Cosmology with Cluster Structural Properties

by

Yuba Amoura

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	August Evrard Professor, Dept. of Physics and Astronomy University of Michigan
Supervisor(s):	James Taylor Associate Professor, Dept. of Physics and Astronomy University of Waterloo
Internal Member:	Will Percival Professor, Dept. of Physics and Astronomy University of Waterloo
Internal-External Member:	Francis Poulin Professor, Dept. of Applied Mathematics University of Waterloo
Other Member(s):	Brian McNamara Professor, Dept. of Physics and Astronomy University of Waterloo

Author's Declaration

This thesis consists of material I authored or co-authored: see the Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapter 2 of this thesis is based on the following article

Amoura, Y., Drakos, N. E., Berrouet, A., & Taylor, J. E. (2021). Cluster assembly times as a cosmological test. Monthly Notices of the Royal Astronomical Society, 508(1), 100-117.

The article is published in the peer-reviewed journal *Monthly Notices of the Royal Astronomical Society* (MNRAS). I am the first author of the article, and performed all the analysis using simulations run by Anael Berrouet. My supervisor James Taylor is a co-author of the paper and wrote most of the Introduction, Conclusion and Section 5.

Chapter 3 of this thesis is based on the following article

Amoura, Y., Drakos, N. E., Berrouet, A., & Taylor, J. E. (2023). Halo Growth and Merger Rates as a Cosmological Test. Submitted to MNRAS.

The article has been submitted for publication in MNRAS and is currently under review. I am the first author of the article, and performed most of the analysis using simulations run by Anael Berrouet. My supervisor James Taylor did the analysis in Section 5 and wrote most of the section.

Chapter 4 is original work where I set up, ran and analysed all the simulations and did all the writing and analysis.

Abstract

Galaxy clusters are massive objects composed of hundreds or thousands of galaxies, hot gas and an extended dark matter (DM) halo. They are the largest gravitationally bound structures in the Universe. As such, they result from the growth and collapse of the highest peaks in the density perturbation field. Because fundamental properties of the Universe determine the statistics and evolution of those peaks, clusters can be used for cosmology. The most common method to do that is to use their abundance, which requires counting them in bins of mass and redshift (distance). The main challenge in using this method is accurately estimating their mass through observational proxies. This method has the advantage of requiring only the mass and redshift of each cluster, but neglects other cluster properties that could contain cosmological information. Since clusters are complex multi-component objects that can be detected in optical, X-ray, and radio wavelengths, a wealth of information is discarded and compressed into a single mass estimate. Moreover, the abundance of clusters is influenced in the same way by increased power in the initial spectrum of density fluctuations, or by increased growth at later times. This induces a degeneracy in estimates of the amount of matter in the Universe Ω_m—governing the growth—and how "clumpy" the matter distribution is, σ_8—governing the height of initial peaks.

In this thesis, I explore another way to exploit clusters to gain cosmological insight. Inspired by original work from Richstone et al. (1992), Evrard et al. (1993) and Mohr et al. (1995), I study how clusters' structural properties and dynamical state can provide useful constraints on Ω_m and σ_8.

In the first part of this thesis, I study how cluster formation time varies as a function of Ω_m and σ_8 through analytical models and DM-only simulations, and find that cluster age constraints are almost orthogonal to typical abundance and structure-growth ones. Moreover, the formation redshift varies by a factor of 2 from one end to the other of the relevant part of the Ω_m and σ_8 plane. Using cluster concentration as a tool to measure age, we would need about 10,000 clusters to get measurements of σ_8 up to 1% accuracy.

I continue in the second part to study the cosmological dependence of cluster and dark matter halo dynamical states but focus on the instantaneous equivalent of age, specifically, the halo growth and merger rates. Here again, I use analytical models and DM-only simulations and find that growth rates' dependence on Ω_m and σ_8 is also orthogonal to cluster abundance constraints. However, the sensitivity is weaker in this case, ranging between 60% to 90%. Using cluster and galaxy growth and merger rates becomes challenging when considering the current scatter in observational measurements.

Finally, the last part of this thesis is to link the dynamical state of haloes to their structural properties in detail. I use a set of high-resolution simulations that I have run for different cosmologies to study what information is contained in the accretion history of haloes, and show that the mass accretion history (MAH) of haloes is mainly a one-dimensional quantity. Moreover, a halo's age is well-described by its state during the middle of its growth, such as the time it reaches 50% of its final mass z_{50}.

The halo concentration and the offset between the centre-of-mass and peak of the density contours are particularly good age predictors when used individually. If we use them together and/or with another structural property, such as shape, they become even better by reducing scatter in individual relations and segregating very young and very old halo populations.

In conclusion, in this thesis, I show how we can use cluster structural properties to constrain Ω_m and σ_8. First, I show how cluster age and growth rate vary with cosmology, and then how age correlates with cluster structural properties.

Acknowledgements

First and foremost, I would like to thank my supervisor James Taylor for his guidance, patience and insight throughout my Ph.D., as well as for his valuable advice on writing, presentation and organisation. I would also like to thank the members of my advisory committee, Brian McNamara and Will Percival, for their helpful comments and suggestions, as well as other members of the Waterloo Centre of Astrophysics for the many scientific discussions and their camaraderie and support during good and bad times. A special thanks to Liza and Roan for their help and friendship and to Prathamesh for being my pandemic work partner and friend.

This work heavily relied on the resources provided by the Digital Research Alliance of Canada (ex-Compute Canada) for running and analysing N-body simulations.

Finally, I would like to thank my parents, who supported me during every step of my journey and made countless sacrifices to allow me to pursue the path that I have chosen.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Standard Cosmological Model

In 1922, Alexander Friedmann, a Russian physicist, published for the first time a solution to Einstein's equations of General Relativity that describes an expanding Universe (Friedmann, 1922). Five years later, independently from Friedmann, a Belgian cosmologist and priest named George Lemaître published similar solutions proposing a model of an expanding Universe; in the same article, he uses the redshifts of galaxies (then known as "Nebulae") as measured by Vesto Slipher (Slipher, 1915) to link their distance to their receding velocities for the first time, proving observationally the validity of his model (Lemaître, 1927). This leads Lemaître to propose a "hypothesis of the primaeval atom" where the Universe started from a tiny particle exploding and expanding. This theory would be ridiculed by the British cosmologist Fred Hoyle during a radio show labelling it "The Big Bang"

A century later, Friedmann's solutions and Lemaître's model are still the basis of the current standard model of cosmology, but we have come a long way in learning more about its composition and the processes by which the observed matter and structures formed. The current general consensus describes the Universe starting from a period of rapid expansion, *Inflation*, necessary to explain the apparent Euclidean geometry of the Universe and its isotropy on very large scales. This is followed by a period of slowing expansion until the current epoch, where the expansion accelerates again. The rate of this expansion is typically given by the Hubble parameter $H(t)$, which describes the rate of relative change in the Universe's "size" Its current value $H_0 \equiv H(t_0)$ is one of the cosmological parameters of the standard cosmological model.

Because the Universe is in expansion, light from distant objects is systematically shifted to longer wavelengths due to the Doppler effect. This "redshift" is denoted z

$$1 + z \equiv \frac{\lambda_0}{\lambda_e} \ , \tag{1.1}$$

and is straightforward to measure with spectroscopic or even photometric data. It is typically used as a proxy for the distances of astronomical objects and the age of the Universe when those objects were observed.

The Universe has different components with very different physical properties. Non-relativistic (Cold) matter, which is composed of particles from the Standard Model of particle physics, and is referred to as "Baryonic" Matter. This accounts for all the ordinary matter in stars, planets gas etc. It has an energy density $\epsilon_b(t)$. The second "Dark" component is not part of the particle physics Standard Model and manifests its existence through gravitational effects. The energy density of the total (baryonic + dark) matter is denoted $\epsilon_m(t)$.

The second main component of the Universe is relativistic matter, referred to as "radiation", composed mainly of photons (light) and neutrinos; it has energy density $\epsilon_r(t)$. Finally, most of the Universe's energy density today is in the form of a negative pressure fluid with the equation of state $p = -\rho$, accelerating the Universe's expansion, called "dark energy", often assumed to be a cosmological constant Λ with energy density $\epsilon_\Lambda(t)$. The total energy density if the Universe were to have a purely Euclidean geometry is the critical density $\epsilon_c(t)$. It is convenient to express the densities of each component "i" relative to the critical density $\Omega_i(t) \equiv \frac{\epsilon_i(t)}{\epsilon_c(t)}$. These constitute another set of parameters of the standard cosmological model.

Although the standard cosmological model is built on the key principle of homogeneity and isotropy, this can obviously not be completely true; otherwise, the wide variety of observed structures could not have been formed. Initial inhomogeneities in the density field must have existed at the end of the inflation era. The spatial distribution of the overdensities and underdensities $\delta(\boldsymbol{x})$ is assumed to follow an isotropic and homogeneous Gaussian random field. Hence, according to the Wiener-Khintchine theorem (Wiener, 1930; Khintchine, 1934), it is completely determined by a unique function, the power spectrum $P(k)$. In the simplest case, the initial post-inflation power spectrum is a power law with index n_s: $P(k) \propto k^{n_s}$. A common way to express the amplitude of the power spectrum is through the r.m.s. of the overdensity field today at a specific scale of 8 Mpc/h: σ_8.

My work takes place within the general framework described above, called the ΛCDM model (Λ Cold Dark Matter). It will specifically focus on the structures of the Universe,

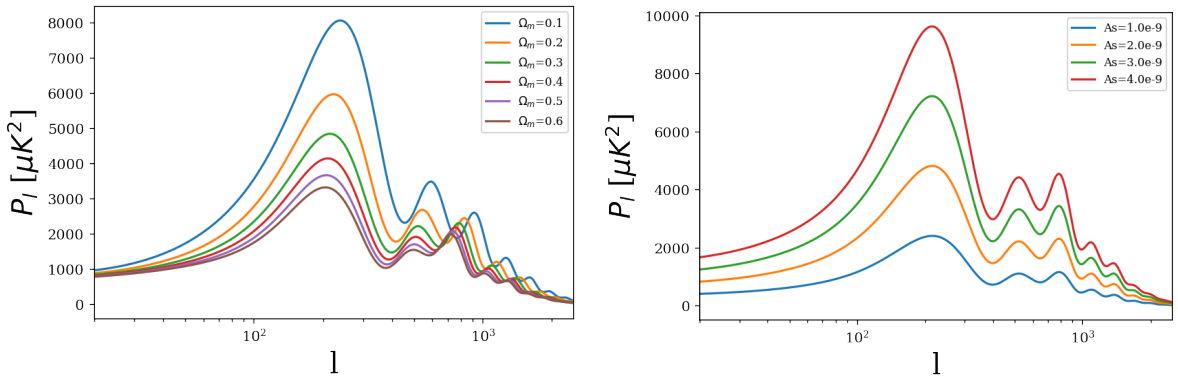

Figure 1.1: CMB power spectrum generated using CAMB (Lewis et al., 2000). The left panel illustrates how Ω_m influences the peaks, while the right panel shows the change in the normalisation A_s, which relates to σ_8.

their growth and the cosmological parameters that influence them the most, namely the matter density Ω_m and the amplitude of density fluctuations σ_8. But first, I summarize briefly the formalism of the structure growth within the standard cosmological model.

1.1.1 The Growth of Structures

Gravitationally bound structures such as galaxies and galaxy clusters have formed from the gravitational collapse of dense regions. However, the Universe started from a highly homogeneous state with small perturbations; these proceeded to grow due to the effect of gravity. This means that the formation of structures is determined by two different factors, the initial perturbation field and its subsequent growth.

The perturbation field is described by $\delta(\boldsymbol{x}) \equiv \frac{\rho(\boldsymbol{x})-\bar{\rho}_m}{\bar{\rho}_m}$ where ρ is the (cold) matter density field and $\bar{\rho}_m$ is the average matter density in the Universe. In Fourier space, the variance of the Fourier transform of the perturbation field is the **power spectrum**

$$P(k) \equiv \left\langle |\delta(\boldsymbol{k})|^2 \right\rangle \ . \tag{1.2}$$

This is enough to describe the properties of an isotropic Gaussian random field fully (Bardeen et al., 1986).

Quantum fluctuations result in an initial power spectrum, right after the inflationary period, that is assumed to be a power law with index n_s

$$P_i(k) \propto k^{n_s} \ . \tag{1.3}$$

At that epoch, the Universe was hot enough for baryonic matter to couple with radiation, inducing radiative pressure pushing dense regions out. The interplay of gravity pulling in and pressure pushing out provoked acoustic oscillations in the density field that can be observed in the temperature distribution of the **Cosmic Microwave Background** (CMB) and the spatial distribution of large-scale structures. The effect of baryons and radiation coupling at different scales on the shape of the matter power spectrum is typically accounted for through the *Transfer function* $T(k)$. This gives a matter power spectrum after decoupling with a form

$$P(k) = P_i(k)T(k)^2 \ . \tag{1.4}$$

The form of the transfer function is determined by the complex physics of baryon-radiation fluid, which depends on the composition of the Universe, notably, the total amount of matter Ω_m (for gravitational effects) and the amount of baryonic matter Ω_b.

If we fix every cosmological parameter necessary to get the transfer function and the initial power spectrum, the overall amplitude of the matter power spectrum is still a free parameter. It is typical in modern cosmology to use the variance of density fluctuations observed today to fix the normalisation of the power spectrum

$$\sigma^2(R) \equiv \left\langle \delta_R^2(\mathbf{x}) \right\rangle = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) \widetilde{W}_R^2(k) dk \ , \tag{1.5}$$

where the perturbation field $\delta_R(\boldsymbol{x})$ is smoothed at a scale R using the filter $\widetilde{W}_R(k)$.

The value of $\sigma_8 \equiv \sigma(R = 8\text{Mpc/h})$ is used to fix the amplitude of the power spectrum.

