence of Dhalo mass on environment is one of the reasons why instead of directly using the Dhalo mass as input to GALFORM semi-analytic model we instead force the halo masses in the halo merger trees to increase monotonically so that they do not artificially decrease, just prior to mergers, due to such environmental effects.

4.4 The mass-concentration relation

Here we compare the mass-concentration relation for FoF haloes that we find in the high resolution MSII simulation with that found by Neto et al. (2007) in the lower resolution Millennium Simulation.[1] We then go on to compare this relation with the relation we find for the secondary/non-bijective Dhaloes. There is no need to separately look at the bijective Dhaloes as their M_{200} and c are necessarily the same as that of the corresponding FoF haloes as they have the same centre and all the surrounding mass is used in the fit. As in Neto et al. (2007) the mass we use in these relations is the M_{200} of the NFW fit rather than the directly measured value. Fig. 4.4 shows concentration as a function of mass for the range $10^{10.5} < M_{200}/h^{-1}\mathrm{M}_{\odot} < 10^{13.75}$ for our catalogue of FoF

[1] As a precise test of our methods we first applied our analysis to FoF haloes in the milli-MillenniumII simulation, which has the same volume, initial conditions and data format as MillenniumII (Boylan-Kolchin et al., 2009), but lower mass resolution, equal to that of the Millennium Simulation (Springel et al., 2005) analysed by Neto et al. (2007). We found precise agreement with the mass-concentration relationship published in Neto et al. (2007).

haloes. The top panel is for our *relaxed* FoF halo sample, while the bottom panel shows results for all the FoF haloes, including systems that do not meet our equilibrium criteria. In each case we find a significant spread in concentration at fixed mass with a weak trend for decreasing concentration with increasing mass. This is generally interpreted (Navarro et al., 1995, 1996a, 1997; Bullock et al., 2001; Eke et al., 2001; Neto et al., 2007; Gao et al., 2008) as reflecting the typical formation time of the halo with the lowest mass haloes forming earliest and having high density cores which reflect the density of the universe at the time they formed. The dependence of the median concentration of FoF haloes on mass is well described by the power-law fit

$$c_{200} = 5.45 \pm 1.00 \left(M_{200}/10^{14} h^{-1}\,\mathrm{M}_{\odot} \right)^{-0.084 \pm 0.002}, \tag{4.4.4}$$

for *relaxed* haloes and by

$$c_{200} = 5.01 \pm 1.03 \left(M_{200}/10^{14} h^{-1}\,\mathrm{M}_{\odot} \right)^{-0.094 \pm 0.003} \tag{4.4.5}$$

for all haloes. These fits were performed only over the mass range $10^{10.5} < M_{200}/h^{-1}\mathrm{M}_{\odot} < 10^{13.75}$ due to poor statistics at higher masses and are shown by the blue solid lines in Fig. 4.4. Also shown on Fig. 4.4 is the fit for the median concentration for relaxed haloes found by Neto et al. (2007). We plot these green lines only for $M_{200} > 10^{12}/h^{-1}M_{\odot}$ corresponding to the resolution limit of their study. We see that over the overlapping mass range our median concentrations agree very well with those of Neto et al. (2007) indicating that the mass profiles over the fitted radial range, $-2.5 < \log(r/r_{200}) < 0$, are not affected by mass resolution. Our fit is also similar to the relation $c_{200} = 5.6(M_{200}/10^{14}h^{-1}M_{\odot})^{-0.098}$ found by Macciò et al. (2007) for relaxed haloes. The small difference could be because they fit the mean rather than median of the relation or due to differences in the criteria used to select relaxed haloes. Like us and Neto et al. (2007), Macciò et al. (2007) find unrelaxed haloes have systematically lower concentrations.

Having demonstrated that for FoF haloes we recover a mass-concentration relation which is in very accurate agreement with previous work (Neto et al., 2007; Macciò et al., 2007), we now want to compare mass-concentration relations for our bijective and non-bijective Dhaloes. The mass-concentration relation we

find for the bijective Dhaloes is practically identical to that of the FoF haloes plotted in Fig. 4.4 and so we have chosen not to effectively repeat the same plot. The similarity is inevitable as Fig. 3.1 shows that for masses greater than $10^{10.5}\ h^{-1}M_{\odot}$, for which we can measure concentrations, the fraction of FoF haloes that have bijective matches with Dhaloes is greater than 95% and these bijectively matched haloes have identical centres and so identical fitted NFW mass profiles.

In Fig. 4.5 we show the mass-concentration for relaxed and all non-bijective Dhaloes. These haloes are all secondary fragments of FoF haloes and so are a completely disjoint catalogue of haloes to those represented in the FoF mass-concentration relations of Fig. 4.4. To aid in comparing the two sets of relations we plot the power-law fits to the median mass-concentration relations of Fig. 4.4 as dashed lines in Fig. 4.5. It can be seen that these are very similar to the power-law fits to the median relations

$$c_{200} = 4.90 \pm 1.00 \left(M_{200}/10^{14}h^{-1}\,\mathrm{M}_{\odot}\right)^{-0.093\pm0.003}, \tag{4.4.6}$$

for *relaxed* and

$$c_{200} = 5.01 \pm 1.00 \left(M_{200}/10^{14}h^{-1}\,\mathrm{M}_{\odot}\right)^{-0.095\pm0.004} \tag{4.4.7}$$

for all the non-bijective Dhaloes which are shown by the solid lines in Fig. 4.5.

Comparison of the bars and whiskers in Fig. 4.4 and Fig. 4.5 show that the not only do the median mass-concentration relations for FoF and non-bijective Dhaloes agree very well, but the distribution of concentrations about the medians are also quite similar. The large number of haloes we have in the MII simulation enables us to look at these distributions in more detail and in Fig. 4.6 we show histograms of the concentration, distributions along with log-normal approximations

$$P(\log_{10} c) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2}\left(\frac{\log_{10} c - \langle \log_{10} c \rangle}{\sigma}\right)^2\right], \tag{4.4.8}$$

for two mass bins centred on 10^{11} and $10^{12}\ h^{-1}\mathrm{M}_{\odot}$. We see in all cases that the non-bijective Dhaloes have a very similar distribution of concentrations as the

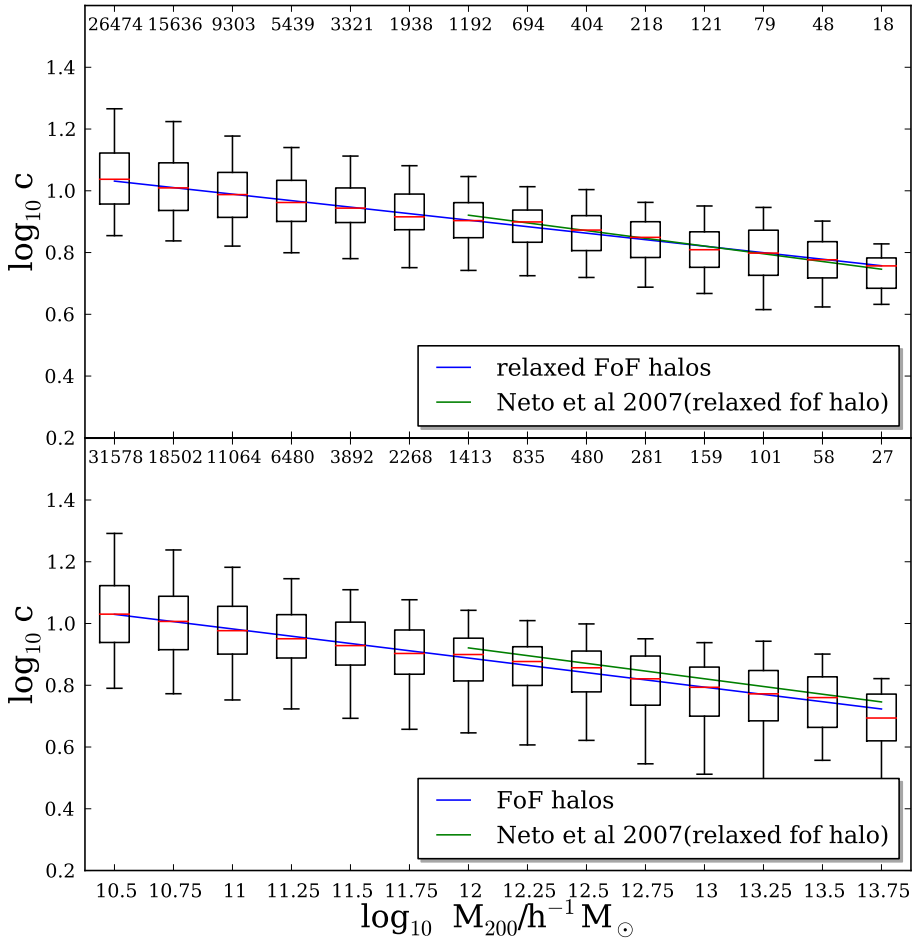

Figure 4.4: The mass-concentration relation for relaxed FoF haloes in MSII (top panel) and for all the FoF haloes (bottom panel). The boxes represent the 25% and 75% centiles of the distribution, while the whiskers show the 5% and 95% tails. The numbers on the top of each panel indicate the number of haloes in each mass bin. The median concentration as a function of mass is shown by the blue solid line and is well fit by the linear relations given in equations 4.4.4 and 4.4.5. The green lines in each panel correspond to fits of Neto et al (2007).

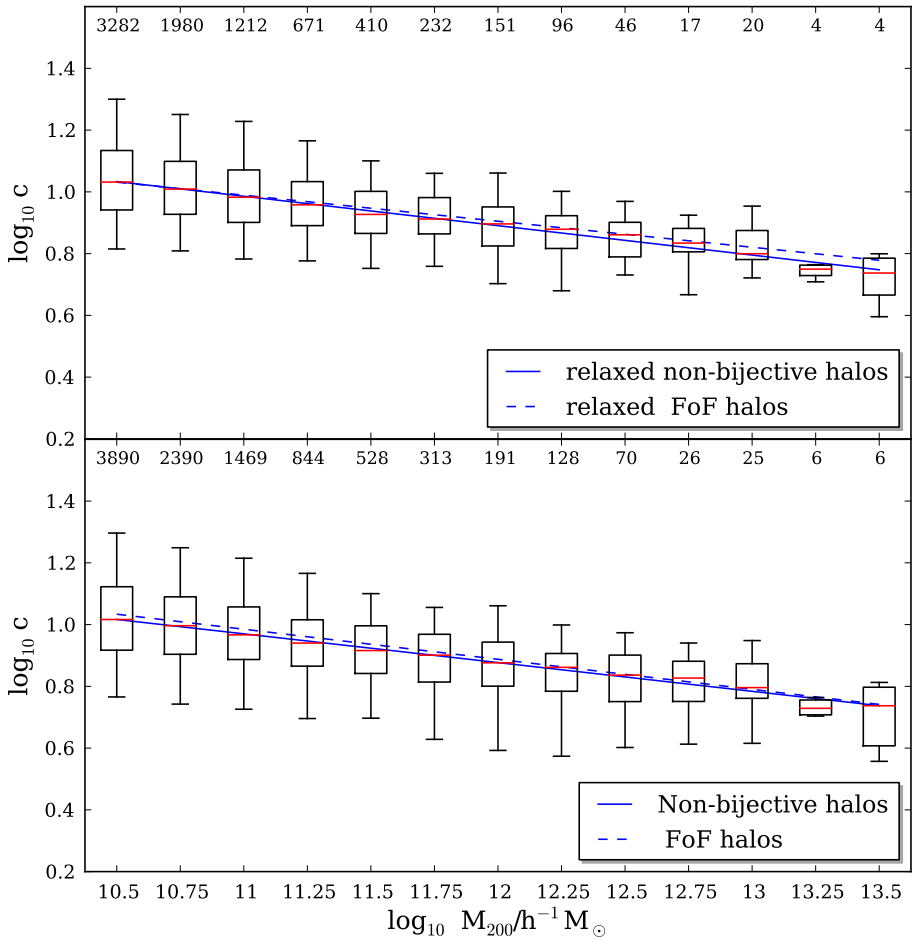

Figure 4.5: The mass-concentration relation for relaxed non-bijective Dhaloes in MSII (top panel) and for all the non-bijective Dhaloes (bottom panel). The boxes represent the 25% and 75% centiles of the distribution, while the whiskers show the 5% and 95% tails. The numbers on the top of each panel indicate the number of haloes in each mass bin. The median concentration as a function of mass is shown by the blue solid line and is well fit by the linear relations given in equations 4.4.6 and 4.4.7. The blue dashed line in each panel repeats the fits to the median mass-concentration relation for FoF haloes shown in Fig. 4.4

distribution of the corresponding FoF sample and that both are approximated accurately by log-normal distributions. Note that in both cases we are binning haloes by the M_{200} of their fitted NFW profile and so we are affected by the Dhalo mass being perturbed and suppressed in non-bijective Dhaloes. We recall that the FoF sample is essentially the same as the sample of bijectively matched Dhaloes and so we conclude that concentration distribution is essentially the same for both the primary Dhaloes and those that are secondary fragments of FoF haloes. In all cases the concentration distributions for the relaxed samples have slightly higher median concentrations and smaller dispersions than the corresponding complete mass selected samples.

Also of interest is the fraction of both FoF haloes and non-bijective Dhaloes that satisfy the equilibrium criteria. From the number of objects per mass bin given in the labels on Figs. 4.4 and 4.5 this can be seen to be in the range of 80 to 85% for both FoF and Dhaloes. One might at first expect that many multi-nucleated FoF haloes would fail both the threshold on the asymmetry, s, and the fraction of mass in sub-structures, $f_{\rm sub}$. However as these statistics are evaluated only using the mass within r_{200} and not across the whole FoF halo, $\gtrsim$ 98% of FoF haloes pass the substructure threshold and $\gtrsim$ 88% the asymmetry threshold. The first of these numbers is slightly lower for the non-bijective Dhaloes, i.e. only $\gtrsim$ 93% pass the substructure threshold. However those passing the more stringent asymmetry threshold is more comparable at $\gtrsim$ 86%, while for both FoF and non-bijective Dhaloes $\gtrsim$ 93% pass the criterion that the virial ratio $2T/|U| < 1.35$. Consequently the fraction of the non-bijective Dhaloes that pass the relaxation criteria is very similar to that for the FoF or bijective Dhaloes. Hence in both cases the mass-concentration distributions that we have quantified are representative of the vast majority of the haloes.