After decoupling, the Universe enters a matter-dominated epoch, and the perturbations start growing effectively due to gravity. When they are small $|\delta| << 1$, they evolve linearly

$$\delta(\boldsymbol{r}, t) = \delta(\boldsymbol{r}, t_0) D(t)/D(t_0)) \ , \tag{1.6}$$

where $D(t)$ is the *growth factor* and can be estimated by applying Newtonian gravity in an expanding Universe. In a universe dominated by matter and dark energy (Λ), the growth factor is given by (Heath, 1977; Peebles, 1980; Hamilton, 2001)

$$D(a) = \frac{5\Omega_m \tilde{H}(a)}{2} \int_0^a \frac{da'}{a'^3 \tilde{H}(a')^3} \ , \tag{1.7}$$

where a is the scale factor, parametrizing the relative expansion of the Universe and varies between 0 at the beginning of the Universe and 1 today, and $\tilde{H}(a)$ is the Hubble parameter

normalized to today's value. In a Universe with cold dark matter (CDM) and dark energy (Λ)-the ΛCDM model- the Hubble parameter evolves following:

$$\tilde{H}(a) = \frac{H(a)}{H_0} = \sqrt{\Omega_m a^{-3} + \Omega_\Lambda + (1 - \Omega_m - \Omega_\Lambda)a^{-2}} \ . \tag{1.8}$$

Note that for a "flat" ΛCDM, we have $\Omega_m + \Omega_\Lambda = 1$, which means that the growth factor is completely determined by Ω_m alone. This is an important property; the growth of linear perturbations in such a Universe is governed solely by the matter density Ω_m.

When perturbations become large enough that $\delta \sim 1$, they enter the non-linear regime of gravitational instability and eventual collapse and virialize to form bound structures. With some assumptions, this process can be analysed analytically using the **spherical collapse model** (Gunn & Gott, 1972).

The model considers spherical perturbations of amplitude δ with radius R and density $\rho = \bar{\rho}(1+\delta)$. This perturbation is taken to be only dark matter (only subject to gravity) and composed of many individual mass shells evolving in an Enstein-de-Sitter (EdS) universe ($\Omega = \Omega_m = 1$). The mass enclosed by any mass shell will be

$$\begin{aligned} M(< r) &= \frac{4}{3}\pi r^3(t)\rho(t) \\ &= \frac{4}{3}\pi r^3(t)\bar{\rho}[1 + \delta(t)] \ , \end{aligned} \tag{1.9}$$

and will be subject to equations of motion. In Newtonian gravity it is simply

$$\frac{d^2 r}{dt^2} = -\frac{GM}{r^2} \ . \tag{1.10}$$

Integrating the equation and assuming negative total energy (collapse), we find parametric solutions of the form

$$\begin{aligned} r &= A(1 - \cos\theta) \\ t &= B(\theta - \sin\theta) \ . \end{aligned} \tag{1.11}$$

In this scenario, a shell expands from $r = 0$ at $t = 0$, reaches a maximum $r_{\max} = r(\theta = \pi)$ at $t_{\max}$, then collapses to $r = 0$ at $t_{\rm col} = t(\theta = 2\pi)$. The maximum radius and corresponding time are often referred to as the **turn-around** radius and time.

Now that we have the equations of motion, we can easily calculate the evolution of the overdensity $\delta(t)$

$$1 + \delta = \frac{\rho}{\bar{\rho}} = \frac{9(\theta - \sin\theta)^2}{2(1 - \cos\theta)^3} \ ,$$

which means that at turn-around we have $1 + \delta(\theta = \pi) = \frac{9\pi^2}{16} \simeq 5.55$.

The overdensity never fully collapses; otherwise, it would result in a black hole. In reality, it reaches a virial equilibrium at $r_{\rm vir}$. Using the virial theorem, one can easily find that $r_{\rm vir} = r_{\rm ta}/2$, which means that the density at virial radius is 8 times the density at the turn-around radius. In addition to that, it takes the overdensity $t_{\rm col} = t_{\rm ta}/2$ to virialise, time during which the average density of the Universe has dropped with a rate of $\bar{\rho} \propto a(t)^{-3}$. In Einstein-de-Sitter Universe (EdS) we have $a \propto t^{2/3}$ giving that the average density drops by

$$\bar{\rho}(t_{\rm col}) = \bar{\rho}(t_{\rm ta}) \left(\frac{t_{\rm col}}{t_{\rm ta}}\right)^{-2} = \bar{\rho}(t_{\rm ta})/4 \ . \tag{1.12}$$

The combination of the fact that the density at the virial radius is eight times larger and the average density of the Universe is four times lower by the time of virialisation results in the average virial overdensity

$$1 + \Delta_{\rm vir} \simeq 32 \times 5.55 \simeq 178 \ . \tag{1.13}$$

This calculation can also be done for non-EdS cosmologies (Bryan & Norman, 1998), and the results are typically used as a definition of the overdensity at the virial radius in simulations and analytical models.

Since we assumed an EdS Universe, we also know that linear overdensity evolves as $\delta \propto a \propto t^{2/3}$. This allows us to find the value of the overdensity at $t_{\rm col}$ if the perturbation continues to be linear and we find $\delta_c \simeq 1.686$. This is the value of the linear overdensity that would have reached when the overdensity collapses.

These two values, δ_c and $\Delta_{\rm vir}$ will be important throughout this work because δ_c provides a condition on the **linear** field for a structure to collapse and $\Delta_{\rm vir}$ gives the average density, and consequently also the size and mass of the newly-formed structure.

In summary, the formation of structures is influenced by the state of perturbations at recombination and their subsequent growth, during the matter and then dark energy-dominated eras. The matter power spectrum and growth factor contain most of the information necessary to predict the statistical properties of the observed distribution of matter in the Universe. Consequently, studying those properties allows one to estimate the power spectrum and growth factor, constraining the cosmological parameters to which they are sensitive. In this specific case, Ω_m influences the power spectrum's growth, form, and amplitude and σ_8 its amplitude. These two parameters are typically referred to as the *growth of structure* parameters, and the heart of this thesis is constraining their value.

1.1.2 Measuring Ω_m and σ_8 : current methods

Most current methods that are used to constrain the values of Ω_m and σ_8 are ways to probe the power spectrum. A variety of observable quantities can be used to measure the power spectrum using different objects subject to very different types of physics at different epochs of the Universe's history. Most of these observables are sensitive to a combination of Ω_m and σ_8 and are not able to disentangle their respective effects. The most common degeneracy follows contours of constant $S_8 \equiv \sigma_8\sqrt{\Omega_m/0.3}$ where S_8 is often called the *Structure Growth Parameter*.

CMB Power Spectrum

The Cosmic Microwave Background (CMB) is the first light in the Universe after it became transparent. The temperature anisotropies observed in the CMB are the result of the density fluctuations when the CMB light is emitted; consequently, the CMB power spectrum is determined by the pre-recombination physics. The physical matter density $\rho_m \propto H^2\Omega_m$ influences the amplitude and location of the peaks as seen in the left panel of Fig. 1.1, while its amplitude A_s is trivially related to σ_8.

Cluster Counts

Galaxy clusters are the largest bound objects in the Universe. They are at the heart of this thesis, and I will talk in detail about them in Sec. 1.4. For now, we will focus on how we can use their abundance to constrain Ω_m and σ_8 (Allen et al., 2011).

For small enough redshift ranges, over which the cluster mass function is assumed to not vary, the number of clusters between masses $M_i - M_{i+1}$ and redshifts $z_j - z_{j+1}$ can be written as

$$N(M_i, z_j) = \frac{\Delta\Omega}{4\pi}\int_{z_j}^{z_{j+1}} dz \frac{dV}{dz}\int_{M_i}^{M_{i+1}} \frac{dn}{dM}\, dM \quad , \tag{1.14}$$

Where $\Delta\Omega$ is the total solid angle that the survey spans, dn/dM is the cluster mass function, which can be found either through analytical models (Press & Schechter, 1974, e.g.), N-body simulations (Tinker et al., 2008, e.g.) or both. The cluster mass function depends on both the power spectrum and the growth factor, hence, on both Ω_m and σ_8, roughly following contours of constant S_8.

The main challenge in these methods is getting the cluster mass, often obtained through observable proxies such as the cluster richness $N(M)$ or luminosity $L(M)$, which need to be calibrated.

Weak Gravitational Lensing

When light passes through a gravitational potential, it follows the local geodesics, which changes its direction (Dyson et al., 1920). The source of the potential effectively acts as a lens. Consequently, by looking at statistics of lensed light, we probe the underlying gravitational potential (Walsh et al., 1979). One way to find lensed light is to look at the statistics of galaxy shapes; this is called the *cosmic shear*. Looking at correlations between galaxy shapes in bins of redshift gives a direct estimation of the underlying density correlation function, and hence the matter power spectrum.

Instead of correlating galaxy shapes, one could also look at the correlation between galaxy shapes and galaxy positions using a method called galaxy-galaxy lensing (e.g. Brainerd et al., 1996). The idea is similar to cosmic shear since, in effect, we are looking at overdense regions which are more likely to contain galaxies and to bend the light coming from background galaxies. This method necessitates linking galaxy number densities to matter density and introduces an additional bias parameter, because galaxies are more likely to populate the highest-density regions.

Additionally, one can look at galaxy clustering, that is, correlations in galaxy positions, as a probe of the underlying matter density. Typical weak lensing surveys combine all three methods to obtain better constraints and break the degeneracy between Ω_m and σ_8; this combination is referred to as a 3x2pt correlation function (Abbott et al., 2022, e.g.).

Peculiar Velocities

Real Space

The gravitational effect of the matter density field on galaxies can be seen through their "Peculiar Velocity", which is the velocity relative to the expansion of the Universe called the "Hubble flow" The statistics of the galaxies' peculiar velocity relate directly to the power spectrum and growth factor through (Peebles, 1980; Gorski, 1988)

$$\langle v^2 \rangle = \frac{H^2 f^2 a^2}{2\pi^2} \int P(k) dk \ , \tag{1.15}$$

where $f \equiv \frac{d \ln D}{d \ln a}$ is the dimensionless growth factor, H the Hubble parameter and a the scale factor. Since the power spectrum is directly proportional to σ_8^2 by design, this method can only constrain a combination of the parameters $f\sigma_8$. Additionally, accurate distance measurement remains challenging in astronomy beyond the local Universe.

Redshift-Space Distortion (RSD)

The RSD method uses the systematic change in galaxies' redshifts due to their peculiar velocities to probe the underlying matter density field. As for the peculiar velocities method, the RSD can only constrain a combination of $f\sigma_8$.

Results and the S_8 Tension

Recent results have highlighted a discrepancy between the value of S_8 inferred from the CMB power spectrum and through probes of the low-redshift matter density field. Specifically, the latest results from the *Planck* collaboration Planck Collaboration et al. (2020) provide a value of $S_8 = 0.832 \pm 0.013$, while the most accurate weak lensing results from the Kilo Degree Survey (KiDS-1000) (Asgari et al., 2021) find a value of $S_8 = 0.766^{0.020}_{-0.014}$, that is a three standard deviation difference away from the Planck result. Moreover, other recent weak lensing results from the three-year data collection from Dark Energy Survey (DES Y3) also find a significantly lower value of S_8. Results from cluster counts, peculiar velocities and redshift space distortions also tend to favour a lower value of S_8. Abdalla et al. (2022) summarize all recent results in Fig. 1.2 which clearly shows a systematic preference of late-Universe studies to lower value of S_8 compared to early-Universe ones.

1.2 Dark Matter Haloes

When dark matter overdensities grow and then collapse, following the process described in section 1.1.1, they form bound structures in virial equilibrium called haloes. These structures represent the gravitational backbone of large structures such as galaxies and clusters. In a "Cold" dark matter (CDM) universe, they develop and grow by accretion and mergers through a bottom-up process, giving rise to a hierarchical structure made of haloes that host clusters and large galaxies, subhaloes that host satellite galaxies and sometimes even subhaloes within subhaloes (Springel et al., 2008).

Density Profile

Navarro et al. (1996) initially proposed a universal, spherically symmetric, two-parameter profile to fit the density of CDM haloes found in their simulations. The two parameters are an overall normalisation ρ_s and a characteristic radius r_s or r_{-2} called the scale radius which represents the radius at logarithmic slope -2. The profile is universal; that is, it has the same shape independently of halo mass, initial power spectrum or cosmology; it

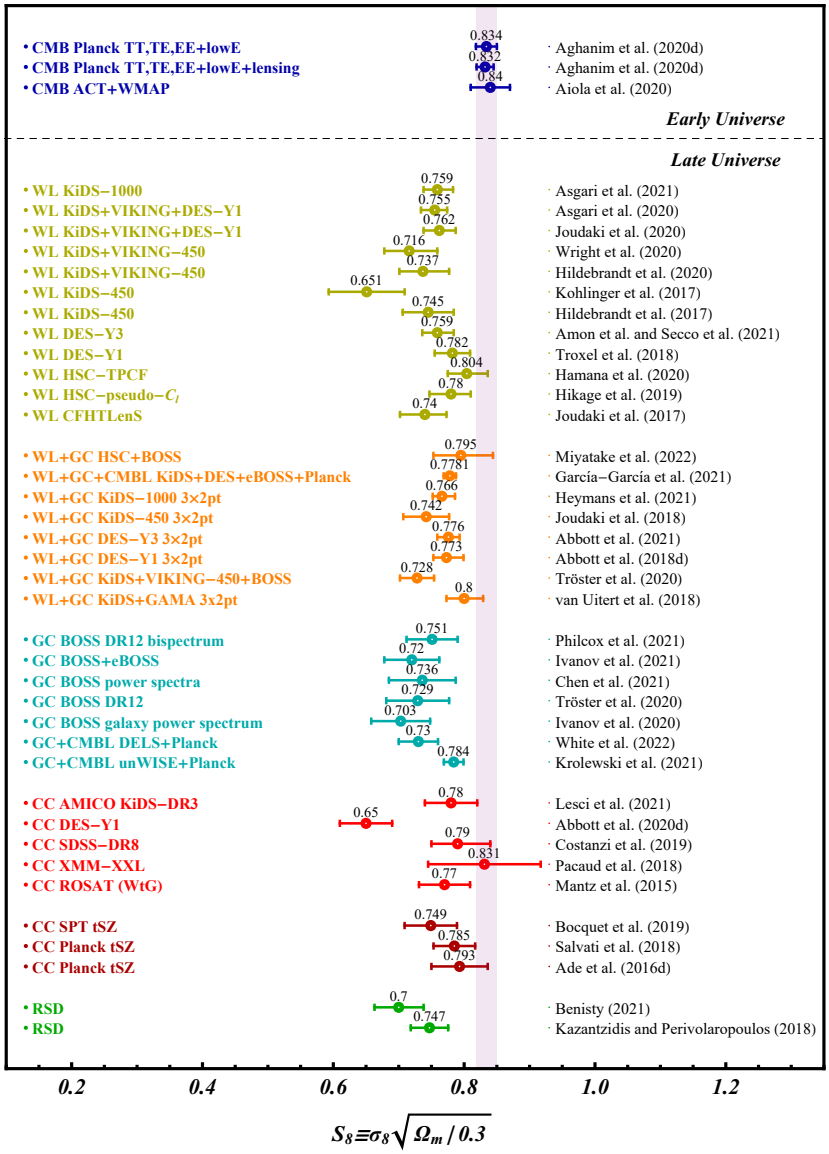

Figure 1.2: Summary of the most recent results of the estimation of the growth of structure parameter $S_8 = \sigma_8 \sqrt{\frac{\Omega_m}{0.3}}$ from different surveys (Abdalla et al., 2022). Studies probing the matter density distribution tend to systematically favour a lower value than studies based on the CMB power spectrum.

has a logarithmic slope of -1 at small radii where $r/r_s \ll 1$ and decreases to -3 at large radii where $r/r_s \gg 1$; the transition is set by r_s

$$\rho_{NFW}(r) = \frac{\rho_s}{(r/r_s)\,(1+r/r_s)^2} \ . \tag{1.16}$$

The halo size $r_{\rm vir}$ and mass $M_{\rm vir}$ are typically set by the expected overdensity at radius of virialisation

$$\rho_{\rm vir} \equiv (1+\Delta_{\rm vir})\bar{\rho}_m \ , \tag{1.17}$$

where $\Delta_{\rm vir}$ is the virial overdensity that is the result of the spherical collapse model we described in section 1.1.1. In practice, other definitions of virial radius and halo mass are used in analytical models and simulations (see White, 2001, for a comparison); in this work, I mostly use $\rho_{\rm vir} \equiv 200\rho_c$ where ρ_c is the critical density.

Given a halo size $r_{\rm vir}$ one can define a measure of how concentrated a given halo is with $c \equiv r_{\rm vir}/r_s$, which measures how extended the core of the halo is.

I will use this profile in Chapter 4 to perform density profile fits, measure concentrations, and calculate the χ^2 of the profiles.

Mass Profile

Integrating the NFW profile is straightforward since it is assumed to be spherically symmetric

$$\begin{aligned} M(<r) &= 4\pi \int_0^r \rho(r) r^2 dr \\ &= 4\pi r_s^3 \rho_s \left[\ln(1+r/r_s) - \frac{r/r_s}{1+r/r_s}\right] \ . \end{aligned} \tag{1.18}$$

Note that for a given definition of virial radius, ρ_s is actually fixed for a given concentration since

$$\begin{aligned} \rho_{\rm vir} &= \frac{M(r_{\rm vir})}{4/3\pi r_{\rm vir}^3} \\ &= \frac{4\pi r_s^3 \rho_s \left[\ln(1+c) - c/(1+c)\right]}{4/3\pi r_{\rm vir}^3} = \rho_s \frac{3\left[\ln(1+c) - c/(1+c)\right]}{c^3} \\ \rho_s &= \rho_{\rm vir} \times c^3/3f(c) \ , \end{aligned} \tag{1.19}$$

where $f(c) \equiv \ln(1+c) - c/(1+c)$ and $\rho_{\rm vir}$ is any chosen definition of virial radius, e.g.

$$\rho_{\rm vir} = (1+\Delta_{\rm vir})\bar{\rho}_m \text{ or } 200\rho_c \ .$$

Projected Surface-Density Profiles

When we observe the density or mass profiles of dark matter haloes or clusters, we observe the density of the halo integrated along the line of sight, which makes the observed surface-density profile different from the 3D one. This gives a surface-density profile (see Coe, 2010, for details):

$$\Sigma(X) = 2\rho_s r_s \frac{1 - f_c(1/X)/\sqrt{|1-X^2|}}{X^2 - 1} \ , \tag{1.20}$$

where $X \equiv R/r_s$ and $f_c(x)$ is $\cos^{-1}(x)$ if $X < 1$ and $\cosh^{-1}(x)$ if $X > 1$. Note that in the limit $X = 1$, $\Sigma(X) = 2\rho_s r_s/3$.

The projected mass profile is

$$M(X) = 4\rho_s r_s^3 \left[\ln \frac{X}{2} + \frac{f_c(1/X)}{X\sqrt{|1-X^2|}} \right] \ . \tag{1.21}$$

Three-Parameter Profiles

Subsequent studies found that three-parameter profiles were more accurate in fitting the simulations data; in particular, profiles allowing the inner density profile parameter γ to be a free parameter were proposed as the generalised Navarro-Frenk-White (gNFW) profile

$$\rho_{gNFW}(r) = \frac{\rho_s}{(r/r_s)^\gamma \left(1 + r/r_s\right)^{3-\gamma}} \ . \tag{1.22}$$

Moore et al. (1998) found that the haloes of their simulation had a cusp $\gamma = 1.5$, significantly steeper than the NFW cusp $\gamma = 1$. This was in contradiction with observations of the rotation curves of dwarf galaxies, which suggests that DM haloes have a core $\gamma = 0$. This is commonly known as the cusp-core controversy, and it is believed to be the result of baryonic effects that flatten the inner profiles of DM haloes.

Further improvement of the fits was found with profiles where the logarithmic slope is a power-law $d\log\rho/d\log r = r^\alpha$, the most universally used being the Einasto profile (Einasto, 1965)

$$\rho_{Ein}(r) = \rho_s \exp\left\{ -\frac{2}{\alpha} \left[\left(\frac{r}{r_{-2}} \right)^\alpha - 1 \right] \right\} \ . \tag{1.23}$$

A step further, recent studies found that the shape parameter α can vary with halo mass (Gao et al., 2008) and/or peak height. In particular, Klypin et al. (2016) find the following fit for haloes in the *MultiDark* simulations

$$\alpha(\nu) = 0.115 + 0.0165\nu^2 \ . \tag{1.24}$$

Concentration

One of the key properties describing a halo is its concentration c. It measures how dense the inner region of a halo is relative to the outer part, or equivalently, what fraction of the halo mass is enclosed in the central part.

Early work using N-body simulations studied the evolution of the halo concentration and found that not only does it evolve with redshift and halo mass, but is also tightly linked to cosmology, the growth history of the halo and environmental effects. (e.g. Navarro et al., 1997; Wechsler et al., 2002; Zhao et al., 2003; Neto et al., 2007; Gao et al., 2008; Zhao et al., 2009).

Evolution with mass and redshift

Initial studies found that the median concentration of haloes is a monotonically decreasing function of mass and redshift. Navarro et al. (1997) modelled the concentration dependence on redshift and mass by relating the density at the scale radius r_s to the density of the Universe when 50% of the halo mass was in progenitors of a mass of 1% or more. Bullock et al. (2001) showed that this model failed to reproduce their data and provided a model where $c \propto 1/(1+z)$. Later on, Zhao et al. (2009) noticed a flattening in the concentration evolution at high z for the most massive haloes; this happens because, for low-mass haloes, the structural parameters vary very little at low-z, suggesting very stable inner structures, and since, by definition, the virial radius increases with time, the concentration of these haloes increases with decreasing redshift. However, for larger mass haloes, the inner structure continues to evolve quickly by accreting matter. Hence the concentration evolution is slower.

Prada et al. (2012) used the ratio $V_{\rm max}/V_{\rm vir}$ as their prescription for studying halo concentration, where $V_{\rm max}$ is the maximum circular velocity and

$$V_{\rm vir} = \sqrt{\frac{GM_{\rm vir}}{r_{\rm vir}}} \ , \tag{1.25}$$

where $M_{\rm vir}$ and $R_{\rm vir}$ are the virial mass and radius. They argue that since NFW is not accurate for all haloes, particularly the most massive ones, the concentration defined for the NFW profile does not reflect the true halo concentration and depends on the profile one fits. So they propose $V_{\rm max}/V_{\rm vir}$ as a profile independent measure of concentration, which for NFW can be expressed as

$$\frac{V_{\rm max}}{V_{\rm vir}} = \left(\frac{0.216c}{f(c)}\right)^{1/2}, \tag{1.26}$$

where $f(c) \equiv \ln(1+c) - c/(1+c)$ as defined above.

With this method, they find an upturn in the concentration-mass-redshift relation at high z for the most massive halos, taking a step further the flattening observed by Zhao et al. (2009). They also show that, despite the apparent complexity of the evolution of the median concentration as a function of z and M, the relation is much simpler if expressed in terms of the rms of the density fluctuation field $\sigma(M,z)$ (or equivalently the peak height $\nu = \delta_c/\sigma(M,z)$), and that $c(\log \sigma^{-1})$ has a clean U-shape.

Concentration and growth history of halos

Haloes grow through mergers and smooth accretion; consequently, the inner density of a halo is tightly linked to the density of the Universe when this central part was first accumulated. Empirical models relating the concentration of haloes to their growth history have emerged since the initial works of Wechsler et al. (2002) and Zhao et al. (2003), who linked concentration to transitions in growth rates of haloes. They argued that the concentration is set when a halo transitions from the fast growth phase to a slower one. Going a step further, Ludlow et al. (2013) claimed that the accretion history and the mass profile of a halo are self-similar if we express the accretion history in terms of the density of the Universe $M(\rho_{crit}(z))$ instead of $M(z)$ and the halo profile using the mass as a function of the average enclosed density $M(\langle\rho(r)\rangle)$. This can be understood as each "layer" of the observed halo has a density related to the density of the Universe when this "layer" was accreted.

Correa et al. (2015b) further explored the link between concentration and halo mass accretion history, relating the concentration to the formation redshift defined as z_s the redshift at which the halo mass was equal to the mass inside the r_s sphere where r_s is the NFW scale radius. They found a correlation between the density within the NFW scale radius $\rho(< r_s)$ and the critical density at the formation redshift as they defined it $\rho_{crit}(z_s)$ of the form

$$\langle\rho\rangle\,(< r_s) = 900 \pm 50\rho_{crit}(z_s), \tag{1.27}$$

and used this relation to relate the concentration to z_s.

1.3 Numerical Cosmological Simulations

We have seen previously that structure formation happens in the highly non-linear regime of cosmological perturbations, making it impossible to represent accurately through analytical models. It is even more complex to model the physics of galaxy and stellar formation when we account for baryonic effects. As for most domains, we can resort to numerical methods when analytical ones reach their limit. In the case of cosmological simulations that follow the evolution of a large number of galaxies and dark matter haloes in a cosmological-size box, there are two main categories, N-body simulations, where the only physics is gravitational, and hydrodynamical simulations, where baryonic physics is accounted for.

1.3.1 N-body Dark Matter Only Simulations

The physics behind structure formation becomes significantly simpler to model when we only consider dark matter, because we know that it is not affected by the electromagnetic interaction. Therefore it is sufficient to model the evolution of a set of particles subject only to gravity. This allows a good prediction of structure formation, in particular at larger scales, at a cost that permits simulating larger portions of the Universe. In N-body simulations, the matter distribution is discretised into mass elements—particles—, which are subject to the gravitational attraction of all other particles. Two main steps are necessary to run an N-body simulation: setting up the initial conditions (ICs) and then evolving the particles with Newtonian gravity.

Initial Conditions

We mentioned in section 1.1.1 that the matter density field is assumed to be an isotropic Gaussian random field. Thus, it is fully described by its power spectrum $P(k)$. The ICs have, therefore, the role of generating an instance of a Gaussian random field, with a particle distribution having the intended power spectrum at a given initial redshift z_i. This can be done in two steps, arranging the particles "randomly", and then moving them to positions where they satisfy the target power spectrum. Neither of these steps is trivial; the arrangement of particles cannot be homogeneous and isotropic; otherwise, the gravitation force on each will be null, and they cannot be white noise; otherwise, we will

generate dense regions which will quickly enter the non-linear regime without additional perturbation. Grid-like distributions can work but will artificially introduce non-physical intrinsic scales (grid spacing). One way to generate a uniform set of points with none of the problems mentioned before is by placing the particles randomly and then evolving them with a repulsive gravitational force until they reach near-equilibrium (e.g. White, 1996).

Once a very uniform distribution of particles has been created, one can use linear perturbation theory to generate a set of particles with the desired power spectrum. A popular approach is using the Zel'dovich (Zel'dovich, 1970) approximation where displacements scale linearly with the growth factor D. This consists of generating the gravitational potential of the desired perturbation Φ and then moving each particle according to

$$\boldsymbol{x}(t) = \boldsymbol{x_i} - \frac{D(a)}{4\pi G\bar{\rho}_i}\boldsymbol{\nabla}\Phi,$$

which comes from the Zel'dovich approximation.

Recently, more accurate approaches have been developed using second-order Lagrangian perturbation theory (Jenkins, 2010) which adds a second-order term to calculate the displacement of the form $D_2\boldsymbol{\nabla}\Phi_2$, here $D_2 \approx -3D^2/7$ and Φ_2 is the potential of a "second-order overdensity" δ_2 which can be derived from the potential Φ

$$\delta_2 = \frac{1}{2}\left[(tr(\Phi_{,ij}))^2 - tr((\Phi_{,ij})^2)\right] \ , \tag{1.28}$$

where tr is the trace and $\Phi_{,ij} \equiv \frac{\partial^2\phi}{\partial q_i \partial q_j}$ is a tensor.

N-body Evolution

The most straightforward way to evolve a set of particles subject only to Newtonian gravity is the **Particle-Particle (PP)** method, which simply consists of calculating the force (acceleration) on each particle induced by the gravitational attraction of all other particles. This acceleration is then integrated numerically to find the velocity and next position

$$a_i \propto \sum_j \frac{m_j}{r_{ij}^2} \ , \tag{1.29}$$

where r_{ij} is the distance between particles i and j. The computational cost of this method scales as N^2, where N is the total number of particles, which quickly becomes prohibitive. Two approximate methods allow for much faster computation.

The **Tree** method consists of grouping particles in boxes with increasing sizes the further they are from the particle of interest. This allows an accurate calculation of the force from close particles but significantly reduces computation time in the force calculation from distant particles. If the size of the boxes scales exponentially with the distance, e.g.doubles every d units, then the cost of calculating the force on each particle scales as $\log N$, therefore $N \log N$ for all particles. This is significantly less costly than the PP method.

The second widely used alternative is the **Particle Mesh (PM)** method, which takes an Eulerian approach to solve the problem. The idea is to subdivide the space into a grid of a given size L and store the kinematic properties of all the particles in each cell, then move the particles according to the properties of each cell. This method is very fast (it scales with N) but requires a significant amount of memory to store the information of each cell, which typically sets the limit on L. Furthermore, we can easily see how small-scale force contributions in this method can be inaccurate, but it has the advantage of being much more accurate than the Tree method on the contributions of particles that are far away. Some codes adopt a hybrid approach, "**TreePM**". It consists of dividing the force into contributions of close and distant particles, where the close short-range force is calculated through the Tree algorithm and the distant particles' contributions through the PM algorithm.