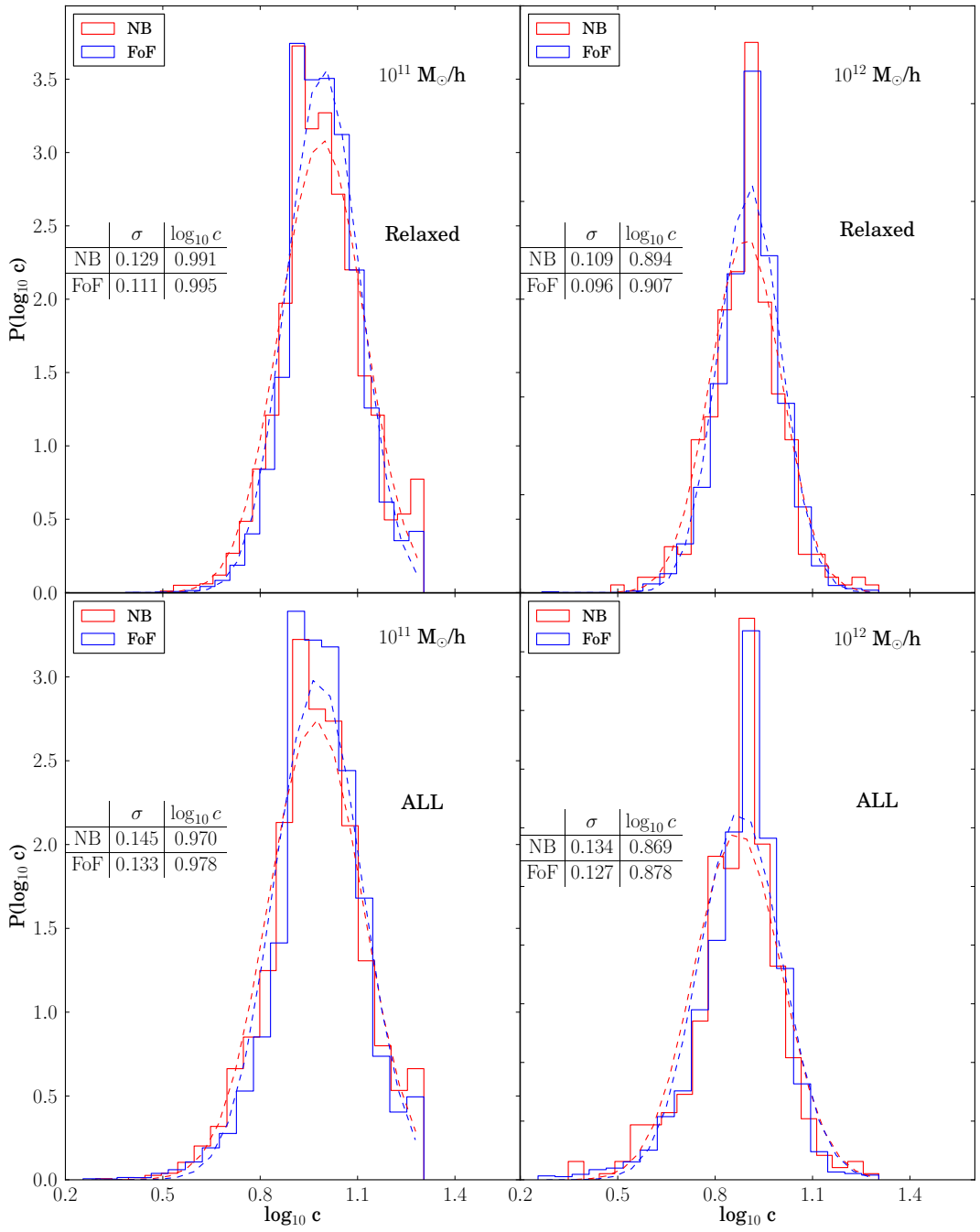

Figure 4.6: The distribution of concentrations for haloes in the two mass bins $10.75 < \log_{10} M_{200}/h^{-1}\mathrm{M}_\odot < 11.25$ and $11.75 < \log_{10} M_{200}/h^{-1}\mathrm{M}_\odot < 12.25$. The upper panels are for samples of relaxed haloes while the bottom panels are for all haloes whether or not they satisfy the relaxation criteria. In each panel the blue histogram is for FoF haloes and the red histogram is for Dhaloes that do not have bijective matches to FoF haloes. The smooth curves are log-normal approximations with the same $\log_{10} c$ and second moment, σ, as the measured distributions. The corresponding values of $\log_{10} c$ and σ are given in the legend.

4.5 Conclusions

We have used the high resolution Millennium Simulation II cosmological N-body simulation to quantify the properties of haloes defined by the Dhalo algorithm. This algorithm is designed to produce merger trees suitable for use with the semi-analytic galaxy formation model, GALFORM. We have presented the properties of the Dhaloes by comparing them with the corresponding properties of the much more commonly used FoF haloes (Davis et al., 1985).

Despite the complex mapping between FoF and Dhaloes, which results in a significant fraction of FoF haloes being broken up into multiple Dhaloes while other FoF haloes get (re)merged into a single Dhalo, we find that the overall mass functions of the two sets of haloes are very similar. The mass functions of our Dhalo and FoF halo catalogues are both reasonably well fit over the mass range of 10^8 to $10^{13.5}$ $h^{-1}\mathrm{M}_\odot$ by currently popular analytic mass functions such as those of Warren et al (2006) and Reed et al. (2007).

The high resolution of the Millennium II simulation has allowed us to study the density profiles and concentrations of both FoF and Dhaloes over a wide range of mass. To avoid contaminating our samples with unrelaxed haloes for which fitting smooth spherically symmetric profiles is inappropriate we exclude unrelaxed haloes following Neto et al. (2007). We find that 80% of both FoF and Dhaloes are relaxed according to these criteria. For FoF haloes we accurately reproduce the mass–concentration distribution found by Neto et al (2007) at high masses and extend the distribution to much lower masses. Combining our results with those of Macciò et al. (2007) and Neto et al. (2007), we find that a single power law reproduces the mass-concentration relation for over five decades in mass. We also find that the mass-concentration distributions for Dhaloes agree very accurately with those for FoF haloes. This is true even for non-bijective Dhaloes which are secondary components of FoF haloes. The properties of such haloes have generally been overlooked in previous studies. We show that the distributions of concentrations around the mean mass-concentration relation are well described by log-normal distributions for both the FoF and Dhaloes.

5

ORBITAL PARAMETERS OF INFALLING SATELLITE HALOES IN THE HIERARCHICAL ΛCDM MODEL

5.1 Introduction

In the current cosmological structure formation model, dark matter haloes grow by the merging of smaller systems (White & Rees, 1978; Davis et al., 1985), leading to hierarchical halo growth. Substructures that are accreted onto a host halo can survive for significant periods of time within the host halo (Chandrasekhar, 1943; Klypin et al., 1999; Moore et al., 1999; Binney & Tremaine, 2008; Boylan-Kolchin et al., 2008; Jiang et al., 2008). These substructures can host satellite gforalaxies, such as those found in the Local Group, and galaxy clusters. Thus, it is important to study the distribution of the initial orbital parameters of subhaloes at the time of infall as they represent the initial conditions which determine the later evolution of the substructures in their host haloes.

Semi-analytic models of galaxy formation rely on prescriptions for dynamical friction survival times and tidal stripping, (see Baugh 2006 for a review). Assuming the halo potential to be spherically symmetric, a satellite orbit can be defined by the plane of the orbit and two further parameters related to the energy and angular momentum such as circularity and pericentre. Previous authors have studied the distributions of such orbital parameters for substructures in numerical simulations (Tormen, 1997; Vitvitska et al., 2002; Benson, 2005;

Wang et al., 2005; Zentner et al., 2005; Khochfar & Burkert, 2006; Wetzel, 2011). Tormen (1997) investigated the infall of satellites into the haloes of galaxy cluster mass, and reported that more massive satellites move along slightly more eccentric orbits, with lower specific angular momentum and smaller pericentres. Benson (2005) presented evidence for a satellite mass dependence of the distribution of orbital parameters, but was unable to characterise these trends accurately due to the limited statistics. Apparently in slight contradiction, Wetzel (2011) reports that the orbital parameters do not significantly depend on the satellite halo mass but depend more on the host halo mass. These studies were hampered by limited dynamic range and sample size. The high resolution and large volume of the simulation we analyse allows us to quantify trends in both satellite and host halo mass.

The two parameters characterising a satellite orbit are, in general, correlated. Wetzel (2011) provides fits to circularity and pericentre, but he stopped short of examining correlations between these parameters which are important if one wants to select representative orbits from the distribution. Khochfar & Burkert (2006) found a tight correlation between pericentre and circularity. Tormen (1997); Gill et al. (2004); Benson (2005) also find correlations between orbital parameters.

In this Chapter, we investigate the correlations between different possible pairs of parameters. We show that to a good approximation total infall velocity and the fraction of this velocity which is in the radial direction are uncorrelated. We present fits to these and show that when transformed these fits provide accurate descriptions of the distributions of other choices of orbital parameters.

Most previous work has focused on orbits only at redshift $z = 0$, or on the satellites that are still identified at $z = 0$. In our work we focus on host haloes that exist at $z = 0$, but we analyse the orbits of all satellites that fall into the host halo after its formation (defined as when its main progenitor had half the final halo mass), regardless of whether the satellite is still identifiable at $z = 0$.

This Chapter is structured as follows. In Section 5.2, we briefly outline the methods including a detailed description of the N-body simulation, the iden-

tification of halo mergers and the measurement of orbital parameters. In Section 5.3, we present detailed analysis of the orbital parameters. We conclude in Section 5.4.

5.2 Methods

5.2.1 Simulation

Our analysis is based on the DOVE simulation, a ΛCDM cosmological dark matter only simulation of a periodic volume with side length 100 Mpc, with cosmological parameters adapted from the WMAP7 analysis of Komatsu et al. (2011). The Hubble parameter, density parameter, cosmological constant, scalar spectral index and linear rms mass fluctuation in 8 h^{-1}Mpc radius spheres were $H_0 = 70.4\ \mathrm{km\,s^{-1}}$, $\Omega_\mathrm{m} = 0.272$, $\Omega_\Lambda = 0.728$, $n_\mathrm{s} = 0.97$ and $\sigma_8 = 0.81$, respectively. The dark matter is represented by $N_\mathrm{p} = 1620^3$ particles of mass $m_\mathrm{p} = 8.8 \times 10^6\ \mathrm{M}_\odot$. Initial conditions were set up using second order Lagrangian perturbation theory (Jenkins, 2010), with phases set using the multiscale Gaussian white noise field *Panphasia* (Jenkins, 2013). These phases were chosen to be the same as in the EAGLE simulation (Schaye et al., 2014) and are fully specified by the *Panphasia* descriptor [Panph1,L16,(31250,23438,39063),S12,CH1050187043, EAGLE_L0100_VOL1]. The initial conditions were evolved to $z = 0$ using the GADGET3 N-body code, which is an enhanced version of the code described in Springel (2005b).

The particle positions and velocities were output at 160 snapshots, equally spaced in $\log(a)$ from $z = 20$. At each output, haloes were identified using a Friends-of-Friends algorithm (FoF; Davis et al., 1985), and the SUBFIND algorithm (Springel et al., 2001) was used to identify self-bound substructures ("subhaloes") within them. We define our FoF haloes by the conventional linking length parameter of $b = 0.2$ (the linking length is defined as b times the mean interparticle separation). Typically the main SUBFIND subhalo contains most of the mass of the original FoF halo, only unbound particles and those bound to

secondary subhaloes are excluded. We keep all haloes and subhaloes with more than 20 particles, corresponding to 2×10^8 $M_\odot$.

5.2.2 Orbital parameters

We define the virial mass, $M_{\rm vir}$, and associated virial radius, $r_{\rm vir}$, of a dark matter halo using a simple spherical overdensity criterion centred on the potential minimum:

$$M_{\rm vir} = \frac{4}{3}\pi\Delta\,\rho_{\rm crit}\,r_{\rm vir}^3 \tag{5.2.1}$$

where $\rho_{\rm crit}$ is the cosmological critical density and Δ is the specified overdensity. We adopt $\Delta = 200$ and include all the particles inside this spherical volume, not only the particles grouped by the adopted halo finder, to define the enclosed mass, M_{200}, and associated radius r_{200}. This choice of $\Delta = 200$ is largely a matter of convention, but has been shown roughly to correspond to the boundary at which the haloes are in approximate dynamical equilibrium (e.g. Cole & Lacey, 1996). We express velocities in units of the virial velocity, V_{200}, of the host halo.

For a spherical potential, the orbit of a satellite can be fully specified by the orientation of the orbit and two non-trivial parameters related to its energy, E, and the modulus of its angular momentum, J. There are various choices for these two parameters. The choice made by Benson (2005) and others of the radial, V_r, and tangential, V_θ, velocities at infall benefits from being directly measurable quantities and being simple. In contrast, Tormen (1997) adopted the *circularity*, defined as the total angular momentum in units of the angular momentum for a circular orbit of the same energy, $J/J_{\rm circ}(E)$, and the infall radius in units of the radius of a circular orbit of the same energy, $r/r_{\rm circ}(E)$. These have the advantage of depending only on the conserved quantities E and J (Note, the r here is the radius at infall and so equals r_{200} in our study.), but require adopting a model of the halo potential. The particular form of these two parameters is motivated by theoretical modelling including that of satellite orbital decay due to dynamical friction (Lacey & Cole, 1993; Jiang et al., 2008). To define these two parameters, we adopt a singular isothermal sphere (SIS) (Cole & Lacey, 1996) as a simple model for the density profile of dark matter haloes. This choice is consistent

with assumptions in Lacey & Cole (1993) and approximates the more realistic NFW potential Navarro et al. (1996b) over a large range of halo radii.

Here we present a derivation of the transformations between these two parametrisations (following e.g. van den Bosch et al 1999). Defining the zero point of the gravitational potential to be at r_{200}, where the circular velocity, V_{200}, is given by $V_{200} = \sqrt{GM_{200}/r_{200}}$, we can express the gravitational potential as

$$\phi(r) = V_{200}^2 \ln(r/r_{200}). \tag{5.2.2}$$

Thus, for a satellite crossing r_{200} with radial and tangential velocities, V_r and V_θ, the total energy per unit mass is

$$E = \frac{1}{2}\left(V_r^2 + V_\theta^2\right). \tag{5.2.3}$$

As the circular velocity is constant for a SIS, the radius, of a circular orbit of the same energy is given by

$$\frac{1}{2}\left(V_r^2 + V_\theta^2\right) = \frac{1}{2}V_{200}^2 + V_{200}^2 \ln(r_{\rm circ}/r_{200}), \tag{5.2.4}$$

implying

$$\frac{r_{\rm circ}(E)}{r_{200}} = \exp\left(\frac{V_r^2 + V_\theta^2 - V_{200}^2}{2V_{200}^2}\right). \tag{5.2.5}$$

As the circular velocity is constant the corresponding angular momentum of a circular orbit is $J_{\rm circ}(E) = M_{\rm s} V_{200} r_{\rm circ}(E)$, we have

$$\frac{J}{J_{\rm circ}(E)} = \frac{V_\theta}{V_{200}} \exp\left(-\frac{V_r^2 + V_\theta^2 - V_{200}^2}{2V_{200}^2}\right). \tag{5.2.6}$$

We also show the reduced mass SIS case in the Appendix A1 but we do not adopt it as the default as in reality the SIS halo will be deformed as the satellite orbits within it. Also we find for a SIS the reduced mass has little effect (see Fig. A.3).

Another useful quantity to define is the composite parameter

$$\Theta = \left(\frac{J}{J_{\rm circ}(E)}\right)^{0.78} \left(\frac{r_{200}}{r_{\rm circ}(E)}\right)^2. \tag{5.2.7}$$

Its utility is that Lacey & Cole (1993) showed that the orbital decay time of a satellite of mass $M_{\rm s}$ due to dynamical friction within a host halo of mass $M_{\rm h}$ is

given by

$$\tau_{\rm mrg} = \Theta\, \tau_{\rm dyn} \frac{0.3722}{\ln(\Lambda_{\rm coulomb})} \frac{M_{\rm h}}{M_{\rm s}}, \tag{5.2.8}$$

where $\tau_{\rm dyn}$ is the dynamical time of the host halo and $\ln(\Lambda_{\rm coulomb})$ is taken to be $\ln(M_{\rm h}/M_{\rm s})$. This formula assumes that the satellite can be treated as a point mass orbiting in a host halo with a SIS density profile and is valid when $\tau_{\rm mrg} \gg \tau_{\rm dyn}$. In this model it is only necessary to know the one-dimensional distribution of Θ values rather than the bivariate distribution of, say, V_r, and V_θ to determine the distribution of orbital decay times.

5.2.3 Identifying halo mergers

We follow the evolution, infall and merging of haloes and subhaloes using merger trees. Our starting point is the catalogue of FoF haloes and their constituent subhaloes at redshift zero. We build subhalo merger trees linking each subhalo to its progenitors and descendants using the algorithm described in Chapter 2. Next, we identify both the progenitors of the FoF haloes and the subhaloes which fall into them. For each FoF halo, we trace its progenitor in the previous snapshot by identifying the main progenitor of its main subhalo. We then define the virial radius of this progenitor halo such that a sphere of this radius centred on the particle at the potential minimum of the main subhalo encloses 200 times the critical density as defined in Eqn. 5.2.1. We trace the main progenitor of each redshift zero FoF halo back in this way until the last snapshot at which its mass is greater than half the final halo mass. We choose not to consider mergers before the formation time of the main halo as we bin our results by the halo mass at $z = 0$ and wish this to reflect (within a factor of two) the mass of the main halo when the merger takes place. To identify subhaloes that merge onto this main halo progenitor we not only trace the progenitors of subhaloes that are in the halo at redshift zero, but also those that were inside progenitors of the main halo at some point but which have since been disrupted, merged or escaped. Hence, we trace every individual subhalo from its formation redshift to the redshift when it first crosses the virial radius of the host halo.