During my thesis, I have used and run simulations made with the publicly available code Gadget 4[1] (Springel et al., 2021), which uses the hybrid **TreePM** method.

1.3.2 Hydrodynamical Simulations

N-body simulations are reasonably good at modelling the formation of structures in the Universe. However, in order to study the complex physics of clusters, galaxies, and gas, such as gas cooling, turbulence and stellar feedback, one needs to solve the hydrodynamical equations governing the evolution of different types of fluids. Both Lagrangian particle-based and Eulerian grid-based approaches have been developed to model the behaviour of the baryonic fluid.

The most widely used Lagrangian approach is the **Smoothed Particle Hydrodynamics (SPH)** developed by Lucy (1977), Gingold & Monaghan (1977) and Evrard (1988, 1990) where a set of N particles samples the fluid. Any specific field $F(\boldsymbol{x})$ (e.g. energy or momentum) can then be calculated by adding all the relevant contributions from particles

[1] https://wwwmpa.mpa-garching.mpg.de/gadget4/

at the position $\boldsymbol{x}$ given a particle smoothing kernel $W(r, h)$

$$F(\boldsymbol{x}) = \sum_{j=1}^{N} m_j \frac{F_j}{\rho_j} W(\boldsymbol{x} - \boldsymbol{x_j}, h) \ , \tag{1.30}$$

where m_j, ρ_j and F_j are the mass, density and field value of the particle j. The smoothing kernel determines how much a particle contributes to the value of the field F in the position $\boldsymbol{x}$ depending on its distance from $\boldsymbol{x}$.

Given this discretisation of the fluid, one can use numerical integration methods to solve the Lagrangian formulation of the hydrodynamical equations.

In contrast to this, Eulerian approaches divide the volume into grid cells and store the physical properties of each cell, calculate the flux of energy, momentum and mass at each boundary and evolve the physical properties at each cell. The most popular grid-based algorithm is the **Adaptative Mesh Refinement (AMR)** first described in Berger & Oliger (1984) and Berger & Colella (1989), which uses grids of adaptative sizes to have higher resolution in dense regions. The Euler equations —typically in integral form to handle shocks— are solved on the grid. Then numerical integration methods are used to find the temporal evolution of the different physical properties. However, this is insufficient to model all aspects of galaxy formation, particularly at smaller scales. "Subgrid" prescriptions are typically added to account for star formation, supernova feedback, Active Galactic Nuclei (AGN) feedback, black hole formation and other physical phenomena that go into galaxy formation.

While hydrodynamical simulations play an important role in the description of the complex physics of galaxy formation, they are computationally expensive, limiting their resolution and box size. The former forces the use of subgrid prescriptions, and the latter forbids their use on the largest cosmological scales.

An alternative to hydrodynamical simulations is the use of semi-analytical and semi-empirical models. These consist of using a large box N-body simulation and populating dark matter haloes with galaxies and clusters either using subgrid prescriptions in the case of semi-analytical models or with population statistics obtained through observations in the semi-empirical case.

1.3.3 Post Processing: Finding Haloes

The output of an N-body simulation, at any given time, is the positions and velocities of all the discretised mass elements —particles— used. This gives no real information

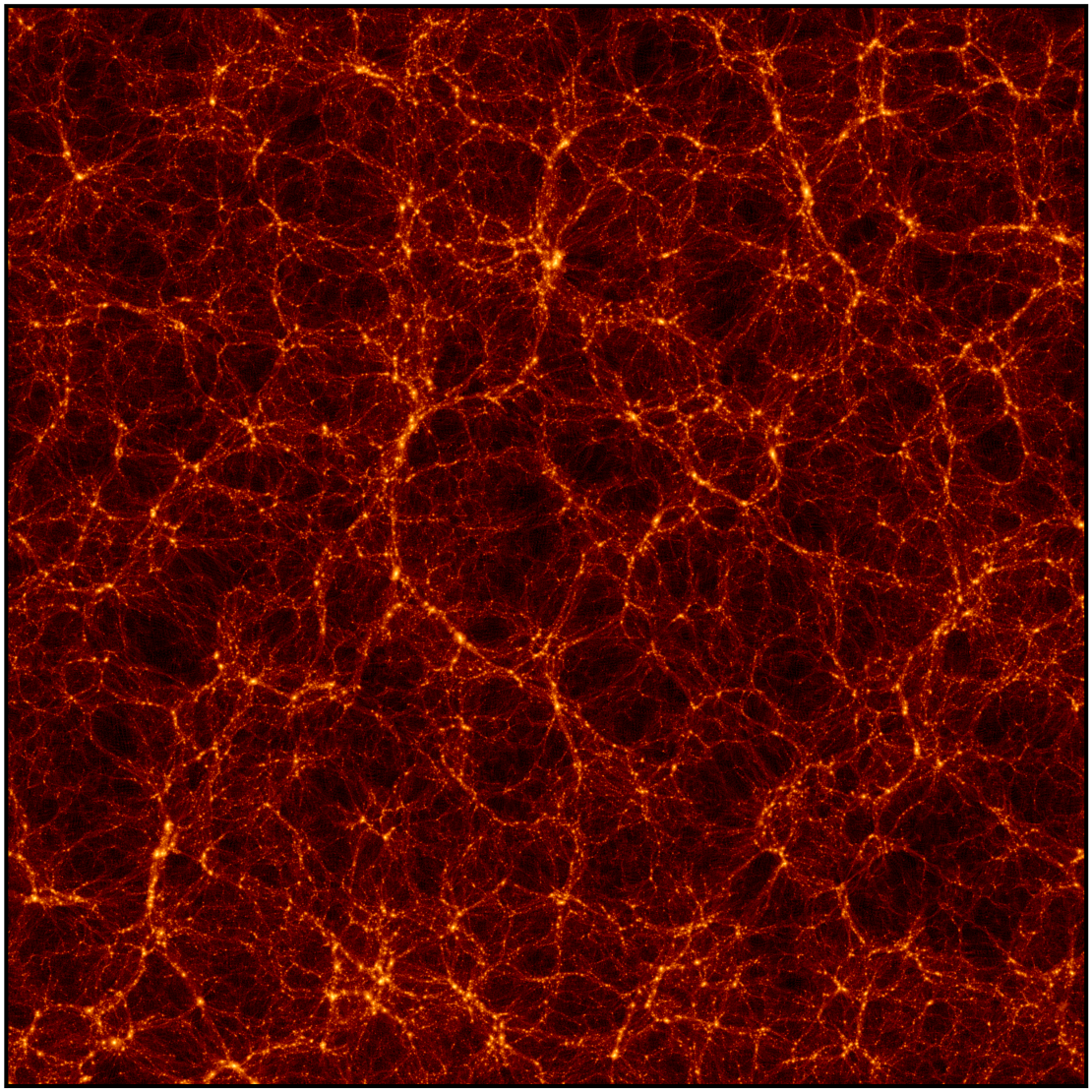

Figure 1.3: Slice of an output of one of the N-body simulations I have run (see Chapter 4.3.1 for details). It illustrates the distribution of large-scale structures. The length of the box is $L = 500$Mpc/h, and the brightness measures the local density. Local bright spots are haloes linked together through thinner filaments of matter.

about the structures in the Universe which, in ΛCDM, organise hierarchically in haloes and subhaloes, linked through filamentary structures as can be seen in Fig. 1.3.

Therefore, we need an algorithm to group particles that are supposed to be part of the same structure. Two main approaches to accomplish this task are using a clustering algorithm to group particles together, or using the local density to assess whether each particle is supposed to be part of a collapsed structure. The **Friends of Friends (FoF)** algorithm (Huchra & Geller, 1982; Press & Davis, 1982; Einasto et al., 1984) is a standard clustering method; in the case of N-body simulations, it links together particles that are within a linking length from each other (typically around 0.2 times the average inter-particle separation). The **Spherical Overdensity (SO)** (Lacey & Cole, 1994) algorithm uses contours of isodensity to detect regions that are above the chosen threshold. For haloes, it is usually chosen to be the density at the virial radius, found through the spherical collapse model (see section 1.1.1). More complex criteria are typically used for the substructure.

Most of the analysis of this work has been done using the publicly available code *Amiga Halo Finder* (AHF)[2] (Knollmann & Knebe, 2009), which uses a spherical overdensity method. I also used data results from the *Subfind* (Rodriguez-Gomez et al., 2015) and *Rockstar*[3] (Behroozi et al., 2013a) codes.

1.4 Clusters of Galaxies

Galaxy clusters are the largest bound structures in the Universe; as such, they represent a unique laboratory for studying galaxies' properties, the Universe's large-scale structure, and the nature of dark matter. Their large gravitational potential results in a hot gas distributed between galaxies in Intra-Cluster Medium (ICM). Thanks to their large mass, hot gas and galactic component, they can be detected in a variety of ways.

X-ray

Diffuse gas accounts for most of the baryons in the Universe, and is usually difficult to observe. But in clusters of galaxies, the gas in the ICM gets heated up to X-ray emitting virial temperatures ($10^7 - 10^8$K). Since clusters are the only large, luminous, spatially extended, and non-varying X-ray source, they are relatively easy to detect in the high-energy range.

[2] http://popia.ft.uam.es/AHF/
[3] https://github.com/yt-project/rockstar

The ICM is a hot, tenuous plasma with densities ranging from 10^{-5} to 10^{-1} cm^{-3} and energies of several keV per particle, which causes the hot plasma to emit the bulk of the thermal energy in the regime of soft X-rays. Bremsstrahlung is the dominant radiative process at high temperatures relevant for massive clusters and has a spectrum of the form (Böhringer & Werner, 2010)

$$e(\nu, T_e) \sim n_e^2 T_e^{-0.5} \exp(-h\nu/kT_e) \ , \tag{1.31}$$

where T_e and n_e are the temperature and density of electrons. The most important feature of the spectrum is the sharp cut-off at high energy that provides a good temperature estimation.

The normalisation is given by the emission measure, E, which is the rate of the free-free events. It is proportional to the integral along the line of sight of the squared plasma density n_e^2

$$E = \int_{los} n_e^2 dV \tag{1.32}$$

On top of the thermal continuum, we also find line emission with increasing contribution with decreasing plasma temperatures.

In general, the X-ray spectrum provides temperature and chemical composition information through the general shape and specific features.

If we want a mass profile, we need to assume that the ICM is in hydrostatic equilibrium. This is justified by the fact that the sound crossing time through the cluster core (10^8 years) is generally much less than the age of the clusters (several 10^9 years). Thus, if we can measure the density and temperature distribution, we can measure the mass distribution. Through simple thermodynamics derivations, we get

$$M(r) = -\frac{kT(r)}{G\mu m_p} \left(\frac{d\ln n}{d\ln r} + \frac{d\ln T}{d\ln r} \right) \ , \tag{1.33}$$

where μm_p the mean molecular mass and $n(r)$ is the gas particle density.

SZ effect

When low-energy photons of the Cosmic Microwave Background (CMB) travel through the hot ICM, they get scattered by the very energetic free electrons of the plasma and experience an energy boost, producing a shift in the CMB spectrum. This is known as the thermal Sunyaev-Zel'dovich effect (SZ or tSZ hereafter) (Sunyaev & Zeldovich, 1972).

In the non-relativistic limit, one can express the SZ distortion as a change in CMB temperature ΔT_{SZ} with a simple formula (Carlstrom et al., 2002)

$$\frac{\Delta T_{SZ}}{T_{CMB}} = -2y \ , \tag{1.34}$$

where y, the Comptonization parameter, is the key measure of the SZ effect and can be interpreted as a dimensionless measure of the time CBM photons spend in the electron cloud, or alternatively the pressure integrated along the line of sight and has the form

$$y \equiv \int n_e \sigma_T dl \frac{kT_e}{m_e c^2} \ , \tag{1.35}$$

where n_e, T_e, and m_e are the electron particle density, temperature and mass, respectively, and σ_T is the Thomson cross-section.

Note that the SZ distortion is redshift independent. Hence, SZ is particularly useful for detecting high-z clusters. We can also notice that it is proportional to the integral along the line of sight of the electron density times its temperature $n_e \times T_e$, whereas X-ray surface brightness is proportional to the integral of $n_e^2 \times T_e^{-1/2}$. Thus, the angular distribution of SZ and X-ray images will likely differ.

Given the different specific properties of the SZ effect seen above, we can draw several conclusions on cluster detection:

- Combining SZ and X-ray data could provide 3D maps of clusters. However, the change in the 3D shape is not sensitive enough to be an efficient tool.
- Combining SZ and X-ray data on clusters provides valuable information on the thermal composition of ICM, as they depend on T_e and n_e in different ways. SZ is a map of electron pressure and is sensitive to n_e while X-ray is sensitive to n_e^2, so SZ is sensitive to the outer part of the cluster where the X-ray signal drops.
- SZ effect is a direct measurement of the projected mass of gas along the line of sight. This means that one can compare it to other total mass (gas + DM) probes (e.g. lensing) and get information on the baryon composition and fraction in clusters.

Gravitational Lensing

Because of their very large mass, clusters are expected to produce effective gravitational lensing phenomena whenever they are in the path of an observed background object. This

allows their detection and mass and density profile measurements (Hattori et al., 1999; Mellier, 1999; Mandelbaum, 2018, for eview).

The observable is usually a source galaxy shape, either through strong lensing events generating giant arcs (assuming a spherically symmetric shape) or through averaging many source galaxy shapes (assuming no bias in the ellipticity distribution).

One method of using arcs and arclets is to assume a mass profile model with one parameter and fit the model. This would be through fitting the shear profile $\gamma(\boldsymbol{\theta})$, which is related to the projected surface-density profile $\Sigma(\boldsymbol{\theta})$. The mass model is generally described as a linear sum of clumps modelled analytically to avoid having too many parameters to constrain compared to the number of observational constraints.

Optical

While X-ray and SZ methods detect the ICM and weak lensing the overall mass distribution, dominated by the dark matter halo, one can also detect a cluster through its galaxies using optical/IR telescopes. For large enough clusters, the radial distribution of the galaxy number density provides very useful information about its density profile and concentration and is a proxy for its total mass.

It is, however, not trivial to accurately associate a given observed galaxy with a cluster, particularly if measured redshifts are not very precise. Moreover, background and foreground galaxies can be falsely associated with a cluster, making richness and profile measurements challenging when using only galaxies.

Current and Forthcoming Surveys

We have seen that clusters can be detected in optical, X-ray and radio wavelengths. At the time of writing, a number of surveys, some of which are already gathering data, are expected to increase the number and accuracy of cluster detections by more than an order of magnitude. I will review some of the current and future prospects of each.

EUCLID

EUCLID[4] is an optical/NIR space telescope which is scheduled to launch around the time of writing this thesis. It will cover 15,000 deg^2, which represents about a third of the sky and is expected to detect O(10^5) clusters up to z=2 (Sartoris et al., 2016) at a 5σ level,

[4] https://www.euclid-ec.org/

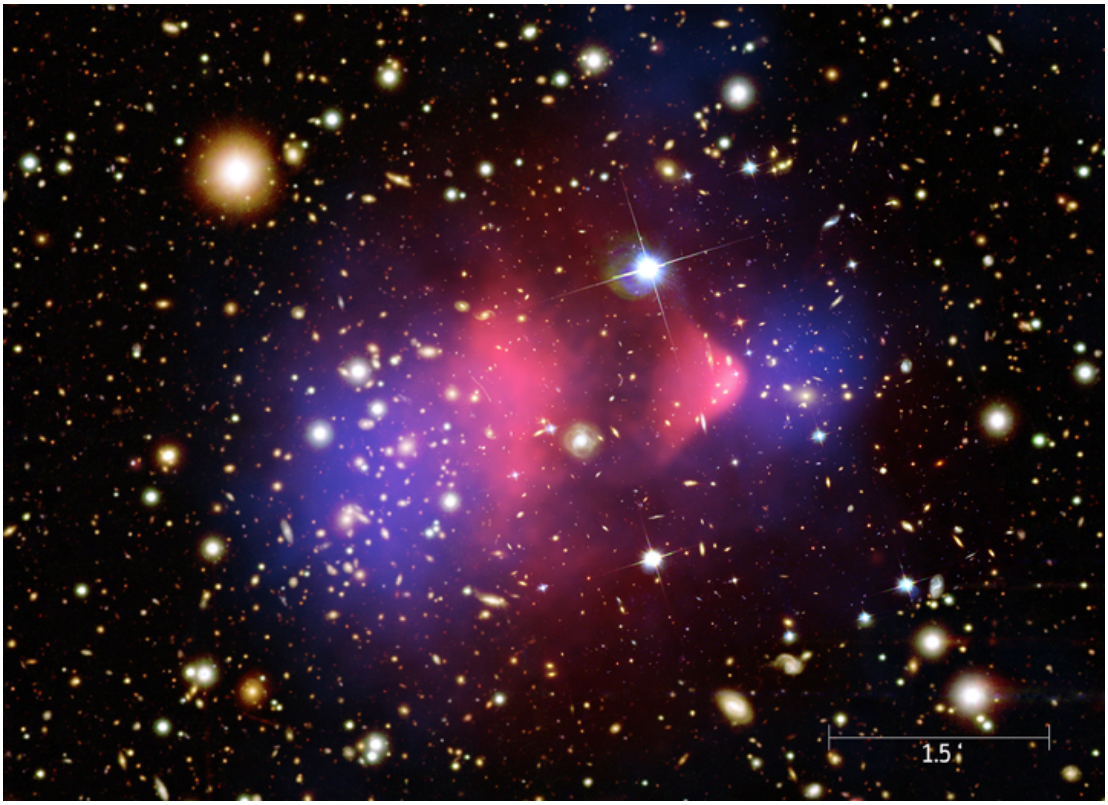

Figure 1.4: The Bullet Cluster (1E 0657-56) observed through different channels. The optical image of the background galaxies is from *Hubble Space Telescope*(HST), while the red and blue images show X-ray emission (from the *Chandra* space telescope) and the mass profile reconstructed from lensing using HST data. The Bullet Cluster is a famous example of two clusters merging, as seen through lensing that follows the mass component. Note that the ICM gas from the two clusters traced by X-ray is interacting, and caught up in a shock between the two mass peaks, while the galaxies and most of the mass have passed through each other. This is one of the strongest pieces of evidence that most of the matter in clusters is collisionless, and thus not baryonic.

which will represent almost all clusters above $10^{14} M_{\odot}$. Of these, almost 20000 will be very high signal-to-noise detections —7σ—. This represents more than an order of magnitude increase from existing cluster surveys.

***Vera C. Rubin Observatory* (LSST)**

The *Rubin*[5] observatory is a ground-based telescope located in the Atacama desert in Chile, which is also set to start collecting light around the time of writing this thesis. It will cover about 18,000 deg^2 of the sky and boasts the largest digital camera ever constructed, observing the optical and NIR. It is expected to detect over 100,000 clusters up to z=1.2 (LSST Dark Energy Science Collaboration (LSST DESC) et al., 2021).

Roman Space Telescope

Roman (ex-WFIRST)[6] (Akeson et al., 2019) is a NIR deep-field survey expected to be launched in the mid-2020s. It will have a primary mirror with a diameter 2.4 meters, the same size as the *Hubble Space Telescope* and will have a Wide Field Instrument that covers a 0.281 deg^2 field, nearly 100 times larger than HST. It will be able to detect lower-mass clusters at high redshift and provide crucial hints on cluster and proto-cluster formation and evolution, which makes it very complementary to wide surveys such as *Euclid* and *Rubin*. In particular, its high density of lensed galaxies (50 galaxies per square arcminute) will produce more detailed weak lensing maps of dark matter in clusters.

The Dark Energy Spectroscopic Instrument (DESI)

DESI[7] is an ongoing ground survey located in Arizona, whose goal is to provide a 3D map of the sky by measuring spectra of about 30 million galaxies and cover 14,000 deg^2 over a five-year span (DESI Collaboration et al., 2016). It will provide accurate spectroscopic redshifts of galaxies and clusters complementing photometric data from *Roman*, *Euclid* and *Rubin*.

CMB-S4

CMB-S4[8] is a next-generation radio survey consisting of 21 ground-based telescopes at the South Pole and in the Chilean Atacama desert. It will be able to detect clusters through analysis of the change of the CMB spectrum through the SZ effect. It will detect clusters up to z=3, and will increase the number of SZ-clusters from the hundreds currently with SPT-SZ to almost a hundred thousand (Abazajian et al., 2019a). This is particularly complementary with optical surveys detecting galaxies while SZ probes ICM gas.

[5] https://www.lsst.org/
[6] https://roman.gsfc.nasa.gov/
[7] https://www.desi.lbl.gov/
[8] https://cmb-s4.org/

eROSITA

eROSITA (extended ROentgen Survey with an Imaging Telescope Array)[9] is a wide-field X-ray survey on-board of the Russian-German *Spectrum-Roentgen-Gamma* (SRG) space observatory. The survey started collecting light in December 2019 but is paused at the time of writing as a consequence of the Russian invasion of Ukraine. The main science target of eROSITA is the detection of a large sample ($\sim 100,000$) of galaxy clusters out to $z > 1$ (Merloni et al., 2012) targeting the ICM gas in soft X-rays in the energy range 0.2–8 keV. These clusters will also be detected in optical thanks to the surveys listed above, providing a unique opportunity to study how the ICM gas, the galaxies and the dark matter halo interact in each cluster.

1.5 This work

In light of the increase in quantity and accuracy of the expected galaxy cluster data discussed above, we want to provide novel ways to extract information about structure formation and the background cosmology. Specifically, this thesis aims to demonstrate the possibility that cluster structural properties can be used to constrain the cosmological parameters Ω_m and σ_8.

To achieve this, we use the fact that the structure of a galaxy cluster is mainly determined by aspects of its dynamical state; this means how perturbed it is, whether it had a recent merger, and the state of the Universe when it was formed. Because cosmology, more specifically Ω_m and σ_8, will influence the average dynamical state of clusters, we can predict how Ω_m and σ_8 will impact the distribution of cluster structural properties, such as their shapes and density profiles. This allows one to obtain constraints on Ω_m and σ_8 by studying the statistics of the structural properties of a population of clusters.

In Chapter 2, I study how the age at which galaxy clusters form depends on Ω_m and σ_8 and compare this dependence to cluster abundance, the typical way to constrain cosmology with clusters. I do this first with analytical predictions, and then compare them to DM-only cosmological simulations.

I extend the study in Chapter 3 by looking at the differential version of these tests, namely the growth and merger rate of clusters and DM haloes. Here again, I use both analytical predictions and tests using simulations.

[9] https://erosita.mpe.mpg.de/

The first two chapters lay the ground for the theoretical predictions of how cluster dynamical state varies with Ω_m and σ_8; In Chapter 4, I explore how to link the cluster dynamical state to potentially observable structural properties. I do this by studying correlations between different structural properties, using a set of high-resolution DM-only simulations, run for a range of cosmologies.

Finally, I summarize the findings of this thesis and give the main conclusions and outlook for the future in Chapter 5.

Chapter 2

Cluster Assembly Times as a Cosmological Test

2.1 Abstract

The abundance of galaxy clusters in the low-redshift universe provides an important cosmological test, constraining a product of the initial amplitude of fluctuations and the amount by which they have grown since early times. The degeneracy of the test with respect to these two factors remains a limitation of abundance studies. Clusters will have different mean assembly times, however, depending on the relative importance of initial fluctuation amplitude and subsequent growth. Thus, structural probes of cluster age such as concentration, shape or substructure may provide a new cosmological test that breaks the main degeneracy in number counts. We review analytic predictions for how mean assembly time should depend on cosmological parameters, and test these predictions using cosmological simulations. Given the overall sensitivity expected, we estimate the cosmological parameter constraints that could be derived from the cluster catalogues of forthcoming surveys such as *Euclid*, the *Nancy Grace Roman Space Telescope*, *eROSITA*, or CMB-S4. We show that by considering the structural properties of their cluster samples, such surveys could easily achieve errors of $\Delta\sigma_8 = 0.01$ or better.

2.2 Introduction

The 'concordance' Λ-Cold Dark Matter (ΛCDM) cosmological model is now well established as a single theoretical framework that is consistent with many different observational tests. The present-day abundance of dark matter and dark energy and the statistical properties of the matter distribution are increasingly well constrained, as expressed by cosmological parameters with gradually decreasing uncertainties (e.g. Planck Collaboration et al., 2020). As the uncertainties in parameter values shrink; however, they reveal tension in several places in the model. Most notably, the Hubble parameter H_0 appears to differ significantly between high-redshift and low-redshift tests, with the tension in independent measurements of this parameter now exceeding 4σ (Riess et al. 2019; see Verde et al. 2019 for a review). In addition to the H_0 tension, there is also growing evidence that the amplitude of the matter fluctuations (typically expressed as σ_8, the r.m.s. of fluctuations in the matter density on a scale of $8h^{-1}$Mpc) may display a similar tension at the 2–3σ level (or ~ 0.1 in this parameter, e.g. Battye et al., 2015; Douspis et al., 2019; To et al., 2021; Heymans et al., 2021a). More generally, the fundamental natures of dark energy and dark matter remain unknown, raising the possibility of new, exotic physics not yet included in the standard cosmological model.

Given the evidence for tension in the current results, multiple, independent tests of the standard cosmological model are needed, on different mass and length scales and at different redshifts, to either reconcile the current results, or to reveal the physical origin of the disagreements. Current and forthcoming space missions, including *Euclid*, the *Nancy Grace Roman Space Telescope* (*Roman*), and *eROSITA* (Pillepich et al., 2012), together with data from large ground-based surveys such as UNIONS (Chambers et al., 2020), DESI (DESI Collaboration et al., 2016) or Rubin LSST (LSST Science Collaboration et al., 2009), or experiments such as CMB-S4 (Abazajian et al., 2019b), will provide remarkable new datasets, mapping out structure over a significant fraction of the observable universe, out to redshifts of a few. Given, on the one hand, the enormous potential of this data, and, on the other hand, the exacting precision required to resolve current parameter tensions, there is a need for new tests of the cosmological model that make full use of our growing understanding of structure formation.

The measured abundance of massive galaxy clusters is a classic cosmological test. Cluster abundance has been estimated using samples detected in the X-ray (e.g. Henry et al., 2009; Mantz et al., 2010; Böhringer et al., 2014), via the Sunyaev-Zel'dovich (SZ) effect (e.g. de Haan et al., 2016; Planck Collaboration et al., 2020), in optical galaxy redshift surveys (e.g. Abdullah et al., 2020), in weak lensing surveys (Kacprzak et al., 2016), or using combinations of these techniques (e.g. Abbott et al., 2020; Costanzi et al., 2021).

These rare objects evolve from peaks in the matter distribution present at early times, and their present-day abundance places a tight constraint on the function $S = \sigma_8 \Omega_{\rm m}^{\gamma}$, where $\Omega_{\rm m}$ is the present-day matter density parameter, and $\gamma \approx 0.5$ is the growth index. Cluster abundance measurements alone do not place very tight constraints on $\Omega_{\rm m}$ or σ_8 individually, due to the degeneracy between them. Simply counting clusters does not leverage the full potential of the underlying datasets, however. The structural properties of clusters — their projected shape, central concentration, substructure and non-axisymmetry — are all related to their degree of dynamical relaxation, which in turn traces their formation history (see Taylor, 2011, for a review). Thus, measurements of these properties provide a separate constraint on the growth rate. While measurements of structural parameters in individual clusters may be noisy, the sheer number of systems expected in forthcoming surveys should allow us to make robust measurements of the average trends, using the expertise developed in fields like weak lensing.

The idea of using the structural properties of clusters to constrain cosmological parameters is not new (e.g. Richstone et al., 1992; Evrard et al., 1993; Mohr et al., 1995), but the context for these tests has changed radically in the 30 years since the idea was first proposed. First, the size of the datasets has grown enormously, giving better statistics. Second, our understanding of the systematics in individual structural measurements has developed considerably. Third, there is increasing sophistication in understanding and exploiting large, complex datasets. In particular, fields such as cosmic shear have illustrated how it is possible to extract parameter constraints from large sets of noisy measurements, even when the link between parameters and observables is indirect and non-linear. Finally, simulations of structure formation have progressed dramatically, allowing us to calibrate some aspects of non-linear structure formation at the per cent level, even if other aspects remain uncertain. Thus, it seems high time to reconsider cosmological tests based on the internal structure of haloes.