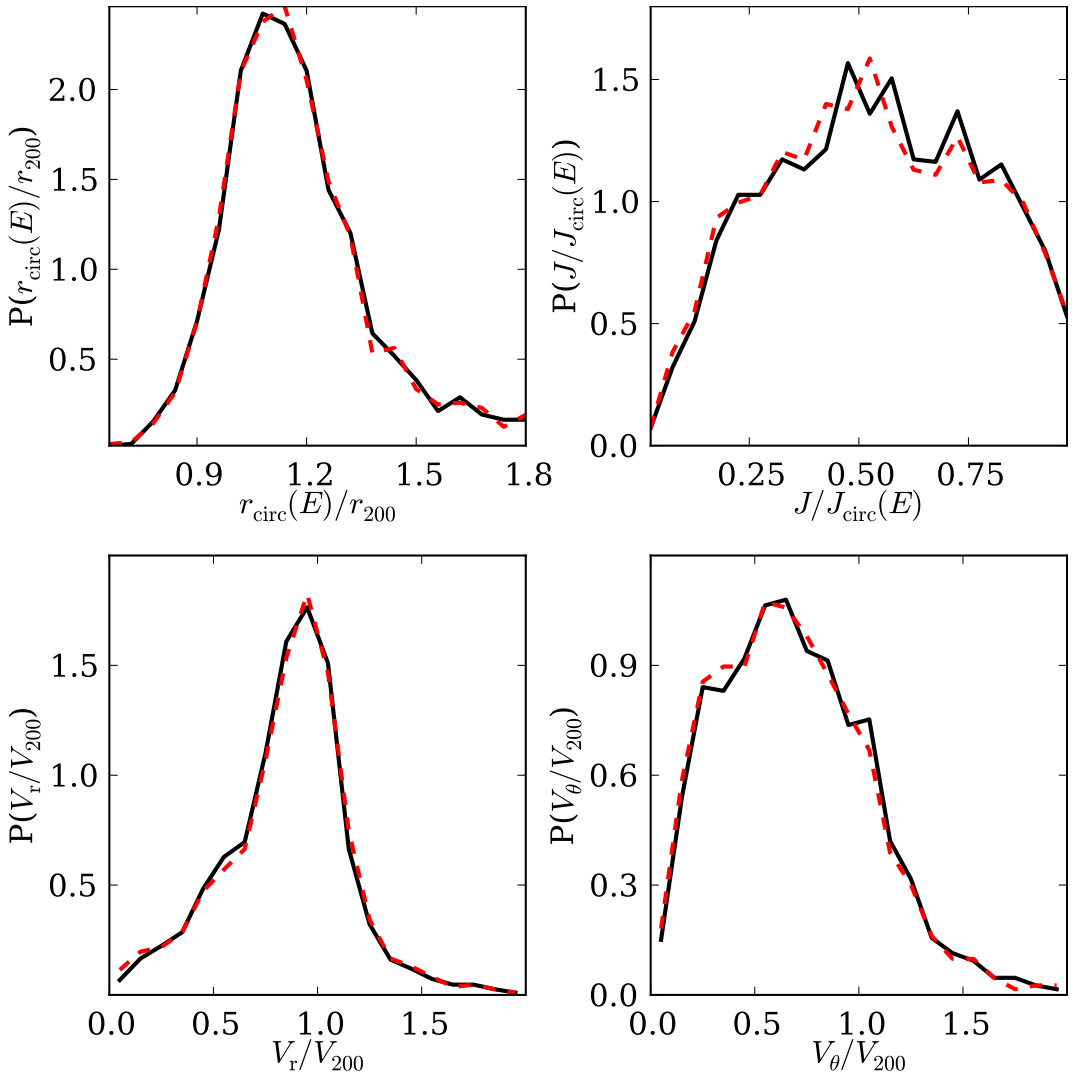

Figure 5.1: Tests of the interpolation scheme on the distributions of the orbital parameters $r_{\rm circ}(E)/r_{200}$, $J/J_{\rm circ}(E)$, V_r/V_{200}, V_θ/V_{200}. The panel shows the differential distribution of orbital parameters in the mass ratio bin: $M_{\rm s}/M_{\rm h} > 0.05$ for all the host haloes in our sample. Solid lines show the results using linear interpolation of energy and angular momentum, dashed lines show results using linear interpolation of velocity and position.

In order to find the precise crossing time, we save the orbital information from the snapshots just before and after a satellite subhalo crosses the virial radius. Then, we interpolate both the satellite position (relative to the halo centre) and the halo virial radius linearly to find the time when the subhalo first crosses the virial radius. To investigate the accuracy of the interpolation scheme we considered two methods of interpolating the satellite orbital parameters to this crossing time:

1. We interpolate the energy (using the singular isothermal sphere approximation of the halo potential described in Section 5.2.2) and angular momentum linearly in redshift to the crossing time. We then compute other orbital parameters such as the radial and tangential velocities from this interpolated energy and angular momentum.

2. Alternatively, we interpolate each component of the satellite's velocity linearly in redshift to the crossing time and then compute the required orbital parameters from the interpolated velocity and position.

Provided our simulation snapshots are sufficiently closely spaced, we would expect these two methods to give very similar results. This is indeed what we find as demonstrated in Fig. 5.1 which compares the distribution of the various orbital parameters for satellites satisfying $M_{\rm s}/M_{\rm h} > 0.05$ at the time of infall in our full sample of haloes. Throughout the rest of this chapter, we show results just from the method that linearly interpolates the energy and angular momentum. We would expect this to be the more accurate method as these two quantities are almost conserved and so only vary slowly with the interpolation parameter.

Accurately defining the orbital parameters at the crossing time is an important issue that has been considered in earlier work. The approach adopted by Benson (2005) and Vitvitska et al. (2002) was to search for pairs of haloes within some separation $r_{\rm max}$ which are about to merge and then predict their crossing time by modelling them as two isolated point masses. A similar approach was taken by Tormen (1997), Khochfar & Burkert (2006) and Wetzel

(2011). When using such schemes one must apply a weighting to correct for the under-representation of satellites with large infall velocities, some of which will be at separations greater than $r_{\rm max}$ at the earlier snapshot. In our work, due to the higher time resolution of our simulation outputs, we do not have to limit the separation between satellite and host halo at the snapshot prior to infall and instead form a complete census of all the infalling satellites.

I tested other interpolation schemes. In particular I tested cubic interpolation. For this I expressed the each Cartesian coordinate of the position of a satellite with respect to the halo centre as a cubic polynomial in time and determined its four coefficients using the position and velocity at the two adjacent snapshots. The cases where the linear and cubic interpolation schemes differed significantly I tracked down to where the motion of the centre of potential of a satellIite was not consistent with the mean velocity of the particles assigned to that satellite. This can happen as in some cases the population of particles belonging to a particular subhalo can change significantly between snapshots. Hence, linear interpolation was determiend to be more robust. To improve the interpolation we could either use a simulation with more frequent snapshots or a substructure finder that enforces more consistent membership.

5.2.4 Formation and infall redshifts

As we want our measured orbital parameter distributions to be directly applicable to semi-analytic galaxy formation models we trace all the infalling subhaloes back to the formation time of the main halo, where its formation time is defined as when its main progenitor has half the final, $z = 0$, halo mass. We bin our halo samples by their mass at redshift $z = 0$ and so by not tracing haloes back further in time we avoid significant ambiguity in the mass of the main halo at the time satellites are accreted, i.e. at all infall events the main halo is always within a factor of two the final halo mass. The distribution of halo formation redshifts, $z_{\rm HF}$, are shown in the top row of Fig. 5.2 for each of our final halo mass bins. As expected we see that lower mass haloes form earlier. The median formation redshift of our 10^{12}, 10^{13} and 10^{14} $\rm M_{\odot}$ haloes are 1.14, 0.92, and 0.66 respectively.

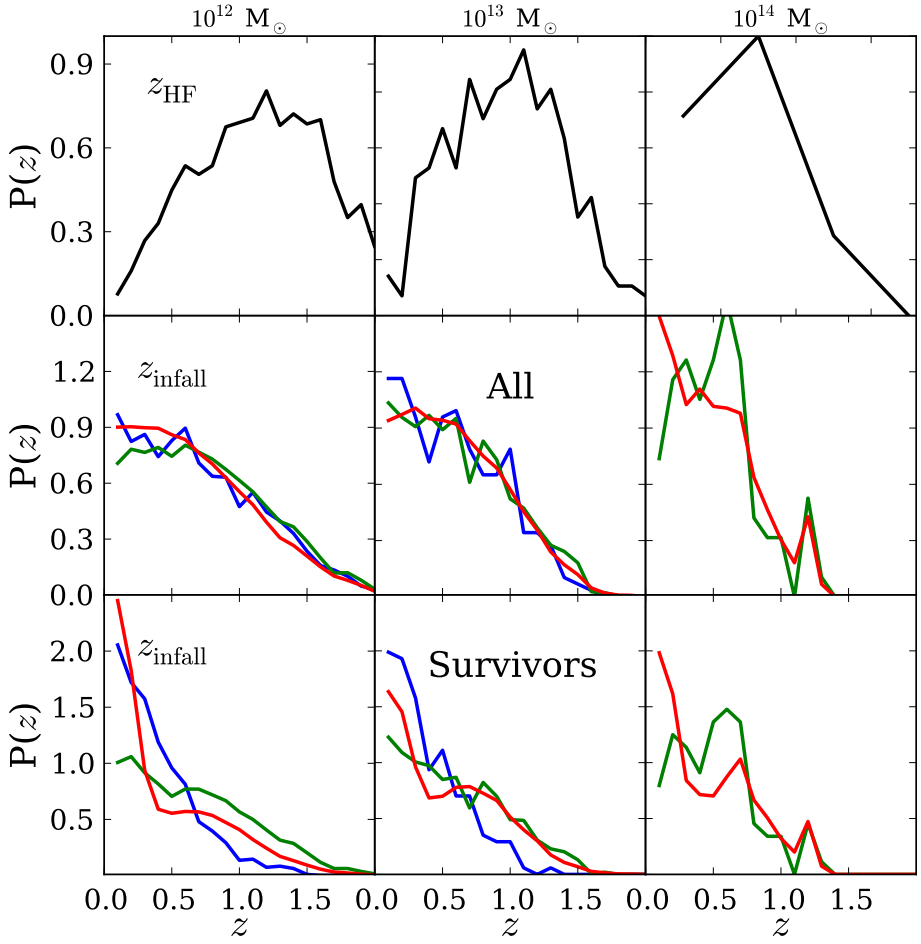

Figure 5.2: The distributions of halo formation redshifts and the redshifts at which satellites fall into these halos. Each column is for a fixed final halo mass as labelled at the top of the figure. The top row is the distribution of halo formation redshifts. The middle row is the distribution of satellite infall redshifts for all infalling satellites, while bottom row is for the subset of these satellites which survive as subhaloes at $z = 0$. In the bottom two rows the line colour indicates the satellite-to-host mass ratio. The red lines are for $0.0001 < M_{\rm s}/M_{\rm h} < 0.005$, green for $0.005 < M_{\rm s}/M_{\rm h} < 0.05$ and blue for $M_{\rm s}/M_{\rm h} > 0.05$.

The middle row of Fig. 5.2 shows the distribution of infall redshifts, $z_{\rm infall}$, split both by final halo mass and by the ratio of satellite-to-host mass at infall. These distributions rise steadily towards redshift $z = 0$ from the upper redshift set by when the first haloes in the sample form. The most interesting aspect is that infall redshift distribution at fixed halo mass is essentially independent of satellite-to-host mass ratio. This is equivalent to the mass distribution of the infalling satellites, measured in units of the host halo mass, being independent of redshift. Given that the distribution of host halo masses is constrained not to vary greatly with redshift (only haloes with mass greater than half the final mass are retained in the sample) then this behaviour is expected in simple excursion set models of hierarchical growth (Lacey & Cole, 1993).

The bottom row of Fig. 5.2 also shows distributions of infall redshifts, but now just for the satellites that survive and are identifiable at redshift $z = 0$. Contrasting these distributions with those from the middle row one clearly sees that the typical infall redshift of surviving satellites is significantly lower than that of the complete sample. This is, at least in part, a resolution effect as we are unable to identify satellites with fewer than 20 particles. Thus the shift to lower infall redshifts is greatest for the lowest mass satellites which are the ones with the smaller satellite-to-host mass ratio in the lower halo mass bins.

5.3 Orbital parameter distributions

5.3.1 Comparison to previous work

Fig. 5.3 compares our orbital parameter distributions with those from Tormen (1997), Benson (2005) and Wetzel (2011). In all panels, the black solid lines show the distributions for satellites with mass ratios in the range $0.05 < M_{\rm s}/M_{\rm h} < 0.5$ averaged over all our analysed haloes which span the mass range $5 \times 10^{11} < M_{\rm h} < 2.5 \times 10^{14} {\rm M}_{\odot}$. In general our results are in good agreement with these published datasets and those of Wang et al. (2005); Zentner et al. (2005); Khochfar & Burkert (2006), despite variations between these studies in the definition of

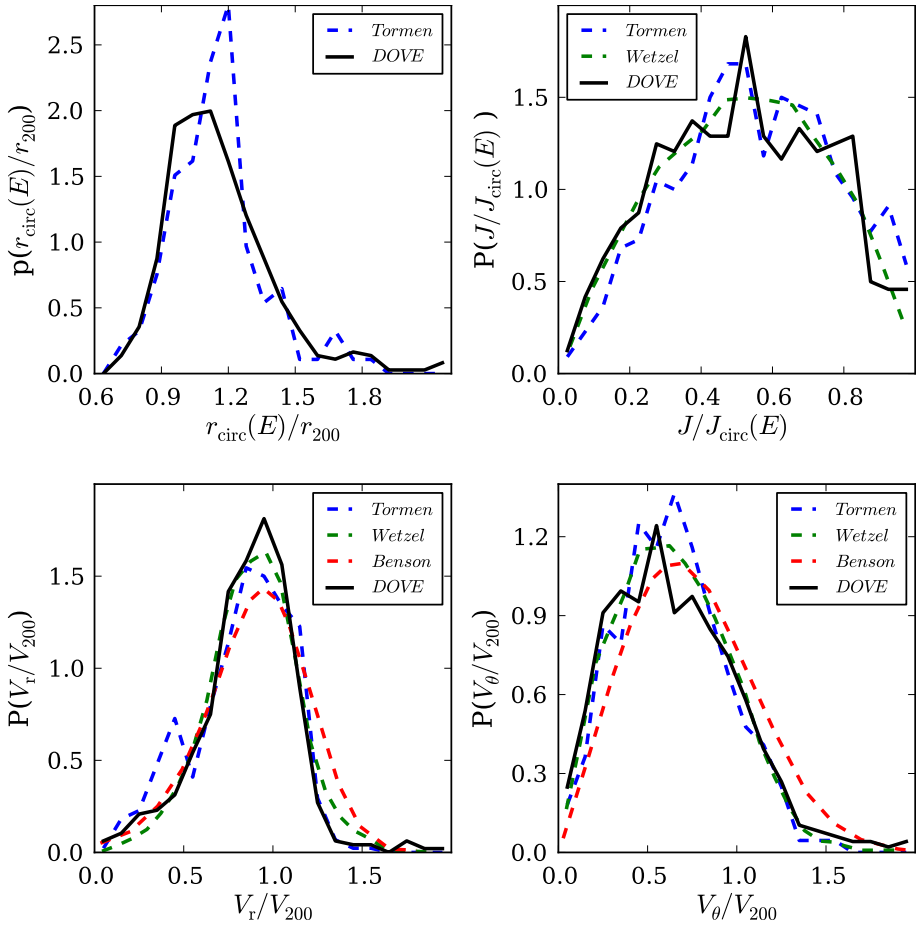

Figure 5.3: Comparison to published distributions of the orbital parameters $r_{\rm circ}(E)/r_{200}$, $J/J_{\rm circ}(E)$, $V_{\rm r}/V_{200}$, and V_θ/V_{200}. In all the panels the black solid line shows the distribution of the satellite orbital parameters for infalling satellites in our analysed host haloes (covering the mass range 5×10^{11} to 2.5×10^{14} $M_\odot$) with satellite-to-host halo mass ratios spanning 0.05 to 0.5. This range is typical of that probed by the samples to which we are comparing. Blue, green and red dashed lines show the results from the work of Tormen (1997), Wetzel (2011) and Benson (2005) respectively.

crossing time and the choice of cosmology.