In this chapter, we consider the possibility of estimating from their structural properties the mean assembly time or formation epoch for a large sample of galaxy clusters. This measurement of mean 'age' would leverage the same data already collected for cluster abundance studies, providing an independent constraint on the cosmological parameters. We focus in particular on the parametric dependence of halo age, and its sensitivity to the parameters $\Omega_{\rm m}$ and σ_8; in a subsequent chapter, we will consider the (non-trivial) path to developing practical observational tests based on age estimates. The outline of the chapter is as follows. In Section 2.3, we review theoretical models of cluster abundance and age, and use them to predict how these properties vary as a function of the cosmological parameters. Given the approximate nature of the theoretical estimates, in Section 2.4 we test these predictions using catalogues from several different N-body simulations. We

show that with some careful analysis, we can reconcile the analytic and numerical results to reasonable accuracy. In Section 2.5, we estimate the sensitivity a realistic observational program could achieve, using concentration as a proxy for age. Finally, in Section 2.6, we review and summarize our results. The details of the analytic calculations and the dependence of several important quantities on the cosmological parameters are discussed in the appendices. We consider a range of cosmologies throughout the chapter, but assume flatness ($\Omega_{\rm m} + \Omega_\Lambda = 1$), and take a model with $\Omega_{\rm m} = 0.3$ as the fiducial case.

2.3 Cosmological Sensitivity of Halo Abundance and Halo Age

We will begin by estimating the potential sensitivity of age tests, using theoretical predictions of how cluster abundance and age depend on the cosmological parameters. We use analytic models of abundance and age based on the extended Press & Schechter formalism, and calculated using standard tools and techniques summarized in Appendix A.1.

2.3.1 Analytic Models of the Halo Mass Function

The Press–Schechter (PS – Press & Schechter, 1974; Bond et al., 1991) and extended Press–Schechter (EPS – Lacey & Cole, 1993) formalisms provide a convenient analytic framework for computing the number density of dark matter haloes and their growth rate, given a background cosmological model. The basic expression for the halo mass function, derived assuming spherical collapse, is

$$n(M)dM = \sqrt{\frac{2}{\pi}}\frac{\rho_0}{M}\frac{\delta_c}{\sigma^2}\exp\left(-\frac{\delta_c^2}{2\sigma^2}\right)\left|\frac{d\sigma}{dM}\right|dM\,, \tag{2.1}$$

where ρ_0 is the matter density at the redshift of interest, $\sigma = \sigma(M)$ is the r.m.s. amplitude of fluctuations in the density field smoothed on mass scale M, and δ_c is the threshold for collapse to a virialized halo. Although fluctuations grow in amplitude as z decreases to zero, the condition for collapse by redshift z can also be considered at some fixed, early redshift, taking $\sigma(M)$ to be a function of mass only, and $\delta_c = \delta_c(z)$ to be a function of the collapse redshift.

The mass function can also be written in a more compact form as

$$n(M,t)dM = \frac{\rho_0}{M}f_{PS}(\nu)\left|\frac{d\nu}{dM}\right|dM\,, \tag{2.2}$$

where $\nu(M,z) \equiv \delta_c(z)/\sigma(M)$ is the height of the collapse threshold at redshift z, relative to typical fluctuations on mass scale M, and

$$f_{\rm PS}(\nu) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{\nu^2}{2}\right) \tag{2.3}$$

is the mass fraction that has collapsed per unit interval of ν[1].

It is well known, however, that this basic form of the mass function fails to reproduce the halo abundance found in N-body simulations, particularly for low-mass haloes (Sheth & Tormen, 1999, 2002; Jenkins et al., 2001). This failure is due to several simplifying assumptions made in the model, the most important one being a fixed threshold for (spherical) collapse δ_c that is independent of mass and environment. To solve this problem, Sheth & Tormen (1999, ST hereafter) considered a mass-dependent collapse threshold (or 'moving barrier'), to derive a functional form that better fits the mass function from simulations

$$f_{ST}(\nu) = A\sqrt{\frac{2a}{\pi}}\left(1 + \frac{1}{(\sqrt{a}\nu)^{2p}}\right)\exp\left(-\frac{a\nu^2}{2}\right) , \tag{2.4}$$

with $A = 0.322$, $a = 0.707$ and $p = 0.3$.

A number of subsequent studies have improved our understanding of the mass function. Tinker et al. (2008) demonstrated that the HMF is not completely universal, but evolves with redshift; allowing the parameters A, a, and p in the fit to evolve as a power-law of $1+z$ provides a better match to simulations. This non-universality has since been confirmed by other groups (e.g. Watson et al., 2013). Despali et al. (2016), argued that it is in fact, an artifact of the halo mass definition, and that the common choices of overdensity of 200 or 178 times the critical density induce much of the non-universality. Finally, several authors (Velliscig et al., 2014; Bocquet et al., 2016; Castro et al., 2021) have studied the impact of baryonic effects on the HMF by measuring halo masses, profiles, and abundance in hydrodynamic simulations. These improvements to the HMF fit are required in precision applications, but are generally secondary in importance ($\leq$ 20% – Tinker et al. 2008; Velliscig et al. 2014; Bocquet et al. 2019a), relative to the large variations in abundance with cosmology shown below. Thus, for simplicity in what follows, we will assume the ST form of the collapsed fraction (Eq. 2.4), in order to study how abundance and age depend on cosmology. We discuss the possible effect of baryons on the *internal* structure of haloes in Section 2.5.4 below.

[1]Note the mass fraction is often defined per unit $\ln\nu$; the expressions for $f_{\rm PS}$ and $f_{\rm ST}$ (Eqs. 2.3 & 2.4) then contain an extra factor of ν.

2.3.2 Cosmological Dependence of Halo Abundance

Cluster abundance depends on the cosmological parameters both through the cluster mass function and through the survey volume. Within a survey volume subtending a solid angle $\Delta\Omega$, the expected number of clusters in the mass bin i: $[M^i, M^{i+1}]$ and redshift bin j: $[z_j, z_{j+1}]$ is

$$N(M_i, z_j) = \frac{\Delta\Omega}{4\pi} \int_{z_j}^{z_{j+1}} dz \frac{dV}{dz} \int_{M_i}^{M_{i+1}} \frac{dn}{dM}\, dM\,, \tag{2.5}$$

where dn/dM is the HMF given above and dV/dz is the volume element per unit solid angle and per unit redshift. As discussed previously, the HMF is calculated as a fraction of the material in a region that has collapsed to form haloes on some mass scale. Thus, rather than relating the halo abundance to the volume probed by the survey, we can express it in terms of the total mass M_V of material in the survey volume:

$$N(M_i, z_j) = \frac{\Delta\Omega}{4\pi} \int_{z_j}^{z_{j+1}} dz \frac{dM_V}{dz} \int_{M_i}^{M_{i+1}} f(\nu) \frac{d\nu}{dM} dM\,. \tag{2.6}$$

The advantage of this form is that we can now separate the cosmological dependence of the first factor, the total mass M_V probed by the survey in a given redshift range Δz, from that of the second factor, which is the fraction $f(\nu)\Delta\nu$ of that mass that has collapsed to form haloes in the mass range $\Delta\nu = (d\nu/dM)\Delta M$ by that redshift.

To make explicit the cosmological dependence of the HMF, in Appendices A.2 and A.3 we consider each of these factors separately. Over a realistic range of ($\Omega_{\rm m}$, σ_8), the survey mass M_V varies by a factor of $\lesssim 2$, while the collapsed fraction can vary by several orders of magnitude, and thus dominates the parametric dependence of the HMF.

As demonstrated in Appendix A.3, the peak height ν varies approximately as $\nu(M, z) \propto (\sigma_8)^{-1}\Omega_{\rm m}^{\alpha(z)}\Omega_{\rm m}^{-\beta(M)}$. The resulting behaviour in the $\Omega_{\rm m}$–σ_8 plane is shown in Fig. 2.1, for several mass/redshift combinations. We see that peak height depends mainly on σ_8; variations in $\Omega_{\rm m}$ introduce a slight tilt in the contours, that goes from negative at low mass/redshift, where $\beta(M) > \alpha(z)$, to slightly positive at high mass/redshift (and low $\Omega_{\rm m}$), where $\alpha(z) > \beta(M)$ (see also Appendix A.3, and the lower right panel of Fig. A.2, which shows the dependence on $\Omega_{\rm m}$ for fixed σ_8, at several different masses and redshifts).

In both spherical collapse (PS – Eq. 2.3) and ellipsoidal collapse (ST – Eq. 2.4) models, the collapsed fraction exhibits power-law growth for $\nu < 1$, and an exponential decay for $\nu > 1$. Thus, there are two main regimes, the first where the abundance of haloes increases with ν, and the second where it decreases rapidly. The effect of the cosmological parameters

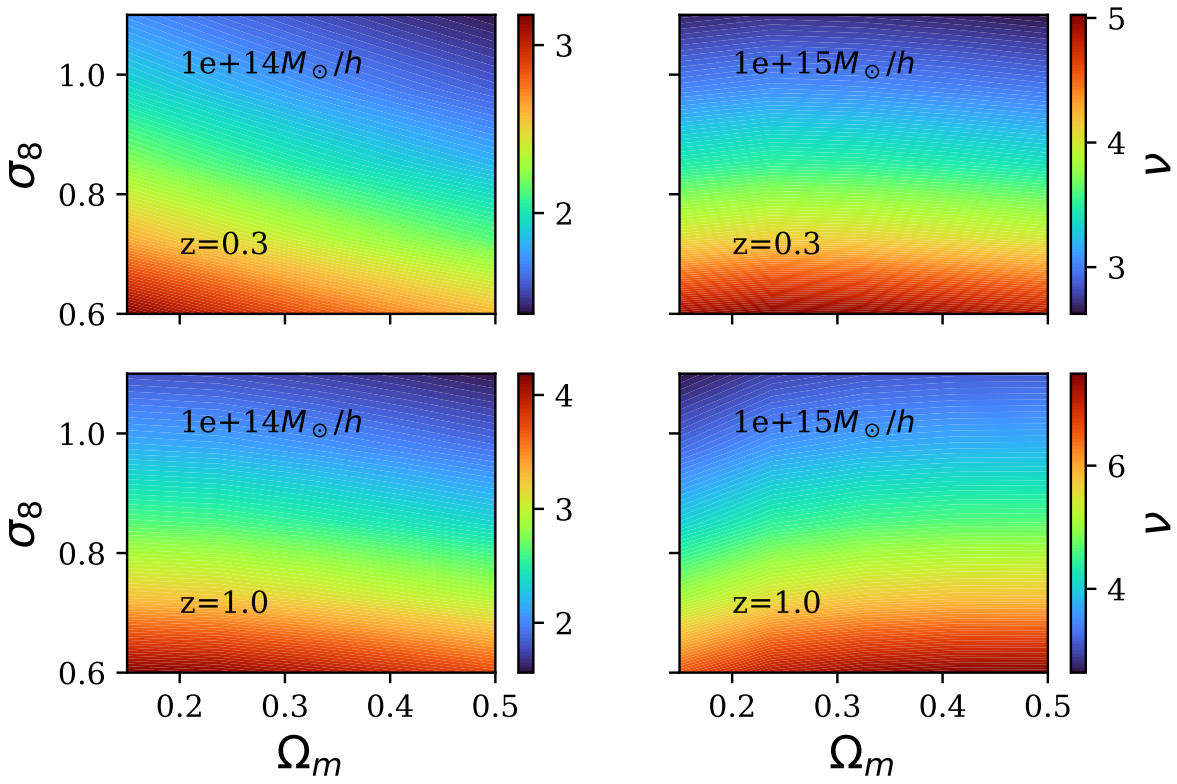

Figure 2.1: Variation of the peak height ν, for a halo of the mass and redshift indicated, in the $\Omega_{\rm m}$–σ_8 plane.

on the two regimes is shown, for instance, in Fig. A.4. The amplitude of fluctuations σ_8 controls the mass at which the transition between regimes occurs, while Ω_m controls the sharpness of the transition. In the case of clusters, we are generally in the second regime, where increases in ν produce an exponential decrease in abundance. Thus, the parametric dependence of the collapsed fraction (shown in the right panel of Fig. A.1) is very similar to the corresponding figure for peak height, Fig. 2.1, but with an inverted and logarithmic colour scale, since $f(\nu) \propto \exp(-\nu^2/2)$ implies that in a log scale, $\ln(f) \propto -\nu^2/2$.

Finally, we can combine the parametric dependence of the survey volume and the collapsed fraction (shown in the left and right panels of Fig. A.1 respectively) to plot the dependence of number counts on Ω_m and σ_8. This is shown in Fig. 2.2, for the same mass/redshift combinations considered previously. We note that the variation of the total mass within the survey volume has a minimal effect, and aside from the change in the overall scale, the contours are almost identical to those for the collapsed fraction.

2.3.3 Theoretical Estimates of Halo Assembly Time

In CDM cosmologies, dark matter haloes grow through repeated, stochastic mergers, gradually assembling their mass from a large number of smaller progenitors. Thus, deciding when a given halo has 'formed' is rather arbitrary. Most definitions in the literature are based on the Mass Accretion History (MAH) (van den Bosch, 2002). This is calculated by tracing the growth of the halo backwards in time and selecting the largest progenitor of each merger, to produce a single monotonic growth sequence $M(z)$; the MAH is then defined as the relative value $M(z)/M(0)$. Given a MAH, the formation epoch is often defined as the redshift by which a halo has reached some fixed fraction f of its final mass (e.g. z_{50} for $f = 0.5$ – Lacey & Cole 1993).

As for the HMF, the EPS formalism provides an analytic framework for exploring the cosmological dependence of halo age. We will first consider the predicted z_{50} distribution derived by Lacey & Cole (1993) assuming spherical collapse, and then give two different models of the ellipsoidal collapse equivalent, derived by Sheth & Tormen (2002) and Zhang et al. (2008) respectively.