The selection of Tormen (1997) data which we plot matches the $M_s/M_h > 0.05$[1] cut used in our own data, but is for host halos with typical masses of 10^{15} $M_\odot$. The good agreement we find with Tormen (1997) is only expected if, as we find below, the distributions depend only weakly on halo mass at fixed M_s/M_h. The Benson (2005) data is based on a wide range of simulations of different volumes and resolutions. In this sample he uses all satellites and haloes with masses greater than 10^{11} $M_\odot$ and states that the typical ratio $M_s/M_h = 0.08$. The smooth radial and tangential velocity distributions we plot in the lower panels of Fig. 5.3 are the fitted distributions presented by Benson (2005). Benson (2005) and also Vitvitska et al. (2002) modelled the radial distribution as a Gaussian and the tangential distribution as a Rayleigh or 2D Maxwell-Boltzmann distribution. The agreement with our results is reasonable. The radial and tangential velocity distributions of Wetzel (2011) are in very good agreement with our results. Like Benson, Wetzel uses all satellites and haloes above a fixed mass cut, 10^{10} $M_\odot$, and so we would expect the mean M_s/M_h ratio to be similar to that of Benson and to our $0.05 < M_s/M_h < 0.5$ sample. The comparison of $J/J_{\rm circ}(E)$ distributions between us and Wetzel is not strictly fair as we compute $J_{\rm circ}(E)$ using the singular isothermal sphere model while he models the satellite and host as two point masses. However while this introduces a bias for satellites for which $M_s/M_h \ll 1$, we find that the resulting distributions are very similar for satellites with $0.05 < M_s/M_h < 0.5$ (see Appendix A.1).

5.3.2 Orbital parameters: mass ratio and mass dependence

Fig 5.4 presents our results for the orbital parameter distributions for three bins of halo mass and three bins of satellite-to-host halo mass ratio. We reiterate that the host halo mass bins are defined by the mass of the host haloes at $z = 0$ while the mass ratio, M_s/M_h, is defined by the values at the infall redshift.

[1] We were able to apply this cut as G. Tormen kindly supplied his catalogue of satellite orbital parameters in electronic form.

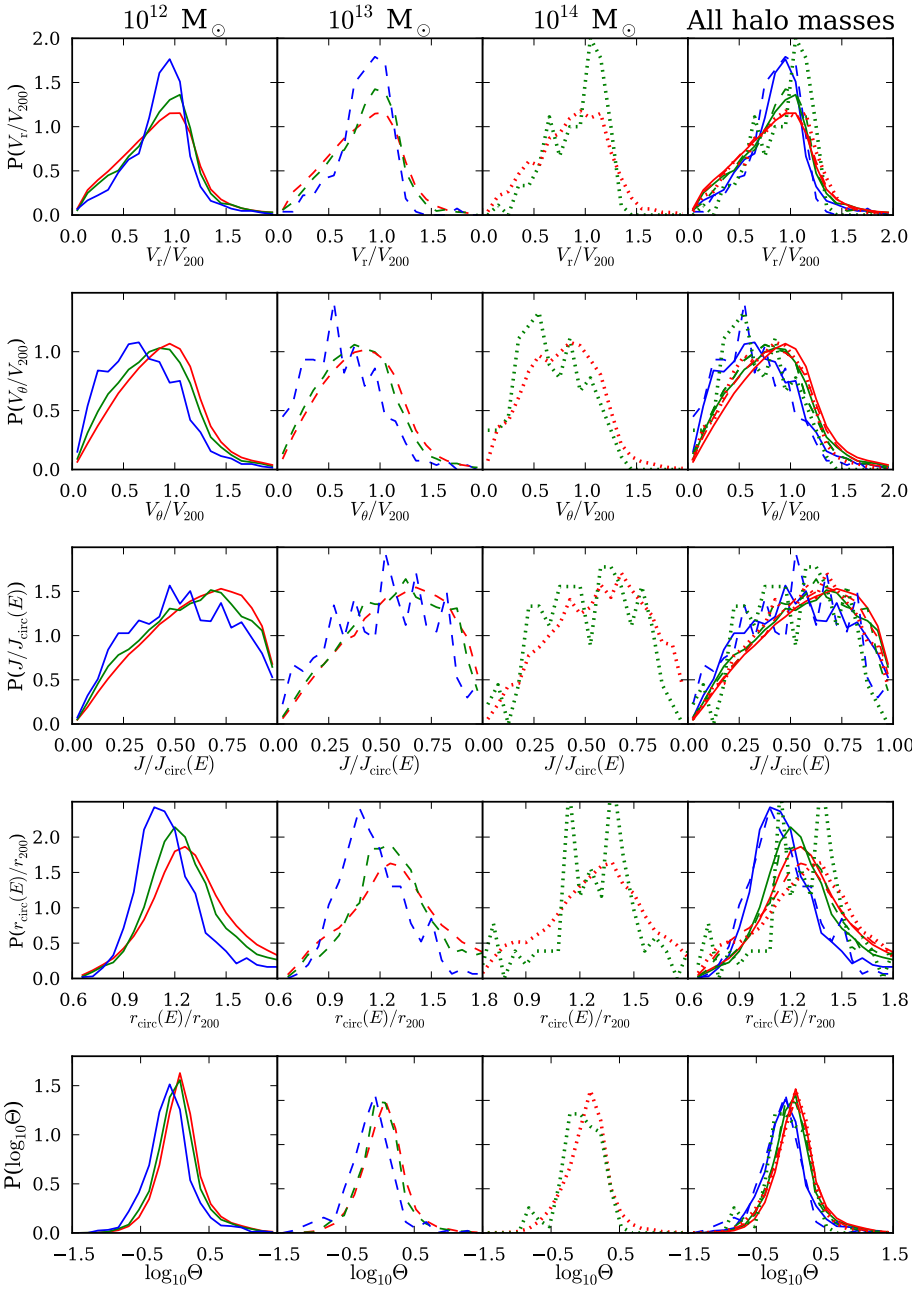

Figure 5.4: Orbital parameter distributions for bins of different final halo masses and satellite-to-host halo mass ratios, M_s/M_h. The central value of the final halo mass bin is indicated at the top of each column, with the rightmost column overplotting the results from each of the three mass bins using the appropriate line type. The red lines are for $0.0001 < M_s/M_h < 0.005$, green for $0.005 < M_s/M_h < 0.05$ and blue for $M_s/M_h > 0.05$. The first two rows show the radial, V_r/V_{200}, and tangential, $V_\cdot/V_{200}$, velocity distributions. The second two rows show the circularity, $J/J_{\rm circ}(E)$, and $r_{\rm circ}(E)/r_{200}$, while the final row shows the distributions of the composite parameter Θ defined in Eqn. 5.2.7. Note that for host haloes in the 10^{14} M$_\odot$ bin, we do not show the $M_s/M_h > 0.05$ distributions due to the low number of subhaloes.

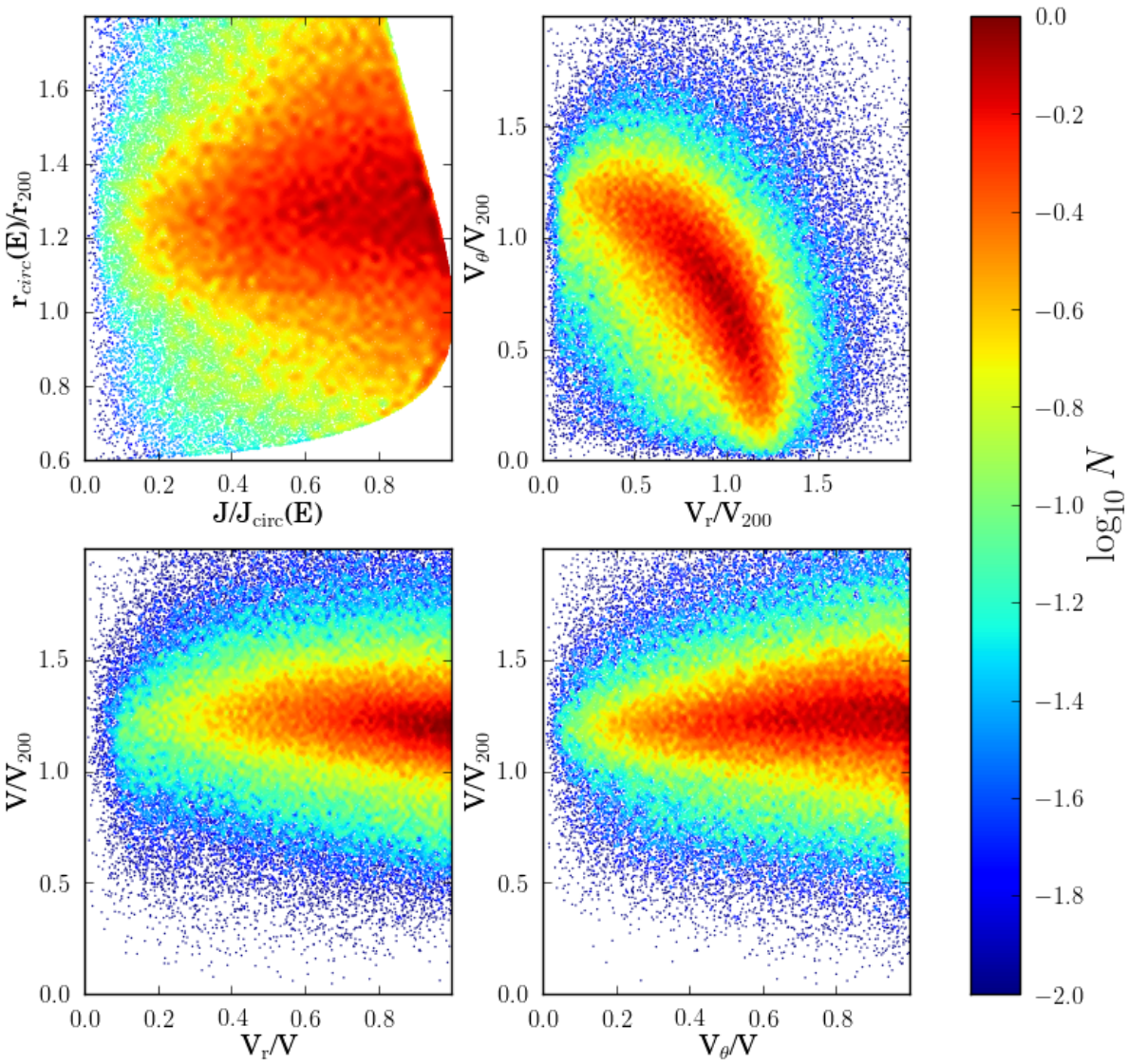

Figure 5.5: The bivariate distributions of orbital parameters for all satellites infalling onto 10^{13} $M_\odot$ haloes. The top panels show the two-dimensional distribution of $r_{\rm circ}(E)/r_{200}$ versus $J/J_{\rm circ}(E)$ and $V_\cdot/V_{200}$ versus $V_{\rm r}/V_{200}$ respectively. The bottom panels show the two-dimensional distributions of V/V_{200} versus V_r/V_{200} and V/V_{200} versus V_θ/V_{200}. The colour bar illustrates the relative density of points (on an arbitrary scale).

Table 5.1: Parameters of the fitted orbital parameter distributions for bins of final halo mass, $M_{\rm h}$, and the satellite-to-host mass ratio at infall, $M_{\rm s}/M_{\rm h}$. The notation for the parameters of the Voigt and exponential fitting functions are as defined in Eqns. 5.3.11 and 5.3.12

$M_{\rm h}/{\rm M}_\odot$	$M_{\rm s}/\ M_{\rm h}$	B	γ	σ	μ
10^{12}	$0.0001 - 0.005$	0.049 ± 0.055	0.109 ± 0.003	0.077 ± 0.002	1.220 ± 0.001
10^{13}	$0.0001 - 0.005$	0.548 ± 0.105	0.114 ± 0.010	0.094 ± 0.006	1.231 ± 0.002
10^{14}	$0.0001 - 0.005$	1.229 ± 0.292	0.110 ± 0.018	0.072 ± 0.007	1.254 ± 0.010
10^{12}	$0.005 - 0.05$	1.044 ± 0.086	0.098 ± 0.005	0.073 ± 0.004	1.181 ± 0.002
10^{13}	$0.005 - 0.05$	1.535 ± 0.255	0.087 ± 0.013	0.083 ± 0.010	1.201 ± 0.005
10^{14}	$0.005 - 0.05$	3.396 ± 1.040	0.050 ± 0.023	0.118 ± 0.025	1.236 ± 0.020
10^{12}	$0.05 - 0.5$	2.878 ± 0.200	0.071 ± 0.010	0.091 ± 0.007	1.100 ± 0.004
10^{13}	$0.05 - 0.5$	3.946 ± 0.578	0.030 ± 0.030	0.139 ± 0.021	1.100 ± 0.013
10^{14}	$0.05 - 0.5$	2.982 ± 4.646	-0.012 ± 0.035	0.187 ± 0.019	1.084 ± 0.052

The top two rows of Fig. 5.4 show the distributions of radial and tangential velocities at infall. The radial distributions peak close $V_{\rm r} = V_{200}$ and the tangential at a lower value of around $V_\theta = 0.65\,V_{200}$. Both distributions only have small tails beyond $1.5\,V_{200}$. Independently of host halo mass, we see that the distributions of radial velocities become broader for lower mass satellites with little change in the location of the peak of the distribution. In contrast for the tangential velocities the mode of the distribution shifts to higher values for less massive satellites. The most massive satellites are on the most radial, low angular momentum, orbits, The dependence of these distributions on halo mass at fixed $M_{\rm s}/M_{\rm h}$ is much weaker. This can be seen in the righthand panels where, to a first approximation, the lines of the same colour (same $M_{\rm s}/M_{\rm h}$) coincide. There is some residual dependence on halo mass (different line styles), with orbits becoming more radial – the V_θ/V_{200} distributions peaking at lower values – for more massive haloes, but this trend is much weaker.

The middle row of Fig. 5.4 shows the distributions of circularity, $J/J_{\rm circ}(E)$. The distributions are broad with those for the $M_{\rm s}/M_{\rm h} > 0.05$ bin peaking at close to a circularity of a half. In each bin of halo mass, we again see the trend,

for higher mass satellites to have less circular, more radially biased orbits. Also, once again, the trends with satellite-to-halo mass ratio are much stronger than those with halo mass.

The penultimate row of Fig. 5.4 shows the distributions of $r_{\rm circ}(E)/r_{200}$. This is essentially a measure of the energies of the orbits, with higher $r_{\rm circ}(E)/r_{200}$ corresponding to less bound orbits. At each halo mass, there is a strong trend for the more massive satellites to be more strongly bound. Again, the variations of the distributions with halo mass, at fixed satellite-to-halo mass ratio, are much weaker.

These trends are consistent with the observation that within the filaments of the cosmic web that surround an accreting dark matter halo, the most massive infalling haloes move along the central spines of the filaments. In this way the filamentary structures act as focusing rails which direct massive satellites onto predominantly radial orbits. Perhaps more simply, the force on the most massive satellites is dominated by the central halo while lower mass satellites can be significantly perturbed by other more massive satellites.

We show the distribution of the composite orbital parameter Θ in the bottom row of Fig. 5.4. We see a clear shift in the distributions towards higher values of Θ with decreasing values of $M_{\rm s}/M_{\rm h}$ and negligible dependence on host halo mass. According to Eqn. 5.2.8 this will contribute to lower mass satellites having longer merger timescales but this effect is subdominant to the explicit $M_{\rm h}/M_{\rm s}$ term in that equation which also acts in the same sense.

5.3.3 2D distribution of orbital parameters

As described in the Benson (2005) paper, the radial and tangential velocity distributions are tightly correlated. Consequently the 1-dimensional distributions presented in Fig. 5.4 are not a sufficient characterisation of the orbital parameter distributions. We emphasise this in Fig. 5.5 which shows bivariate distributions of various orbital parameter combinations.

The top left-hand panel of Fig. 5.5 shows the bivariate distribution of $r_{\rm circ}(E)/r_{200}$ and $J/J_{\rm circ}(E)$. The first thing to note in this distribution is that there are

excluded regions at high value of $J/J_{\rm circ}(E)$ both for low and high values of $r_{\rm circ}(E)/r_{200}$. These arise from our stipulation that we are characterising the orbits of satellites when they first cross r_{200}. The plotted distribution touches the right hand axis at $r_{\rm circ}(E)/r_{200} = 1$ and $J/J_{\rm circ}(E) = 1$. This point corresponds to a circular orbit with $r = r_{200}$. Circular orbits of either larger or smaller radius would not be included in our sample as they never cross r_{200}. Hence, $r_{\rm circ}(E)/r_{200}$ either increases or decreases away from unity for increasingly eccentric orbits in our sample. This defines the complicated boundary to the measured bivariate distribution.

The top right-hand panel of Fig. 5.5 shows the correlated bivariate distribution of radial and tangential velocities. This is similar to that presented and parametrised in Benson (2005). We note that the ridge line of this distribution is approximately circular, i.e. it corresponds to a fixed total velocity $V = \left(V_r^2 + V_\theta^2\right)^{1/2}$.

The lower two panels of Fig. 5.5 show the two dimensional distributions of the total velocity versus either the ratio V_r/V or V_θ/V. We see to a good approximation these pairs of parameters appear uncorrelated. This suggests that we can construct a simple model for the full bivariate distribution of orbital parameters by modelling the individual independent distributions of V/V_{200} and V_r/V. This will then provide a simple parametrised model that can be used in semi-analytic galaxy formation models.

5.3.4 Fitted distributions

To build a complete model of the bivariate distribution of parameters we perform fits to the marginalised distributions of both the total velocity, V/V_{200}, and the radial-to-total velocity ratio, V_r/V. Assuming these to be independent we can then transform variables to generate model predictions for the distributions of any of the other choices of orbital parameters such as $J/J_{\rm circ}(E)$ and $r_{\rm circ}(E)/r_{200}$. Here we present these fits as a function of halo mass and satellite-to-halo mass ratio.

The distributions of V/V_{200} for each of our samples are shown in Fig. 5.6

along with Voigt profile fits. The distributions of V/V_{200} are reasonably symmetric about their means but much more centrally peaked than Gaussians of the same rms width (leptokurtic). We find that the distributions can be fitted well by Voigt profiles, convolutions of a Lorentz profile,

$$P_L(x;\gamma) \equiv \frac{\gamma}{\pi(x^2+\gamma^2)}, \tag{5.3.9}$$

and a Gaussian

$$P_G(x;\sigma,\mu) \equiv \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \tag{5.3.10}$$

$$P_V(x;\sigma,\gamma,\mu) = \int_{-\infty}^{+\infty} P_G(x';\sigma,\mu) P_L(x-x';\gamma) dx' \tag{5.3.11}$$

where $x = V/V_{200}$. We determine the best fitting Voigt profiles by finding the parameters that maximise the likelihood, $\mathcal{L} = \Pi_i P_V(x_i;\sigma,\gamma,\mu)$, where the index i runs over all the satellites in the sample. The resulting fits are shown in Fig. 5.6 and their parameters σ, γ and μ are listed in Table. 5.1.

We find that the distributions V_r/V are well fit by exponential distributions of the form:

$$P(V_r/V) = A\left(\exp\left(\frac{BV_r}{V}\right) - 1\right). \tag{5.3.12}$$

Here A is simply a normalisation constant and B is the single free parameter. The distributions of V_r/V and the corresponding maximum likelihood fits are shown in Fig. 5.7. (In appendix A.2 we show that the distributions of V_θ/V can also be fit with this form of distribution function.) The distribution is almost linear, $B \ll 1$, for the combination of low $M_{\rm h}$ and low $M_{\rm s}/M_{\rm h}$. The distributions become increasingly radially biased, peaked at $V_r/V = 1$ (high B), for both increasing $M_{\rm s}/M_{\rm h}$ and $M_{\rm h}$, consistent with our earlier discussion.

The trends of the distributions of V/V_{200} and V_r/V with halo mass and satellite-to-halo mass ratio are depicted more clearly in Fig. 5.8, which shows all the fitted distributions on a single panel. In the lower panel we see the tendency for the distributions to become more radially biased for satellites with higher $M_{\rm s}/M_{\rm h}$. In the upper panel, it is clear that the V/V_{200} distributions have very little dependence on halo mass at fixed $M_{\rm s}/M_{\rm h}$. There is a stronger dependence on $M_{\rm s}/M_{\rm h}$ with samples of larger $M_{\rm s}/M_{\rm h}$ ratios having narrower distributions

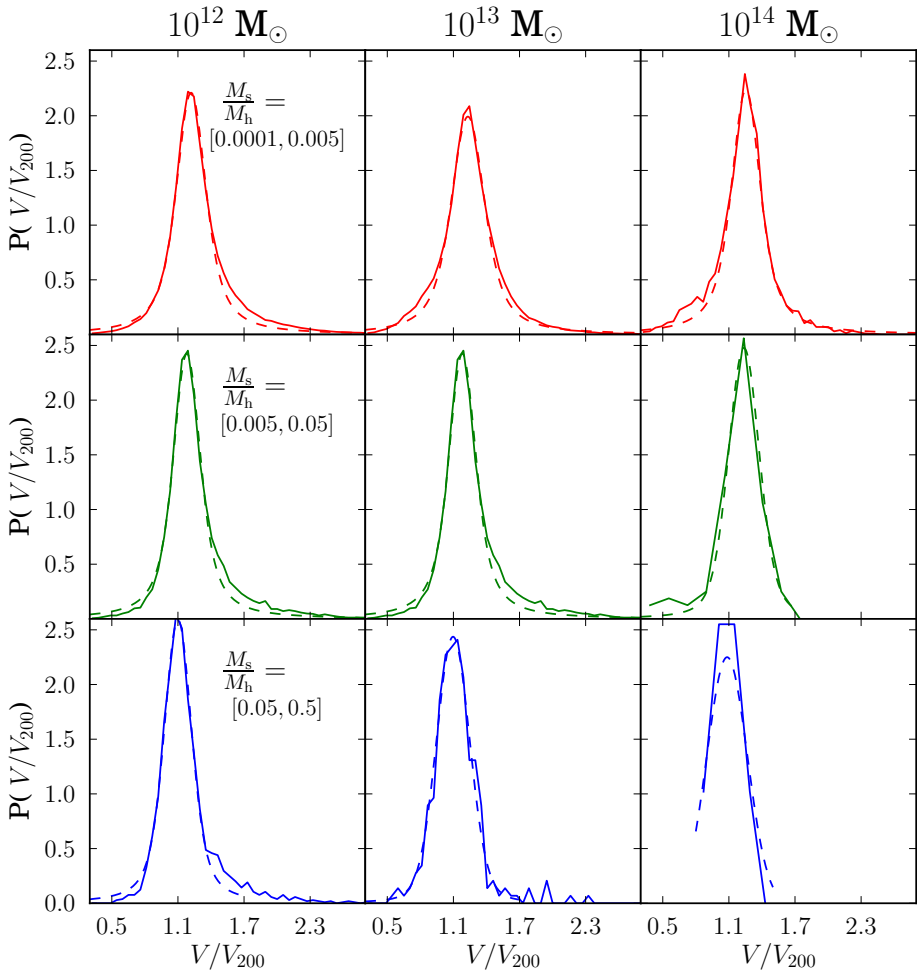

Figure 5.6: Probability distribution of the total infall velocity, V/V_{200}, as a function of both the satellite-to-host mass ratio at infall and the host halo mass. Each column is for a fixed final halo mass as labelled at the top of the column. Each row is for a different bin in satellite-to-host mass ratio: top (red lines) $0.0001 < M_s/M_h < 0.005$, middle (green lines) $0.005 < M_s/M_h < 0.05$ and bottom (blue lines) $M_s/M_h > 0.05$. The dashed lines are the Voigt profile fits whose parameters, μ, γ and σ are listed in Table 5.1.

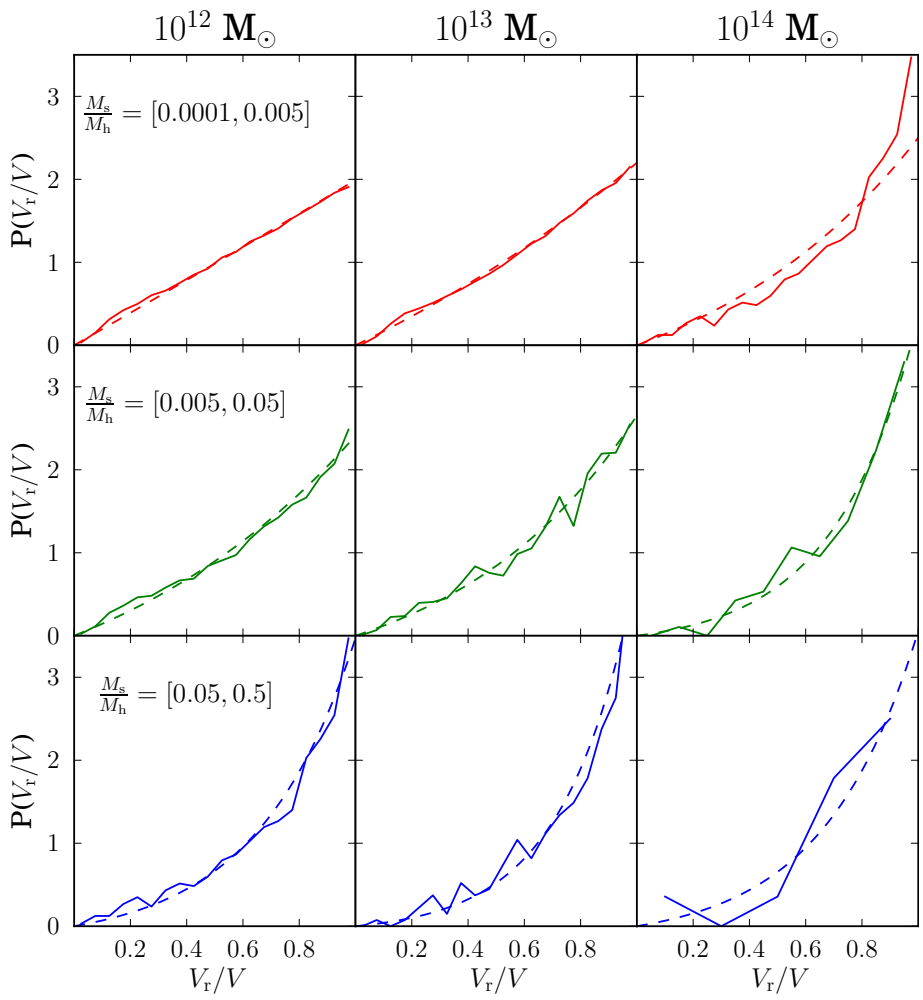

Figure 5.7: Dependence of the orbital parameters V_r/V on the mass ratio between the satellite halo mass and the host halo mass. Each column is for a fixed final halo mass as labelled at the top of the figure. Each row is for a different bin in satellite-to-host mass ratio, top (red lines) $0.0001 < M_s/M_h < 0.005$, middle (green lines) $0.005 < M_s/M_h < 0.05$ and bottom (blue lines) $M_s/M_h > 0.05$ The dashed curves are the best fitting exponential distributions and the corresponding value of the parameter B in Eqn. 5.3.12) is shown on each panel and in Table 5.1.

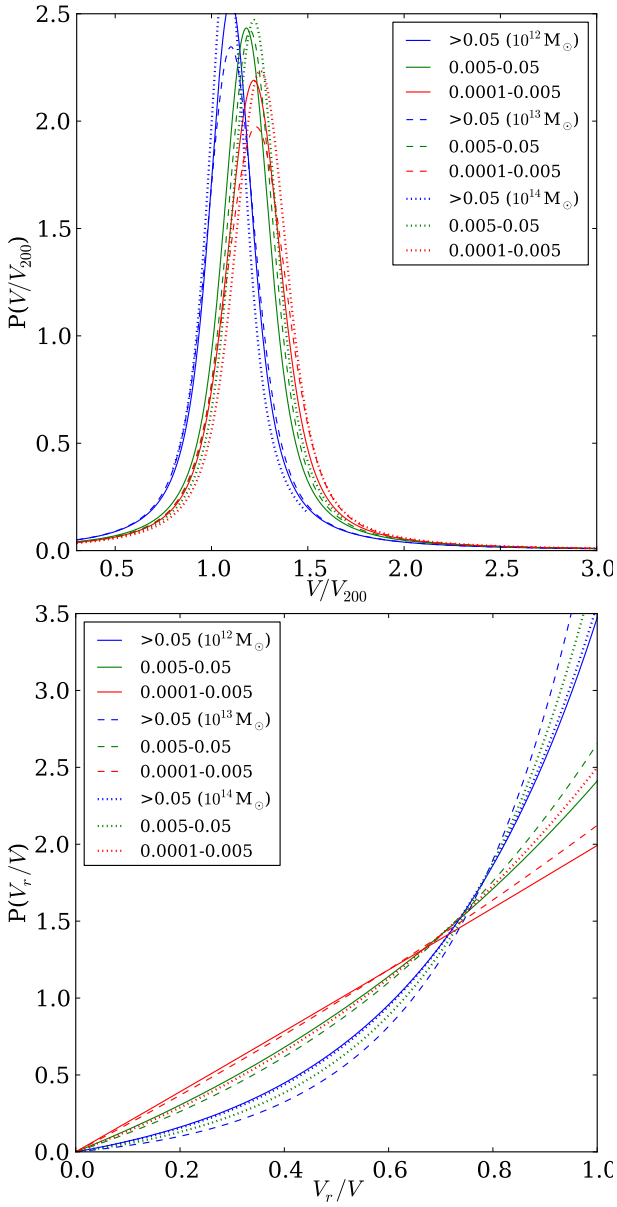

Figure 5.8: The fitted distributions of the orbital parameters V/V_{200} (top) and V_r/V (bottom) for the different values of both the satellite-to-host mass ratio and the host halo mass. Line colour denotes satellite-to-host mass ratio, red $0.0001 < M_s/M_h < 0.005$, green $0.005 < M_s/M_h < 0.05$ and blue $M_s/M_h > 0.05$. The line style indicates the host halo mass, solid 10^{12} M$_\odot$, dashed 10^{13} M$_\odot$ and dotted 10^{14} M$_\odot$.

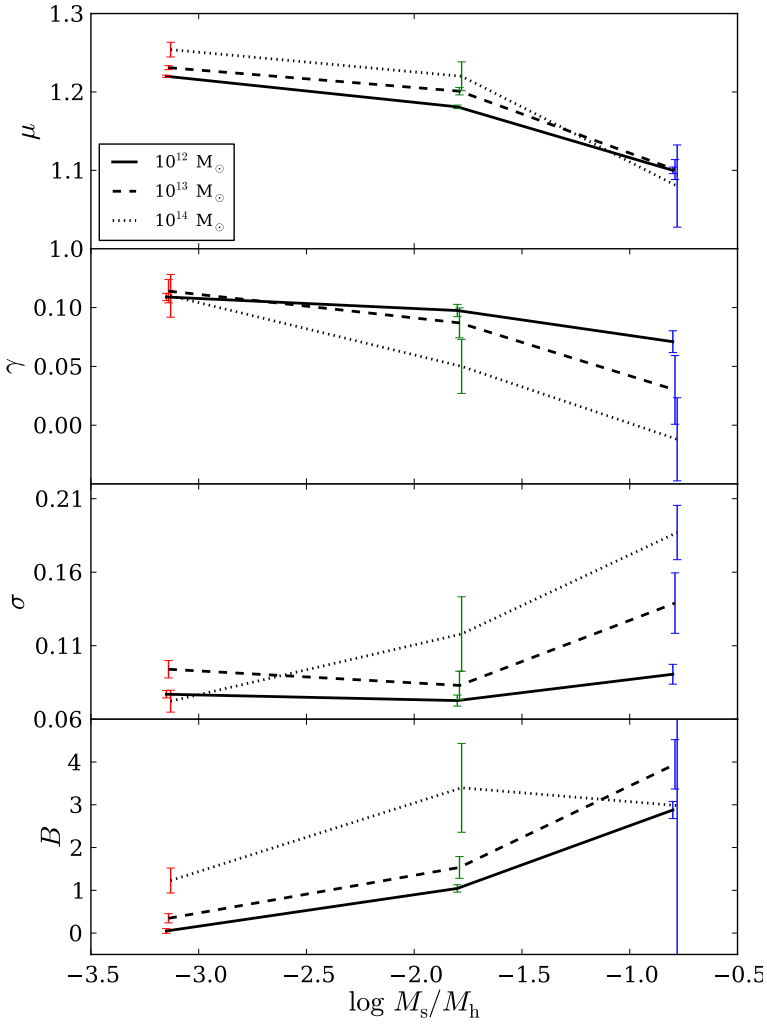

Figure 5.9: Dependence of fitting parameters μ, γ, σ and B on the satellite-to-host mass ratio at infall. The different line styles denote different host halo masses, as indicated in the legend. The colours of the error bars denote satellite-to-host mass ratio bins: red $0.0001 < M_\mathrm{s}/M_\mathrm{h} < 0.005$, green $0.005 < M_\mathrm{s}/M_\mathrm{h} < 0.05$ and blue $M_\mathrm{s}/M_\mathrm{h} > 0.05$. The errors are estimated by bootstrap sampling of the $z = 0$ halo catalogue.

and lower mean values. This is consistent with the similar trends in the distribution of $r_{\rm circ}(E)/r_{200}$ that we saw in Fig. 5.4. These trends can also be seen in Fig. 5.9, where we plot the dependence of the fit parameters on $M_{\rm s}/M_{\rm h}$. In all halo mass bins the mean, μ, decreases strongly for the highest values of $M_{\rm s}/M_{\rm h}$. The narrower width of the V/V_{200} distributions for high $M_{\rm s}/M_{\rm h}$, which we see in Fig. 5.8, is reflected in a decreasing value of γ (the width of the Lorentzian) with increasing $M_{\rm s}/M_{\rm h}$, which has greater effect on the width of the distribution than the corresponding slow increase in σ (the width of the Gaussian). The error bars shown on Fig. 5.9 have been estimated by bootstrap resampling of the $z = 0$ halo catalogue and we have investigated the correlations of all the pairs of parameters. The only significant correlation we find is an anticorrelation between σ and γ. This is to be expected as the overall width of the distribution is determined by $\sigma^2 + \gamma^2$, while their ratio, γ/σ, determines how peaked the distribution is (its kurtosis).