Given a halo of mass M_0 at redshift z_0, the fraction of its mass that was in progenitor

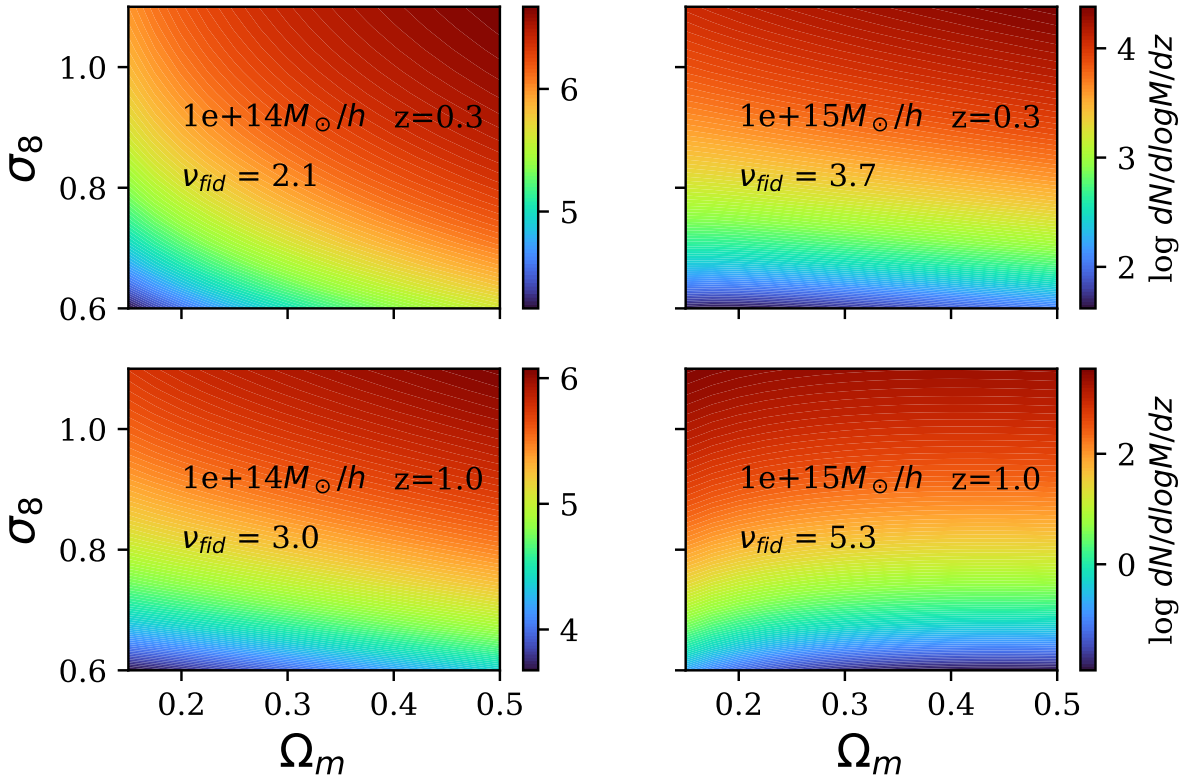

Figure 2.2: The cosmological dependence of the differential cluster number counts, per unit log mass and per unit redshift. The colour scale is logarithmic. The value of the peak height at different masses and redshifts for our fiducial cosmology ($\Omega_m = 0.3$, $\sigma_8 = 0.8$) is shown to illustrate the dependence of the contour shapes on peak height.

haloes of mass $M_1 \pm dM_1/2$ at redshift z_1, is given by the conditional probability

$$f_{SC}(M_1, z_1|M_0, z_0)dM_1 = \frac{1}{\sqrt{2\pi}} \frac{\delta_c(z_1) - \delta_c(z_0)}{(S(M_1) - S(M_0))^{3/2}} \times \exp\left(-\frac{(\delta_c(z_1) - \delta_c(z_0))^2}{2(S(M_1) - S(M_0))}\right) dS_1 \,, \tag{2.7}$$

where $S(M) = \sigma^2(M)$, and the other variables are as in Section 2.3.1

Multiplying by the factor M_0/M_1, we get the progenitor mass function (PMF), that is the number of progenitors of mass M_1

$$PMF(M_1, z_1|M_0, z_o)dM_1 = \frac{M_0}{M_1} f(M_1, z|M_0, z_0)dM_1 \,. \tag{2.8}$$

Following Lacey & Cole (1993), if we integrate the PMF from masses $M_0/2$ to M_0, we are calculating the average number of progenitors at redshift z_1 that have more than half the final halo mass at z_0. Since the halo cannot have more than one progenitor with more than half of its final mass, this quantity is also the probability that the halo had built up at least half of its mass in a single progenitor by redshift z_1. Thus, it gives the cumulative distribution of the formation redshift z_{50}:

$$P(z_{50} > z|M_0, z_0) \equiv \int_{M_0/2}^{M_0} \frac{M_0}{M} f(M, z|M_0, z_0)dM \,. \tag{2.9}$$

(We note that this approach only works for formation redshifts z_f with $f \geq 0.5$; there is no simple analytic way to obtain the distribution of z_f for $f < 0.5$.)

As with the HMF, this estimate of halo formation redshift is limited by the assumption of spherical collapse. Sheth & Tormen (2002) provided a version of the conditional mass function using a Taylor expansion of their moving barrier from (Sheth & Tormen, 1999) that can be used to calculate z_{50} (e.g. Giocoli et al., 2007). Their conditional probability is

$$f_{c,ST}(M_1, z_1|M_0, z_0) = \frac{|T(M_1, z_1|M_0, z_0)|}{\sqrt{2\pi \left(S(M1) - S(M_0)\right)^3}} \times \exp\left(-\frac{(B(M_1, z_1) - B(M_0, z_0))^2}{2(S(M_1) - S(M_0)}\right) , \tag{2.10}$$

where B is the moving barrier

$$B(M,z) = \sqrt{a}\delta_c(z)\left[1+\beta\left(\frac{S(M)}{a(\delta_c(z))^2}\right)^{\gamma}\right], \tag{2.11}$$

with parameters $a = 0.7$, $\beta = 0.485$ and $\gamma = 0.615$, while T is the first terms of a Taylor expansion of the function B

$$T(M_1,z_1|M_0,z_0) = \sum_0^5 \frac{(S_0-S_1)^n}{n!}\frac{\partial^n\left[B(M_1,z_1)-B(M_0,z_0)\right]}{\partial S^n(M_1)}. \tag{2.12}$$

Inspired by the ellipsoidal collapse model, Zhang et al. (2008) also developed a fitting function for the conditional probability based on ellipsoidal collapse:

$$\begin{aligned} f_{EC}(M_1,z_1|M_0,z_0)dM_1 &= \frac{A_0}{\sqrt{2\pi}}\frac{\delta_c(z_1)-\delta_c(z_0)}{(S(M_1)-S(M_0))^{3/2}}\exp\left(-\frac{A_1^2}{2}\tilde{S}\right) \\ \times\left\{\exp\left(-A_3\frac{(\delta_c(z_1)-\delta_c(z_0))^2}{2(S(M_1)-S(M_0))}\right)+A_2\tilde{S}^{3/2}\left(1+2A\sqrt{\frac{\tilde{S}}{\pi}}\right)\right\}dM_1\,, \end{aligned} \tag{2.13}$$

where $A_0 = 0.8661(1-0.133\nu_0^{-0.615})$, $A_1 = 0.308\nu_0^{-0.115}$, $A_2 = 0.0373\nu_0^{-0.115}$, $A_3 = A_0^2 + 2A_0A_1\sqrt{\Delta S\tilde{S}}/\Delta\omega$, $\nu_0 = \omega_0^2/S(M_0)$, $\tilde{S} = \Delta S/S(M_0)$, $\Delta S = S(M_1)-S(M_0)$ and $\omega_i \equiv \delta_c(z_i)$.

The probability distributions obtained using the three models (Eqs. 2.7, 2.10 and 2.13) in Eq. 2.9 are shown in Fig. 2.3. The ellipsoidal collapse models predict earlier formation times z_{50} at all masses, although the difference is largest at low mass. The figure also illustrates a well-known feature of hierarchical structure formation, that massive haloes have formed more recently. The predictions of the two ellipsoidal collapse models are very similar, so we will use the model from Zhang et al. (2008) as our base model in what follows, as it is slightly faster to calculate.

2.3.4 Cosmological Dependence of Halo Assembly Time

Given a prediction for the distribution of halo formation redshifts, we can study how it varies with cosmological parameters. We have tested the dependence of three quantities in particular:

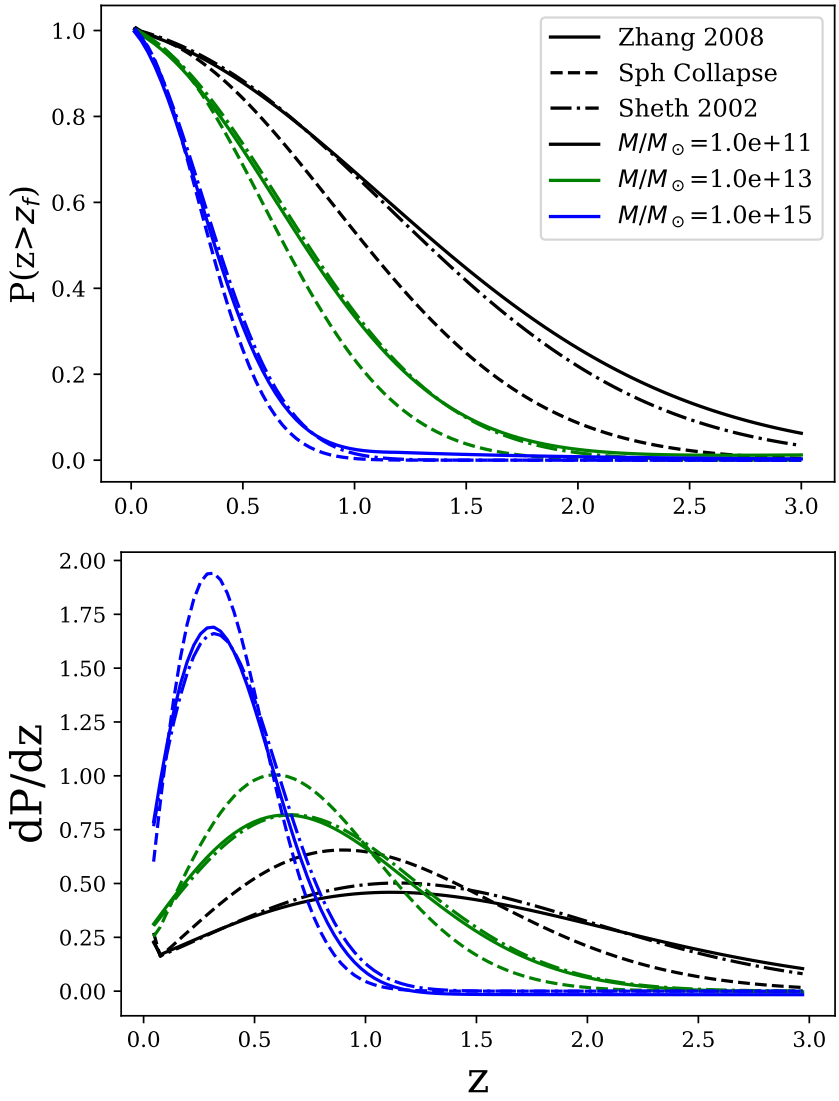

Figure 2.3: Cumulative (top panel) and differential (bottom panel) distributions of the formation redshift z_{50} in our fiducial cosmology, for the SC model (Lacey & Cole, 1993, dashed lines), and the EC models from Sheth & Tormen (2002, dot-dashed lines) and Zhang et al. (2008, solid lines).

- The median formation redshift, defined as $\langle z_{50} \rangle = z_m$ such that $P(Z > z_m) = 0.5$
- The peak of the differential probability distribution, $z_p = \max\left[\frac{\mathrm{dP}}{\mathrm{dz}}(\mathrm{z})\right]$
- The average formation redshift, $z_a = \int_{z_{obs}}^{\infty} z \frac{dP}{dz} dz$.

Of these three, we will focus on the median formation redshift, noting that the average formation redshift is slightly higher.

Both spherical and ellipsoidal collapse models predict the same behaviour of the median formation time, as shown in Fig. 2.4. Haloes form earlier in high–σ_8 cosmologies, since a greater mean fluctuation amplitude causes typical peaks in the density field to cross the threshold for collapse earlier in the process of structure formation. The Ω_{m}-dependence is less trivial and differs between low-mass and high-mass haloes. As explained in Appendix A.3, high-Ω_{m} cosmologies have more power on small scales relative to large ones. Thus at fixed σ_8, low-mass haloes form earlier in higher Ω_{m} universes, while high mass haloes form slightly later. This agrees with the previous findings of Giocoli et al. (2012).

The general dependence of formation epoch z_{50} on Ω_{m} and σ_8 is shown in Fig. 2.5. The main trend is for age to increase with σ_8; since the masses shown here are all beyond the cross-over point in Fig. 2.4, median age also decreases slightly with Ω_{m}, particularly at large masses. We explore how the dependence of z_{50} on Ω_{m} and σ_8 arises in more detail in Appendix B.

Comparing Figs. 2.2 and 2.5 closely, we note an important feature of halo age relative to halo abundance: for lower mass haloes and/or at lower redshift, the contours for the two are fairly orthogonal over much of the Ω_{m}–σ_8 plane. To highlight this point, Fig. 2.6 shows the two sets of contours superimposed, for the ranges of mass and redshift accessible to large cluster surveys. In the region of particular interest, around the concordance model ($\Omega_{\mathrm{m}} = 0.3$, $\sigma_8 = 0.8$), the two sets of contours are almost exactly orthogonal for lower masses and/or redshifts, where $\nu \sim$2–3 (top and middle left hand panels). They only become similar for the most massive clusters, at $z \geq 1$, where $\nu \sim$ 4–6 (bottom right panel). This implies that age or age proxies, measured for clusters with masses $M < 5 \times 10^{14} M_{\odot}$ at $z < 1$, can potentially break the main degeneracy in cluster abundance measurements.