5.3.5 Derived distributions

If the fits we have presented in Section 5.3.4 are accurate and if V_r/V and V/V_{200} are uncorrelated then we can use these distributions to derive model distributions of any other choice of orbital parameter. For instance we can select pairs of values of V_r/V and V/V_{200} of a given probability from the fitted distributions and compute the radial and tangential velocities using

$$\frac{V_r}{V_{200}} = \left(\frac{V_r}{V}\right)\left(\frac{V}{V_{200}}\right) \tag{5.3.13}$$

and

$$\frac{V_\theta}{V_{200}} = \left(\frac{V}{V_{200}}\right)\sqrt{1-\left(\frac{V_r}{V}\right)^2}. \tag{5.3.14}$$

We can also derive $J/J_{\rm circ}(E)$, $r_{\rm circ}(E)/r_{200}$ and Θ from V_r/V and V/V_{200} using the equations in Section 5.2.2. We show all the resulting orbital parameter distributions in Fig. 5.10, which should be compared with Fig. 5.4. Direct comparison of the two figures shows that these are faithful representations of the data and validate the assumption that, to a good approximation, V_r/V and V/V_{200} can

be treated as independent random variables. The model distributions shown in Fig. 5.10, particularly the superimposed distributions in the righthand column, clearly show both the strong dependence on $M_{\rm s}/M_{\rm h}$ and the much weaker dependence on $M_{\rm h}$.

In the bottom panel, we compare our distributions of Θ to the lognormal distribution of Θ adopted in the GALFORM model (Cole et al., 2000). We see that relative to the GALFORM model our distributions are shifted slightly to higher values of Θ with this becoming more pronounced for massive host halos and for lower satellite-to-halo mass ratios. The shifts are not large but will lead to the GALFORM predicting slightly shorter merger timescales and thus few satellite galaxies in groups and galaxy clusters. This in turn will suppress the predicted clustering on small scales.

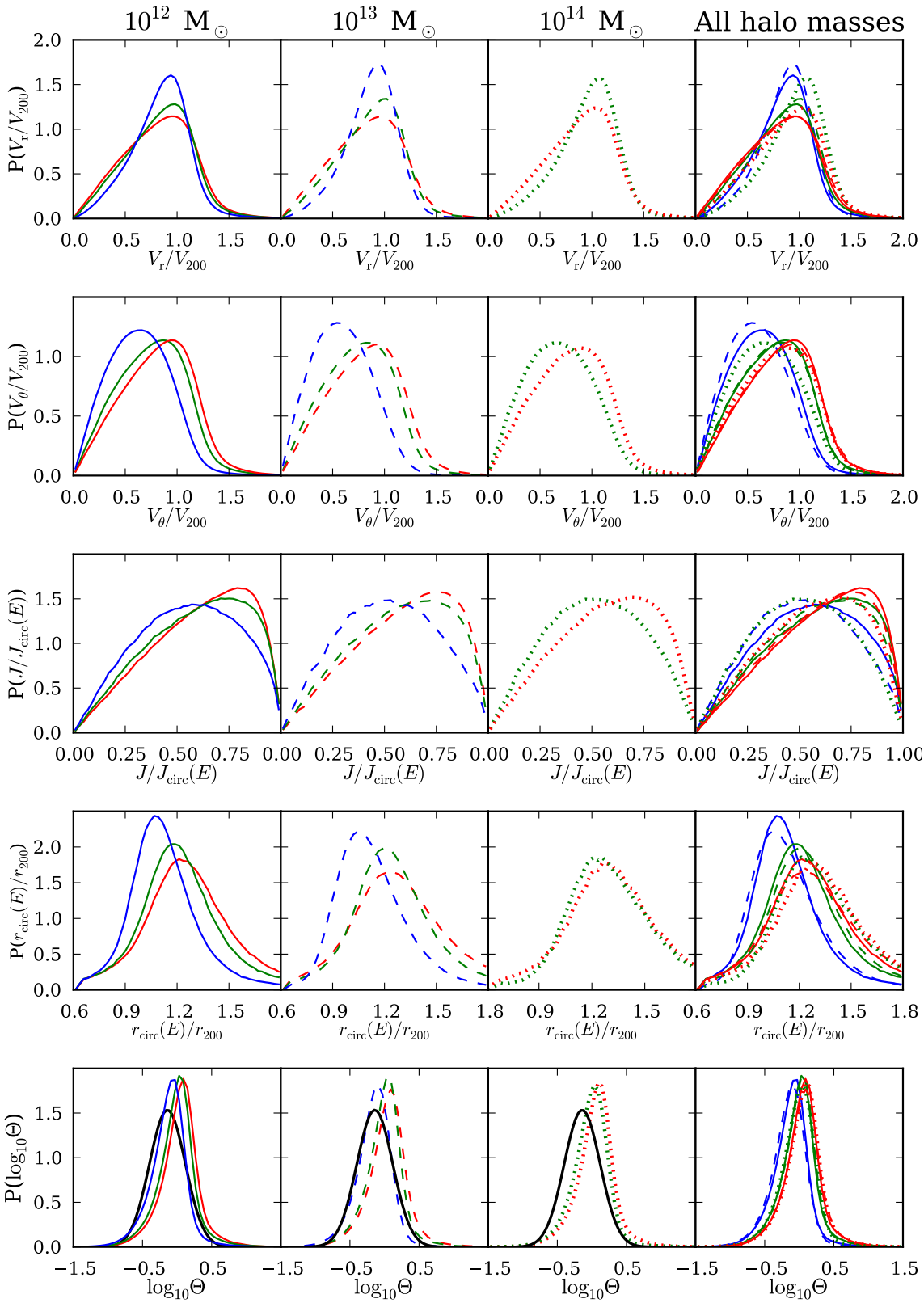

Figure 5.10: Like Fig. 5.4, but showing the distributions derived from the fits presented in Section 5.3.4 rather than the directly measured distributions. The black solid line in the bottom panel is the lognormal distribution from GALFORM (Cole et al., 2000)

5.4 Conclusions

We have employed the DOVE high resolution cosmological N-body simulation with more than 4 billion particles to study the distribution of the orbits of infalling satellites during hierarchical halo formation in the standard ΛCDM cosmology. We study host haloes with masses from 10^{12} to 10^{14} $M_\odot$ and satellites with masses as low as 2×10^8 $M_\odot$. Compared to previous studies (Tormen, 1997; Vitvitska et al., 2002; Benson, 2005; Wetzel, 2011) we have better mass and time resolution and a larger sample of satellite orbits.

There are various choices for the pair of orbital parameters that specify a satellite orbit in a spherical potential. We quantify the distributions of the radial, V_r, and tangential, V_θ, velocities as well as other common alternatives such as the circularity, $J/J_\mathrm{circ}(E)$ and the radius of the circular orbit of the same energy, $r_\mathrm{circ}(E)$.

We have examined the dependence of the distributions of these orbital parameters on both the host halo mass, M_h, and the mass ratio between the satellite and host, $M_\mathrm{s}/M_\mathrm{h}$. We find that the strongest trends are with $M_\mathrm{s}/M_\mathrm{h}$ at fixed M_h. Satellites with larger $M_\mathrm{s}/M_\mathrm{h}$ tend to be on more radial orbits with lower angular momentum and are more tightly bound. At fixed $M_\mathrm{s}/M_\mathrm{h}$ there is a trend for satellites around more massive haloes to also be on more radial orbits, but this trend is weaker. Insofar as previous authors have examined similar relationships, our results are consistent with their data. However, while Wetzel (2011) had not detected a significant dependence of orbital parameters on satellite mass ratio, possibly due to their limited sample size, our larger sample of orbits reveals a dependence, particularly at high mass ratios.

In general we find that complementary pairs of orbital parameters, such as (V_r,V_θ), are non-trivially correlated, making a complete description of their bivariate distribution complex. However, we find that, to a good approximation, the distributions of total infall velocity $V = (V_\mathrm{r}^2 + V_\theta^2)^{1/2}$ and the ratio V_r/V are uncorrelated. We present accurate Voigt and exponential fits to their respective distributions. Assuming them to be uncorrelated, we transform these simple

bivariate distributions and demonstrate that the distributions of other choices of orbital parameter can be successfully recovered.

6 CONCLUSION

My research has centred around establishing the nature of the dark matter haloes by investigating the abundance of haloes as a function of halo mass, the formation history of each halo, commonly called the merger tree, and the internal structure of the halo, in terms of radial density profiles and their angular momentum. During my PhD, I have done in-depth research on Dhalo merger trees which are used for driving the Durham semi-analytic galaxy formation model GALFORM, from the Millennium II Simulation (MSII) Boylan-Kolchin et al. (2009) and orbital parameters in a high resolution N-body cosmological simulation focusing on the orbits of infalling dark matter substructures.

6.1 N-body halo merger trees

To model galaxy formation, one must first have an accurate model of the evolution of dark matter haloes. There are two approaches to generating the merger histories of dark matter haloes. One is the Monte-Carlo approach which samples the distribution of progenitor masses predicted from the extended Press-Schechter theory (Lacey & Cole, 1993), and the other method uses halo merger trees extracted from N-body simulations which have sufficiently frequent outputs. The main drawback of N-body merger trees is their finite mass resolution. But it is now quite common for semi-analytic models to use halo merger trees due to the improvement of the resolution of N-body simulations like MSII. The key point for the N-body halo merger tree approach is to provide halo merger

trees which quantify the hierarchical growth of individual dark matter haloes, it is important to adopt models of the individual haloes that are consistent with the haloes defined in the merger trees. In the Durham GALFORM model, we use the Dhalo algorithm which analyses the N-body simulation snapshots to define a self-consistent set of dark matter haloes and the merger trees describing their hierarchical evolution. The advantages of merger trees based on Dhalo rather than the more common FoF haloes is that the Dhalo algorithm maintains as distinct haloes objects that can be prematurely linked into a single FoF group by tenuous bridges of particles or by the onset of the overlap of their outer diffuse haloes. The Dhalo algorithm has been widely exploited in a range of applications (Bower et al., 2006), but there has been no previous literature on the details of Dhalo properties such as the halo abundance and internal structure.

During my PhD, I have used the high resolution MSII cosmological N-body simulation to quantify the properties of haloes defined by the Dhalo algorithm. I have built Dhalo catalogues from the subhalo merger trees, and found accurate matches between FoF haloes and Dhaloes. I have conducted a detailed study of the abundance of Dhaloes of different masses and the internal structure of Dhaloes.

I have shown that unlike the FoF algorithm, the mass of a Dhalo correlates much more closely with the standard virial mass, M_{200}, than the corresponding FoF mass. I have also illustrated how some Dhaloes can be composed of more than one FoF halo. This occurs as structure formation in CDM models is not strictly hierarchical and occasionally a halo, after falling into a more massive halo, may escape to beyond the virial radius of the more massive halo. For the purposes of the GALFORM semi-analytic model it is convenient to consider such haloes as remaining as satellites of the main halo. I find that such remerged FoF haloes are not uncommon, but contribute very little mass to the larger haloes to which they are (re)attached. Where a FoF halo is split into two or more Dhaloes, I find that the secondary Dhaloes have a lower median ratio between the mass of the Dhalo and the virial mass. This is due to the SUBFIND substructure finder assigning less mass to subhaloes when they move into overdense en-

vironments. Despite the complex mapping between FoF and Dhaloes, I find that the overall mass functions of the two sets of haloes are very similar. With the improvement of mass resolution of MSII over the Millennium Simulation (Springel, 2005a), I am able to study the density profile and concentrations of both FoF and Dhaloes over a wide range of mass. I find the mass-concentration distributions for Dhaloes agree very accurately with those for FoF haloes. This is true even for the Dhaloes which are secondary components of FoF haloes. The properties of such haloes have generally been overlooked in previous studies. I also show that the distributions of concentrations around the mean mass-concentration relation are well decribed by log-normal distributions for both the FoF and Dhaloes.

6.2 Orbital parameters in the DOVE simulation

In current cosmological models, dark matter haloes grow via the merging of smaller systems, leading to hierarchical halo growth. The substructures that merge to become part of the host halo can survive for significant periods of time within the host halo. These substructures can host satellite galaxies, such as those found in the Local Group, and galaxy clusters. Thus, it is important to study the distribution of the initial orbital parameters of subhaloes at the time of merging into the host haloes as it represents the initial conditions which determine the later evolution of the substructures in the host.

Having gained a comprehensive understanding of N-body merger trees, I have started to study the orbital parameters of the dark matter substructures at the time of merging into their host halo. I use a high resolution N-body Simulation of the Standard ΛCDM cosmology, Dove (Dark Matter Only Version of Eagle), which has been carried out in Durham. To measure the orbital parameters, I make use of the 160 N-body snapshot outputs from the DOVE simulation which span a range of redshifts. The dark matter haloes in the simulation are identified by the FoF algorithm. Starting from the haloes at redshift zero, I find all the progenitors of subhaloes in the haloes by linking them through subhalo merger trees at each redshift. Then I trace every individual subhalo from its

earliest redshift to the redshift when it first crosses the virial radius of the host halo. I store the position and velocity of the subhalo for both the redshift when it first crosses the virial radius and the one before its first crossing. Then I find the crossing time by doing linear interpolation. I calculate the radial and tangential velocities, and circularity, by using the interpolated energy and angular momentum. I find how these orbital parameters depend on the subhalo mass, the mass ratio between the subhalo and the host halo, and the redshift in three halo mass ranges, which are $10^{12}\mathrm{M}_{\odot}$, $10^{13}\mathrm{M}_{\odot}$, $10^{14}\mathrm{M}_{\odot}$.

6.3 Future work

In the first project, although detailed comparisons between haloes defined by the FoF algorithm and the Dhalo algorithm have been conducted, it would also be interesting to test whether the properties of Dhaloes relative to FoF haloes, and in terms of their $M_{\mathrm{Dhalo}}/M_{200}$ distributions, are completely generic or have any dependence on cosmological parameters or redshift. Another issue is the appropriate choice of the virial radius of dark matter haloes. I used the conventional choice of r_{200} which Cole & Lacey (1996) showed was appropiate for an $\Omega = 1$ cosmology. However they also pointed out the need for detailed analysis of haloes formed in models with a cosmological constant. Hence it would be extremely interesting to investigate these issues using other recent cosmologies such of that specified by the WMAP7 parameter values. This could be done using the DOVE simulation which has the WMAP7 cosmology. I am also very interested in possible redshift dependences between different halo definitions. The different halo definitions will affect the halo mass and hence the properties of the galaxies that form in semi-analytic models such as GALFORM.

In my second project, the methods can be improved upon in several ways. A higher time-resolution simulation would allow a better determination of the accuracy of the crossing time. Alternatively one might be able to improve the accuracy using cubic rather than linear interpolation in determining the crossing time. The trends with redshift and cosmological parameters could also be

examined in the future work.

BIBLIOGRAPHY

Audit E., Teyssier R., Alimi J.-M., 1998, A&A, 333, 779

Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101

Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., cole S., 2005, MNRAS, 356, 1191

Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013, ApJ, 763, 18

Benson A. J., 2005, MNRAS, 358, 551

Benson A. J., Bower R., 2010, MNRAS, 405, 1573

Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587

Berlind A. A. et al., 2003, ApJ, 593, 1

Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122

Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press

Birkedal A., Noble A., Perelstein M., Spray A., 2006, Phys. Rev. D, 74, 035002

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

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645

Boylan-Kolchin M., Ma C.-P., Quataert E., 2008, MNRAS, 383, 93

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150

Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559

Bundy K., Ellis R. S., Conselice C. J., 2005, ApJ, 625, 621

Busha M. T., Alvarez M. A., Wechsler R. H., Abel T., Strigari L. E., 2010, ApJ, 710, 408

Cacciato M., van den Bosch F. C., More S., Mo H., Yang X., 2013, MNRAS, 430, 767

Chandrasekhar S., 1943, ApJ, 97, 255

Cole S., 1991, ApJ, 367, 45

Cole S., Lacey C., 1996, MNRAS, 281, 716

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168

Colless M. et al., 2001, MNRAS, 328, 1039

Courtin J., Alimi J.-M., Rasera Y., Corasaniti P.-S., Fuzfa A., Boucher V., 2010, in Alimi J.-M., Fuözfa A., eds, American Institute of Physics Conference Series Vol. 1241, American Institute of Physics Conference Series. pp 804–810

Crocce M., Fosalba P., Castander F. J., Gaztañaga E., 2010, MNRAS, 403, 1353

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Dodelson S., Widrow L. M., 1994, Physical Review Letters, 72, 17

Einasto J., Klypin A. A., Saar E., Shandarin S. F., 1984, MNRAS, 206, 529

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

Eisenstein D. J. et al., 2005, ApJ, 633, 560

Eke V. R. et al., 2004, MNRAS, 355, 769

Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 554, 114

Evrard A. E. et al., 2008, ApJ, 672, 122

Evrard A. E. et al., 2002, ApJ, 573, 7

Font A. S. et al., 2011, MNRAS, 417, 1260

Font A. S. et al., 2008, MNRAS, 389, 1619

Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, ApJ, 327, 507

Friedmann A., 1922, Zeitschrift fur Physik, 10, 377

Gamow G., 1948, Physical Review, 74, 505

Gao L., Navarro J. F., Cole S., Frenk C. S., White S. D. M., Springel V., Jenkins A., neto A. F., 2008, MNRAS, 387, 536

Gariazzo S., Giunti C., Laveder M., 2013, Journal of High Energy Physics, 11, 211

Gill S. P. D., Knebe A., Gibson B. K., 2005, MNRAS, 356, 1327

Gill S. P. D., Knebe A., Gibson B. K., Dopita M. A., 2004, MNRAS, 351, 410

Giocoli C., Pieri L., Tormen G., 2008, MNRAS, 387, 689

Gonzalez-Garcia M. C., Nir Y., 2003, Reviews of Modern Physics, 75, 345

Gottlöber S., Yepes G., 2007, ApJ, 664, 117

Grego L., Carlstrom J. E., Joy M. K., Reese E. D., Holder G. P., Patel S., Cooray A. R., Holzapfel W. L., 2000, ApJ, 539, 39

Guo Q. et al., 2011, MNRAS, 413, 101

Han J., Jing Y. P., Wang H., Wang W., 2012, MNRAS, 427, 2437

Harker G., Cole S., Helly J., Frenk C., Jenkins A., 2006, MNRAS, 367, 1039

Hatton S., Devriendt J. E. G., Ninin S., Bouchet F. R., Guiderdoni B., Vibert D., 2003, MNRAS, 343, 75

Heitmann K., Lukić Z., Habib S., Ricker P. M., 2006, ApJ, 642, L85

Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., 2003, MNRAS, 338, 903

Hernquist L., 1990, ApJ, 356, 359

Herrero M. J., 1998, ArXiv High Energy Physics - Phenomenology e-prints

Huchra J. P., Geller M. J., 1982, ApJ, 257, 423

Israel H., Schellenberger G., Nevalainen J., Massey R., Reiprich T., 2014, ArXiv e-prints

Jenkins A., 2010, MNRAS, 403, 1859

Jenkins A., 2013, MNRAS, 434, 2094

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Jiang C. Y., Jing Y. P., Faltenbacher A., Lin W. P., Li C., 2008, ApJ, 675, 1095

Jiang F., van den Bosch F. C., 2014, MNRAS, 440, 193

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

Khochfar S., Burkert A., 2006, A&A, 445, 403

Kim H.-S., Baugh C. M., Benson A. J., cole S., Frenk C. S., Lacey C. G., Power C., Schneider M., 2011, MNRAS, 414, 2367

Klypin A., Gottlöber S., Kravtsov A. V., Khokhlov A. M., 1999, ApJ, 516, 530

Knebe A. et al., 2011, MNRAS, 415, 2293

Knebe A. et al., 2013, MNRAS, 435, 1618

Knollmann S. R., Knebe A., 2009, ApJS, 182, 608

Kolb E. W., Turner M. S., 1990, The early universe.

Komatsu E. et al., 2009, ApJS, 180, 330

Komatsu E. et al., 2011, ApJS, 192, 18

Koposov S. E., Yoo J., Rix H.-W., Weinberg D. H., Macciò A. V., Escudé J. M., 2009, ApJ, 696, 2179

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

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

Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011, MNRAS, 418, 1649

Linder E. V., Jenkins A., 2003, MNRAS, 346, 573

Łokas E. L., Bode P., Hoffman Y., 2004, MNRAS, 349, 595

Ludlow A. D., Navarro J. F., Springel V., Jenkins A., Frenk C. S., Helmi A., 2009, ApJ, 692, 931

Lukić Z., Reed D., Habib S., Heitmann K., 2009, ApJ, 692, 217

Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, MNRAS, 391, 1940

Macciò A. V., Dutton A. A., van den Bosch F. C., Moore B., Potter D., Stadel J., 2007, MNRAS, 378, 55

Macciò A. V., Kang X., Fontanot F., Somerville R. S., Koposov S., Monaco P., 2010, MNRAS, 402, 1995

Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, MNRAS, 432, 1544

Merson A. I. et al., 2013, MNRAS, 429, 556

Molnar S. M., Birkinshaw M., Mushotzky R. F., 2002, ApJ, 570, 1

Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Moreno J., Giocoli C., Sheth R. K., 2008, MNRAS, 391, 1729

Muñoz J. A., Madau P., Loeb A., Diemand J., 2009, MNRAS, 400, 1593

Muldrew S. I., Pearce F. R., Power C., 2011, MNRAS, 410, 2617

Murray S. G., Power C., Robotham A. S. G., 2013, MNRAS, 434, L61

Navarro J. F., Frenk C. S., White S. D. M., 1995, MNRAS, 275, 720

Navarro J. F., Frenk C. S., White S. D. M., 1996a, ApJ, 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1996b, ApJ, 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Neistein E., Dekel A., 2008, MNRAS, 388, 1792

Neto A. F. et al., 2007, MNRAS, 381, 1450

Norberg P. et al., 2002, MNRAS, 332, 827

Parkinson H., cole S., Helly J., 2008, MNRAS, 383, 557

Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144

Percival W. J. et al., 2001, MNRAS, 327, 1297

Percival W. J., cole S., Eisenstein D. J., Nichol R. C., Peacock J. A., Pope A. C., Szalay A. S., 2007, MNRAS, 381, 1053

Percival W. J. et al., 2010, MNRAS, 401, 2148

Perkins D., 2008, Particle Astrophysics, Second Edition. Oxford Master Series in Physics, OUP Oxford

Perlmutter S. et al., 1999, ApJ, 517, 565

Pipino A., Pierpaoli E., 2010, MNRAS, 404, 1603

Planck Collaboration et al., 2014b, A&A, 571, A16

Planck Collaboration et al., 2014a, A&A, 571, A20

Power C., Knebe A., Knollmann S. R., 2012, MNRAS, 419, 1576

Press W. H., Davis M., 1982, ApJ, 259, 449

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

Reed D., Gardner J., Quinn T., Stadel J., Fardal M., Lake G., Governato F., 2003, MNRAS, 346, 565

Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2007, MNRAS, 374, 2

Riess A. G. et al., 1998, AJ, 116, 1009

Rubin V. C., Ford W. K. J., . Thonnard N., 1980, ApJ, 238, 471

Sánchez A. G., Baugh C. M., Percival W. J., Peacock J. A., Padilla N. D., cole S., Frenk C. S., Norberg P., 2006, MNRAS, 366, 189

Sawala T. et al., 2014, ArXiv e-prints

Schaye J. et al., 2014, ArXiv e-prints

Schmidt B. P. et al., 1998, ApJ, 507, 46

Seljak U., 2000, MNRAS, 318, 203

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Sheth R. K., Tormen G., 2002, MNRAS, 329, 61

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481

Somerville R. S., Kolatt T. S., 1999, MNRAS, 305, 1

Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087

Spergel D. N. et al., 2007, ApJS, 170, 377

Spergel D. N. et al., 2003, ApJS, 148, 175

Springel V., 2005a, MNRAS, 364, 1105

Springel V., 2005b, MNRAS, 364, 1105

Springel V. et al., 2008, MNRAS, 391, 1685

Springel V. et al., 2005, Nature, 435, 629

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Srisawat C. et al., 2013, MNRAS, 436, 150

Strauss M. A. et al., 1999, in Bicay M. D., Cutri R. M., Madore B. F., eds, Astronomical Society of the Pacific Conference Series Vol. 177, Astrophysics with Infrared Surveys: A Prelude to SIRTF. p. 390

Summers F. J., Davis M., Evrard A. E., 1995, ApJ, 454, 1

Thomas P. A., Muanwong O., Pearce F. R., Couchman H. M. P., Edge A. C., Jenkins A., Onuora L., 2001, MNRAS, 324, 450

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Tormen G., 1997, MNRAS, 290, 411

van den Aarssen L. G., Bringmann T., Pfrommer C., 2012, Physical Review Letters, 109, 231301

van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, MNRAS, 430, 725

van den Bosch F. C., Tormen G., Giocoli C., 2005, MNRAS, 359, 1029

Velander M. et al., 2014a, MNRAS, 437, 2111

Velander M. et al., 2014b, MNRAS, 437, 2111

Vitvitska M., Klypin A. A., Kravtsov A. V., Wechsler R. H., Primack J. R., Bullock J. S., 2002, ApJ, 581, 799

von der Linden A. et al., 2014, MNRAS, 443, 1973

Wang H. Y., Jing Y. P., Mao S., Kang X., 2005, MNRAS, 364, 424

Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, ApJ, 646, 881

Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230

Weinberg D. H., Bullock J. S., Governato F., Kuzio de Naray R., Peter A. H. G., 2013, ArXiv e-prints

Wetzel A. R., 2011, MNRAS, 412, 49

White M., 2002, ApJS, 143, 241

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Yang X., Mo H. J., van den Bosch F. C., 2003, MNRAS, 339, 1057

Zehavi I. et al., 2011, ApJ, 736, 59

Zentner A. R., Berlind A. A., Bullock J. S., Kravtsov A. V., Wechsler R. H., 2005, ApJ, 624, 505

Zheng Z. et al., 2005, ApJ, 633, 791

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

A

APPENDIX

A.1 Circularity in the Keplerian approximation

To compare the circularity, $J/J_{\rm circ}(E)$, inferred under the assumption that the infalling satellite and host halo are treated as two point masses on a Keplerian orbit with the singular isothermal sphere (SIS) model we need to compare the corresponding expressions for the angular momenta of circular orbits, $J_{\rm circ}(E)$. For the Keplerian case this is easily derived from the angular momentum of a circular orbit of radius r, $J_{\rm circ} = \mu V_{\rm circ} r$, where the circular velocity at separation r is given by $\mu V_{\rm circ}^2 = GM_{\rm h}M_{\rm s}/r$ and the corresponding orbital energy $E = \mu V_{\rm circ}^2/2 - GM_{\rm h}M_{\rm s}/r$. Here μ is the reduced mass, which can be expressed in terms of the satellite and host masses as $\mu = M_{\rm s}M_{\rm h}/\left(M_{\rm s}+M_{\rm h}\right)$. Eliminating both $V_{\rm circ}$ and r from these three equations yields

$$J_{\rm circ}^{\rm Kep}(E) = \sqrt{\frac{(GM_{\rm h}M_{\rm s})^2\mu}{-2E}}. \tag{1.1.1}$$

If V is the velocity difference between the satellite and host when the satellite crosses the virial radius, r_{200}, then

$$E = \frac{1}{2}\mu V^2 - \frac{GM_{\rm s}M_{\rm h}}{r_{200}} = \frac{1}{2}\mu V^2 - M_{\rm s}V_{200}^2, \tag{1.1.2}$$

where the circular velocity, V_{200}, is given by $V_{200} = \sqrt{GM_{\rm h}/r_{200}}$. Using Eqn. 1.1.2 to substitute for E in Eqn. 1.1.1 yields

$$\frac{J_{\rm circ}^{\rm Kep}(E)}{M_{\rm s}V_{200}\,r_{200}} = \frac{1}{\sqrt{2M_{\rm s}/\mu - V^2/V_{200}^2}}. \tag{1.1.3}$$

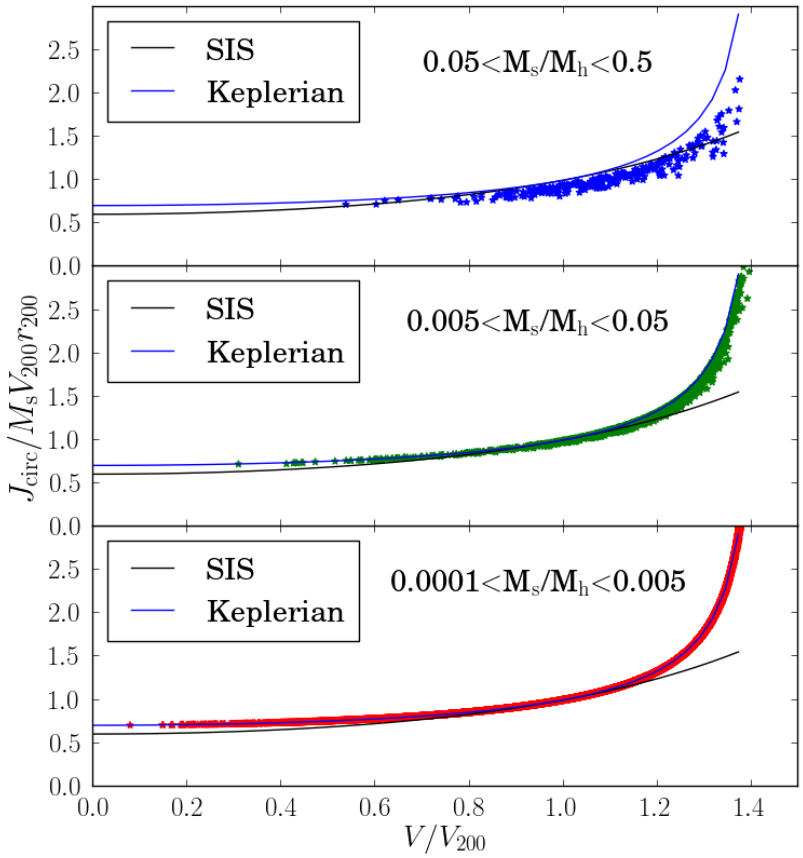

Figure A.1: A comparison of the Keplerian and singular isothermal sphere (SIS) models of $J_{\rm circ}$ in units of $M_{\rm s} V_{200}\, r_{200}$ for satellites with infall velocity, V, at the virial radius r_{200}. In each panel, the black solid line is the SIS expression and the blue solid line is for the Keplerian case in the limit $M_{\rm s}/M_{\rm h} \ll 1$. The stars show the result of the full Keplerian expression including the dependence on the reduced mass, μ, for samples of satellites in different bins of $M_{\rm s}/M_{\rm h}$.

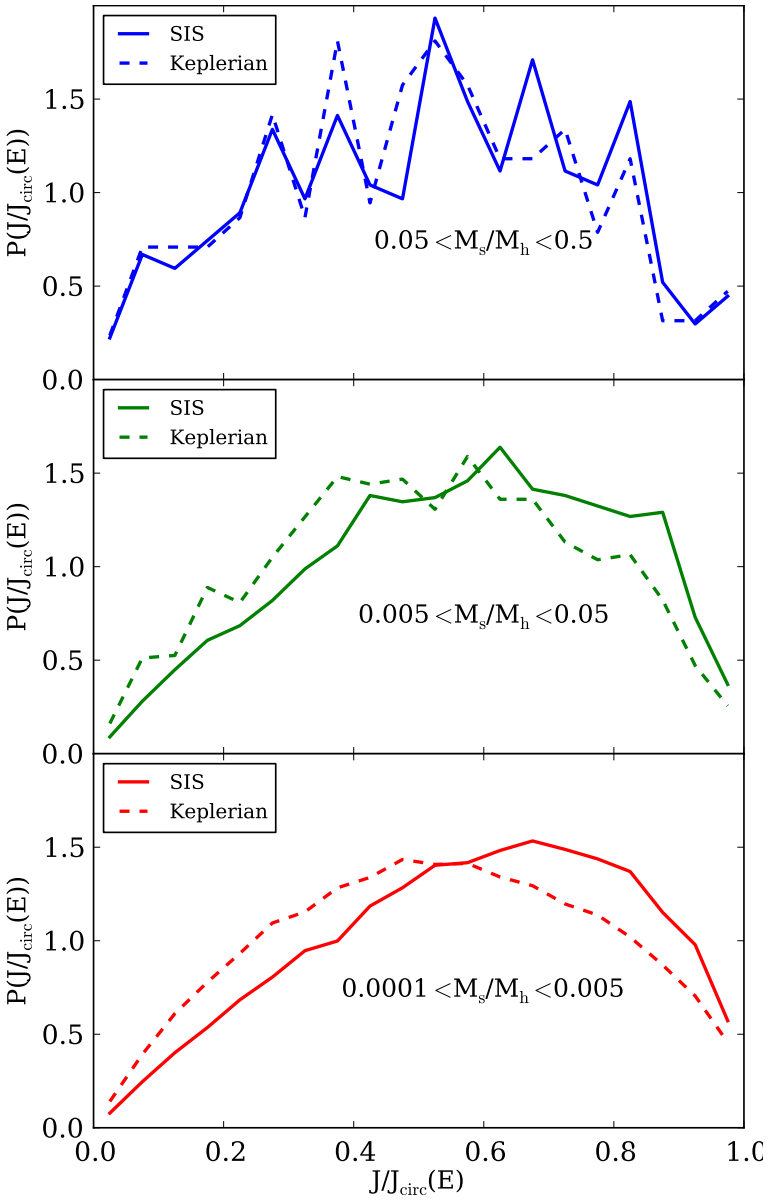

Figure A.2: Distributions of circularity, $J/J_{\rm circ}(E)$, for infalling satellite haloes for host haloes in a mass bin centred on 10^{13} $M_\odot$. Solid curves show the distribution derived assuming a singular isothermal sphere and dashed curves show the distribution derived using the Keplerian model. The three panels are for the same three bins of $M_{\rm s}/M_{\rm h}$ as in Fig. A.1.

This compares with the singular isothermal sphere expression for $J_{\rm circ}$ derived in Section 5.2.2,

$$\frac{J_{\rm circ}^{\rm SIS}(E)}{M_{\rm s} V_{200}\, r_{200}} = \exp\left(\frac{1}{2}\left(\frac{V^2}{V_{200}^2} - 1\right)\right). \tag{1.1.4}$$

In the SIS case we could also take account of the reduced mass by treating the SIS as a rigid potential profile whose centre of mass orbits around the combined centre of mass of the satellite plus halo. In this case the expression for $J_{\rm circ}$ becomes

$$\frac{J_{\rm circ}^{\rm SIS}(E)}{M_{\rm s} V_{200}\, r_{200}} = \exp\left(\frac{1}{2}\frac{1}{M_{\rm s}/M_{\rm h}+1}\left(\frac{V^2}{V_{200}^2} - 1\right)\right). \tag{1.1.5}$$

The solid curves in Fig. A.1 compare, as a function of satellite infall velocity, V, the SIS expression with the Keplerian expression evaluated in the limit $M_{\rm s}/M_{\rm h} \ll 1$, such that $\mu \to M_{\rm s}$. The individual points on the different panels show the results of the full Keplerian expression with its dependence on $M_{\rm s}/\mu$ applied to our satellite sample in different bins of $M_{\rm s}/M_{\rm h}$. The model curves necessarily agree at $V = V_{200}$ because the mass enclosed in a circular orbit at r_{200}, where the circular velocity is V_{200}, is the same by construction. For $M_{\rm s}/M_{\rm h} \ll 1$, the difference between the two models is largest at large V/V_{200} where the orbits extend far beyond r_{200} and hence the mass enclosed in the SIS greatly exceeds the mass assumed in the point mass approximation. The effect of the reduced mass, μ, is small for $M_{\rm s}/M_{\rm h} < 0.05$, but for $0.05 < M_{\rm s}/M_{\rm h} < 0.5$ it has the effect of reducing $J_{\rm circ}^{\rm Kep}(E)$ and produces values closer to the SIS case. This is demonstrated in Fig. A.2 which compares the distribution of circularities, $J/J_{\rm circ}(E)$, evaluated using the two different expressions for three ranges in satellite-to-host mass ratio. Overall, the two models agree well with each other for higher values of $M_{\rm s}/M_{\rm h}$, but they differ for the lowest mass ratio bins.

Fig. A.3 shows how the $J_{\rm circ}(E)$ values are changed in the SIS case if we take account of the reduced mass. We find the difference is very small for all the mass ratio bins and is only at all appreciable for the bin $0.05 < M_{\rm s}/M_{\rm h} < 0.5$. We do not adopt this as the default because the effect is small, we want to be compatible with previous analyses and because treating the SIS as responding rigidly will not be a good approximation.

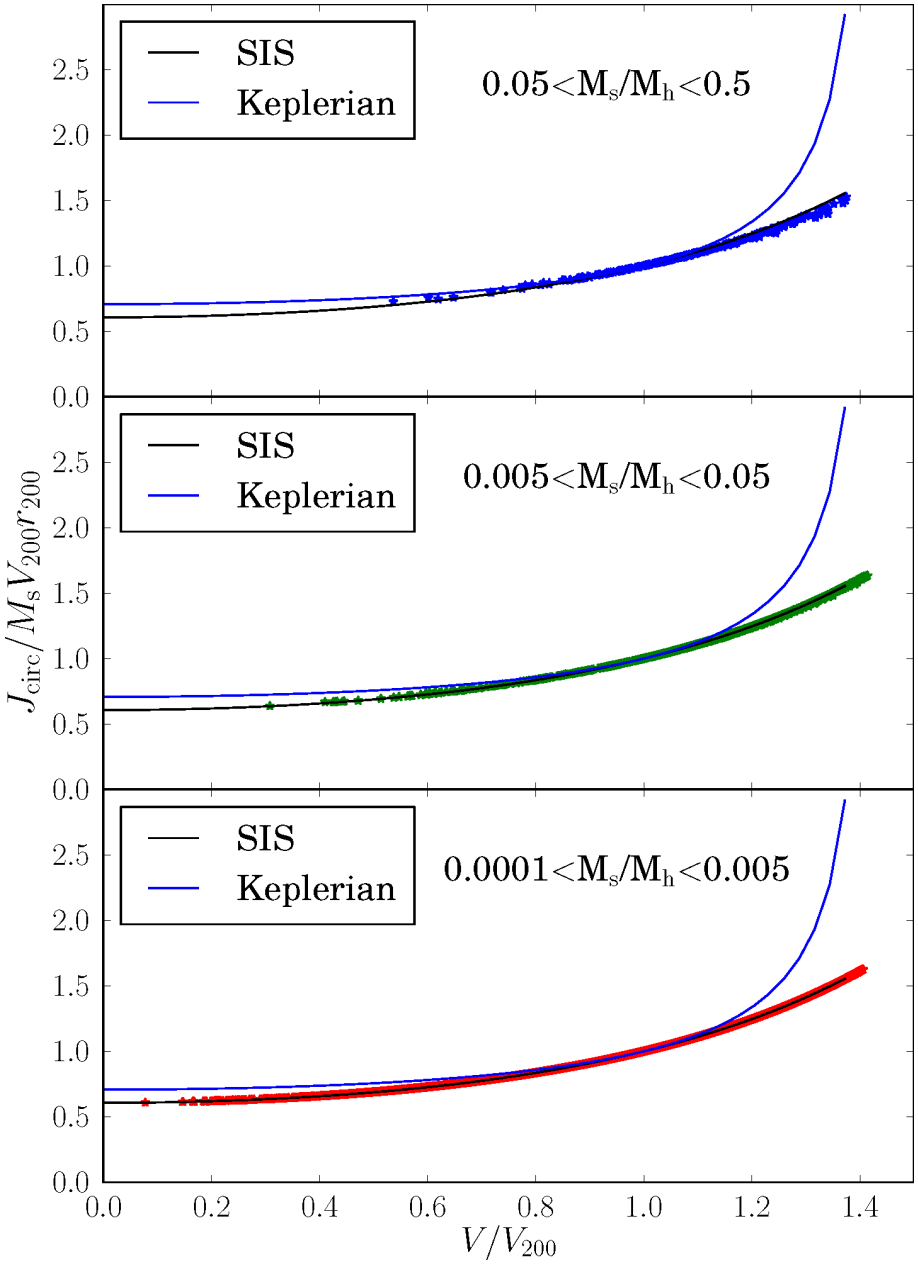

Figure A.3: Like Fig. A.1, but now the stars show the result of using the SIS expression including the dependence on the reduced mass, μ.

A.2 The distribution of tangential velocities

The bivariate distribution in Fig. 5.5 indicates that V_θ/V and total velocity V/V_{200} have no correlation. Consequently we could have equally well fitted V_θ/V rather than V_r/V. Fig. A.4 shows the distributions of V_θ/V and exponential fits of the form

$$P(V_\theta/V) = C\left(\exp\left(\frac{DV_\theta}{V}\right) - 1\right). \tag{1.2.6}$$

The individual fits are compared in Fig. A.5, where we again see the tendency for more massive satellites to be on more radial orbits. The black solid lines in Fig. A.5 show the alternative fits which result from the fits to V_r/V and V/V_{200} presented in Section 5.3.4. Unlike the monotonic exponential fits these curves turn over at large V_θ/V. However within the noise they are equally good if not better fits to the data and this supports our choice that V_r/V and total velocity V/V_{200} are the best orbital parameters to adopt in order to characterise the full bivariate parameter distribution.

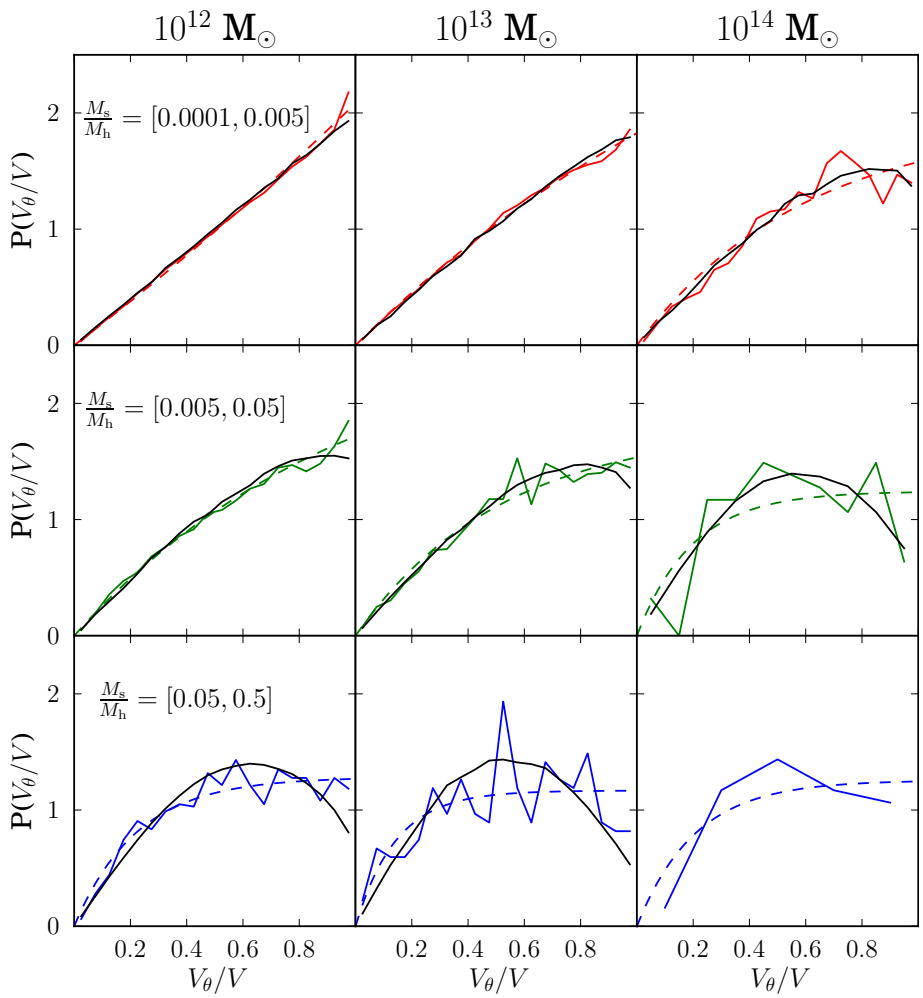

Figure A.4: Probability distribution of the orbital parameter V_θ/V, as functions of both the satellite-to-host mass ratio and the host halo mass. Each column is for a fixed final halo mass as labelled at the top of the column. Each row is for a different bin in satellite-to-host mass ratio: top (red lines) $0.0001 < M_s/M_h < 0.005$, middle (green lines) $0.005 < M_s/M_h < 0.05$ and bottom (blue lines) $M_s/M_h > 0.05$. The dashed curves are the best fitting exponential distributions. The black solid lines show the distributions derived from the fits presented in Section 5.3.4.

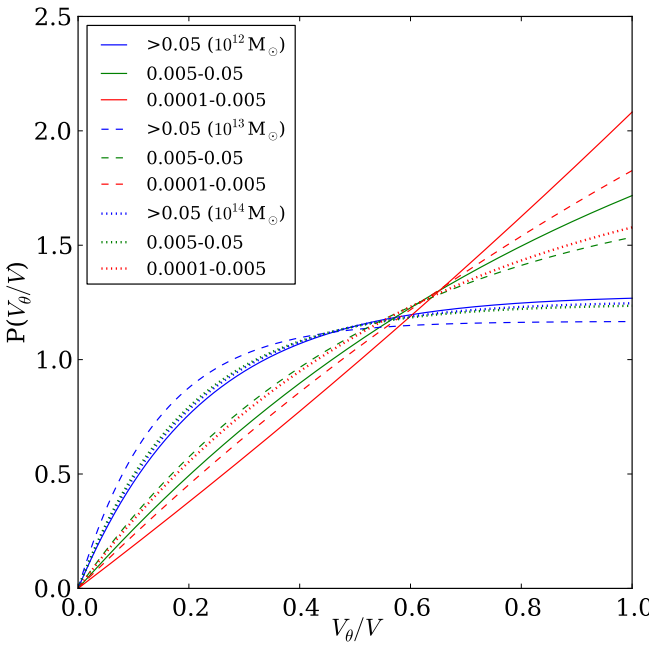

Figure A.5: The fitted distributions of the orbital parameter V_θ/V for different values of both the satellite-to-host mass ratio at infall and the final host halo mass. The line colour denotes satellite-to-host mass ratio: red $0.0001 < M_s/M_h < 0.005$, green $0.005 < M_s/M_h < 0.05$ and blue $M_s/M_h > 0.05$. The line style indicates the host halo mass: solid 10^{12} M$_\odot$, dashed 10^{13} M$_\odot$ and dotted 10^{14} M$_\odot$