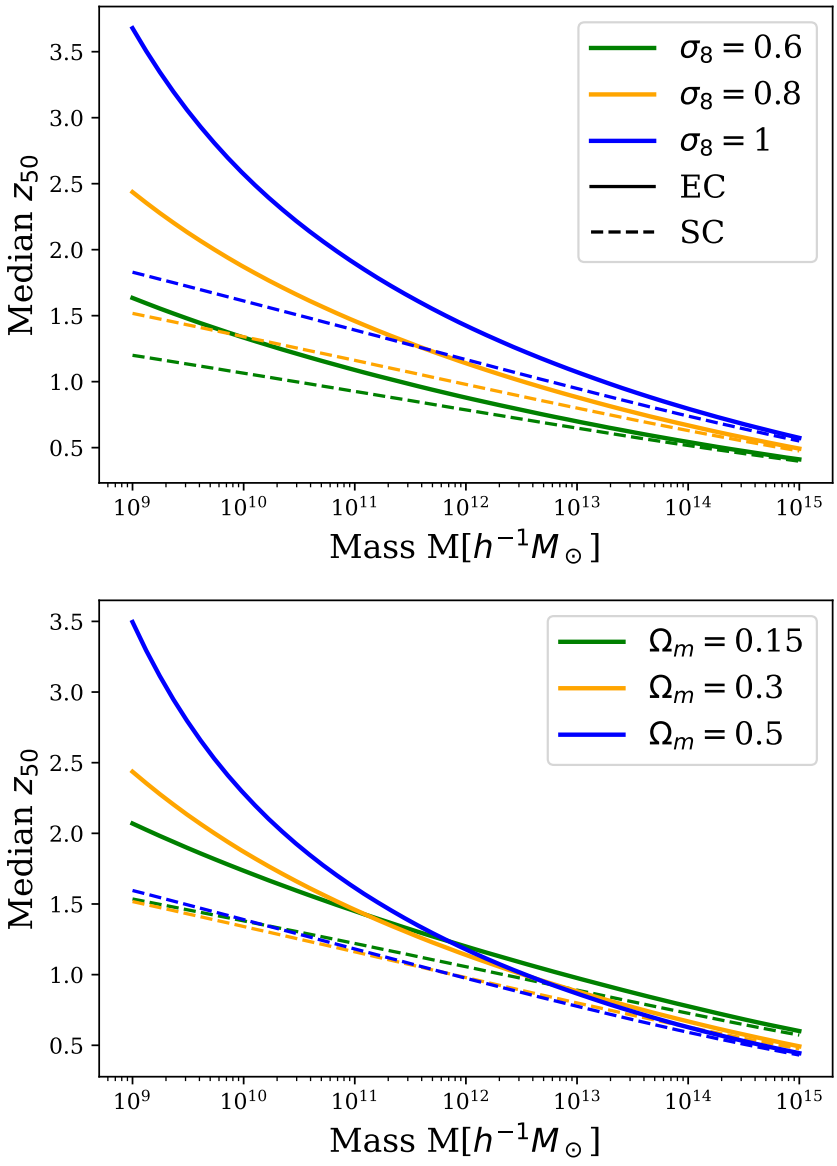

Figure 2.4: Mass dependence of the median formation redshift $\langle z_{50} \rangle$, for haloes at $z = 0.15$, and for different values of σ_8 (left panel) and Ω_m (right panel). Solid lines show the ellipsoidal collapse predictions; dashed lines show the spherical collapse predictions.

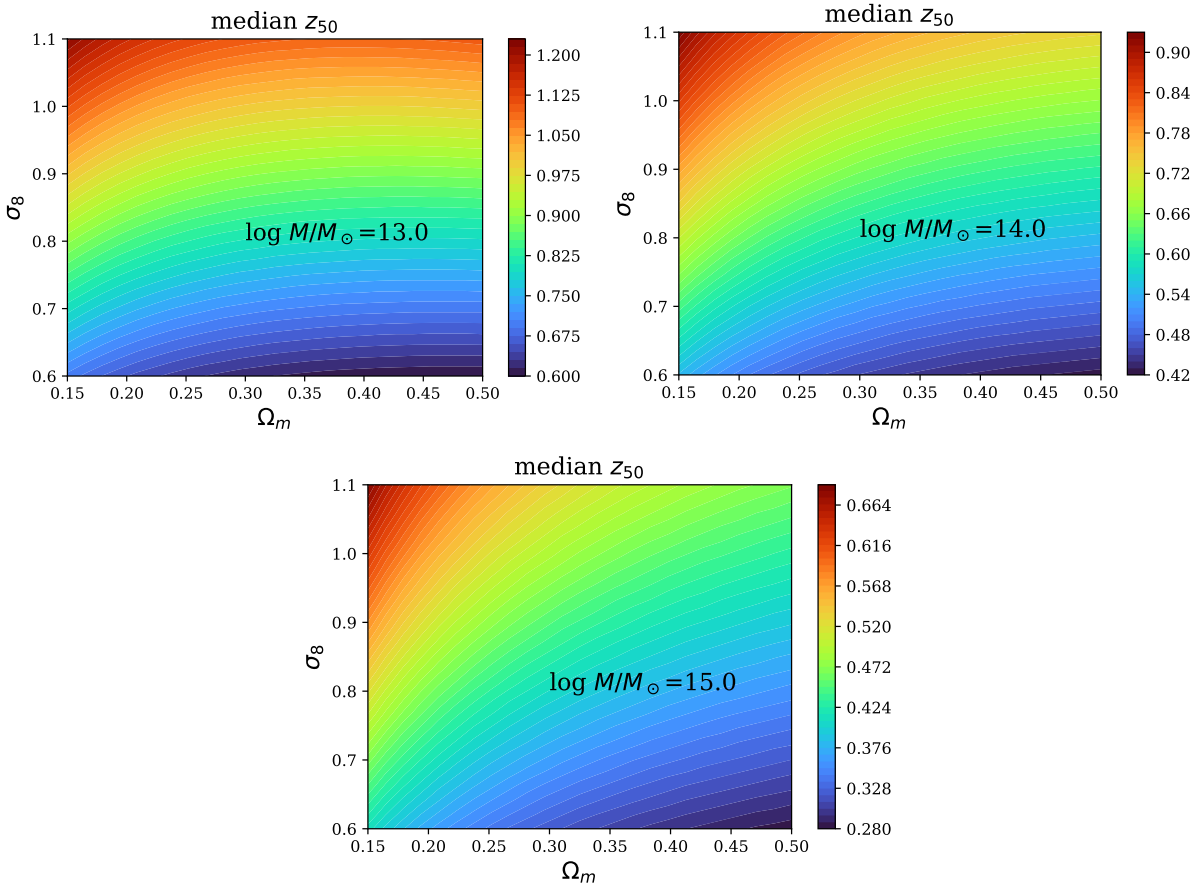

Figure 2.5: Median z_{50} for present-day haloes of the mass indicated, as a function of $\Omega_{\rm m}$ and σ_8. Calculations assume the EC model (cf. Sec.2.3.3).

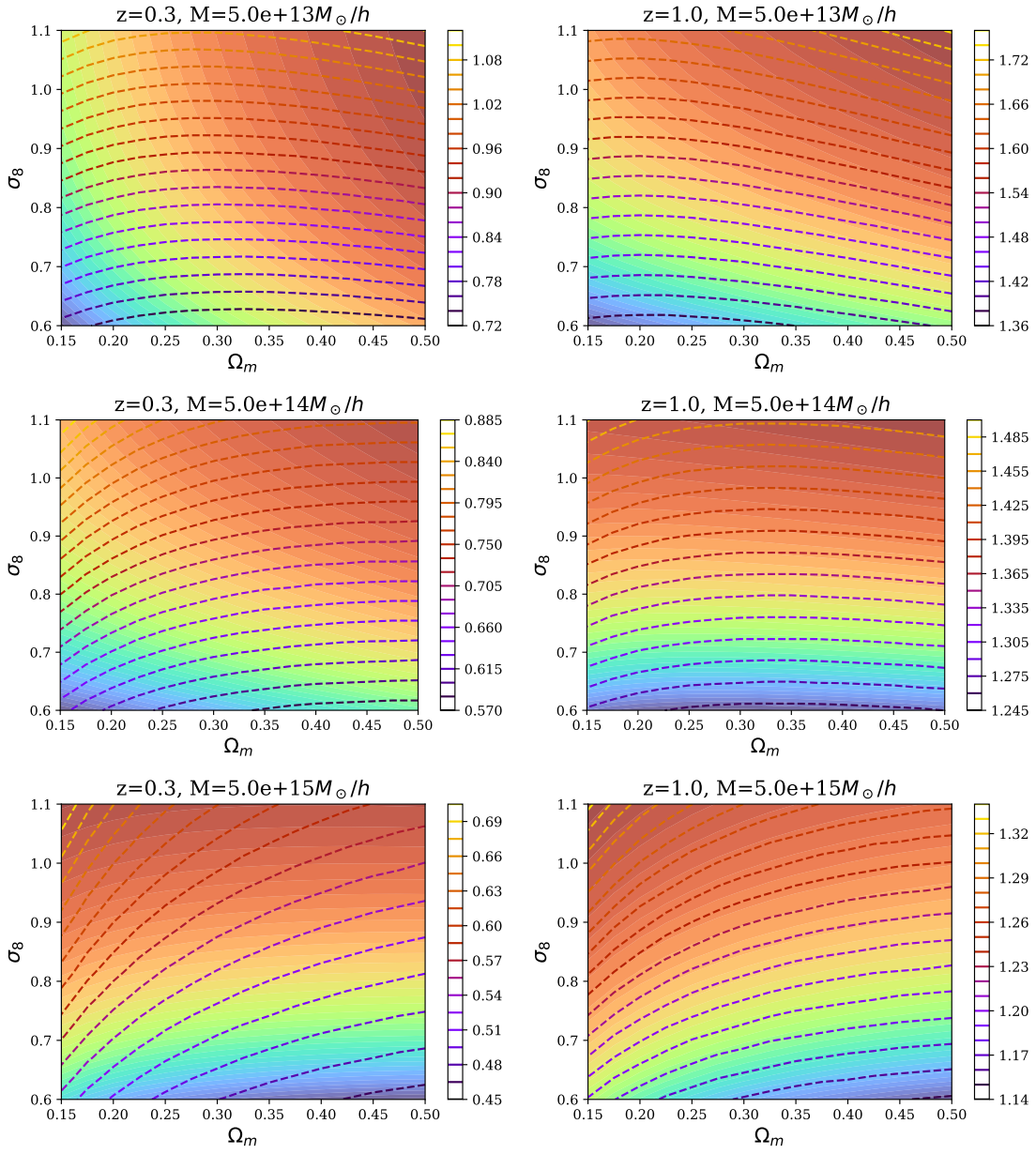

Figure 2.6: A comparison of age and abundance contours in the $\Omega_{\rm m}$–σ_8 plane, for the halo mass and redshift indicated at the top of each panel. Filled contours show curves of constant number density per unit $\log M$. Dashed lines show curves of constant z_{50}, with the scale given by the colorbar.

2.4 Comparison to simulations

2.4.1 Simulation Data

In Section 2.3.3, we provided analytic EPS estimates of how the halo formation time z_{50} depends on $\Omega_{\rm m}$ and σ_8. Given the approximations made in these models, it is worth testing the accuracy of their predictions in N-body simulations.

Previous work has demonstrated that different merger tree algorithms can produce significantly different MAHs (Avila et al., 2014; Srisawat et al., 2013). The exact value of z_{50} may be particularly sensitive to these differences, as discussed in Srisawat et al. (2013). In particular, some merger tree algorithms allow fragmentation events, where haloes lose mass with time, such that MAHs are not always monotonic. Our previous EPS estimates assume strictly hierarchical growth, and thus we anticipate that the numerical results may disagree with them to some degree. To test the effect of different methods of analysis, we consider three sets of simulations (two public, and one of our own), that employed three different merger tree algorithms:

- The Illustris TNG simulation, (Nelson et al., 2019) using the Sublink merger tree algorithm (Rodriguez-Gomez et al., 2015).
- The Bolshoi/BolshoiP simulation, (Klypin et al., 2011) using the Rockstar halo finder (Behroozi et al., 2013a) and the Consistent Trees merger tree code (Behroozi et al., 2013b). (Note that Bolshoi uses WMAP cosmological parameters, whereas BolshoiP uses Planck ones.)
- Our own set of 9 cosmological simulations, spanning a range of cosmological parameters, and analyzed using the AHF halo finder and merger tree code (Knollmann & Knebe, 2009; Gill et al., 2004). These will be labelled MxSy, where $x = 25/3/35$ indicates the value of $\Omega_{\rm m}$, and $y = 7/8/9$ indicates the value of σ_8.

The simulation parameters are summarized in Table 2.1. Data from the TNG and Bolshoi simulations were obtained directly from their respective websites. In particular, we used the Rockstar merger tree data available for the Bolshoi simulation. For the TNG and MxSy simulations, mass accretion histories were calculated using the Sublink and AHF codes, respectively. For the Bolshoi simulations, they were generated by following the main progenitor sequence in the Rockstar files.

Simulation	$\Omega_{\rm m}$	σ_8	part mass $[M_\odot/h]$	$N_{\rm part}$	merger tree	$N_{\rm snap}$
TNG	0.31	0.81	3×10^9	625^3	Sublink	100
Bolshoi	0.27	0.82	1.35×10^8	2048^3	Consistent Trees	181
BolshoiP	0.31	0.82	1.55×10^8	2048^3	Consistent Trees	178
MxSy	0.25/0.3/0.35	0.7/0.8/0.9	4×10^9	512^3	AHF	44

Table 2.1: Summary of the simulations used and their main parameters, including cosmological parameters, particle mass, total number of particles $N_{\rm part}$, merger tree code, and the number of snapshots $N_{\rm snap}$ used to make the merger trees. The MxSy simulations are a set of 9 of our own simulations that span a range of different values of $\Omega_{\rm m}$ and σ_8.

2.4.2 Formation Time Distributions Compared

In each numerical MAH, we define z_{50} to be the lowest redshift at which the mass of the halo has dropped to less than half of the mass at $z = 0$. Fig. 2.7 compares the distribution of these formation redshifts to the analytic (EC) predictions. For all three simulations considered, but particularly for the Bolshoi and M3S8 simulations, we see a clear offset between the numerical results and the EC predictions, that is largest at small masses. In general, the numerical formation redshifts are larger than predicted, by up to 0.1–0.2 on average.

One possible explanation for this shift lies in the different definitions of merger time assumed. Given a particular merger event, EPS theory takes the corresponding collapse redshift (that is, roughly, the time by which newly-accreted mass has first fallen to the centre of the halo) to be the moment at which a halo's virial mass is said to increase. In contrast, numerical group finders may link haloes when their outer virial surfaces first touch. Thus, numerical mergers may occur up to one infall time earlier than analytic ones. Adding a delay equal to the infall time to the numerical results, we obtain the z_{50} distributions in Fig. 2.8. The discrepancy between the numerical and analytic results is greatly reduced, although some differences remain, as seen most clearly in the cumulative distributions.

The remaining differences may have several possible explanations. There are slight offsets between the distributions for the three simulations, suggesting that the different halo finders and merger tree algorithms used to analyze them affect the results. A detailed comparison of halo finders and merger tree algorithms, including AHF, Rockstar/ConsistentTrees and Subfind/Sublink, was presented in Knebe et al. (2011). They highlight a number of significant differences between methods, notably in how they treat fragmentation events and non-monotonic MAHs. We note that the discrepancy between

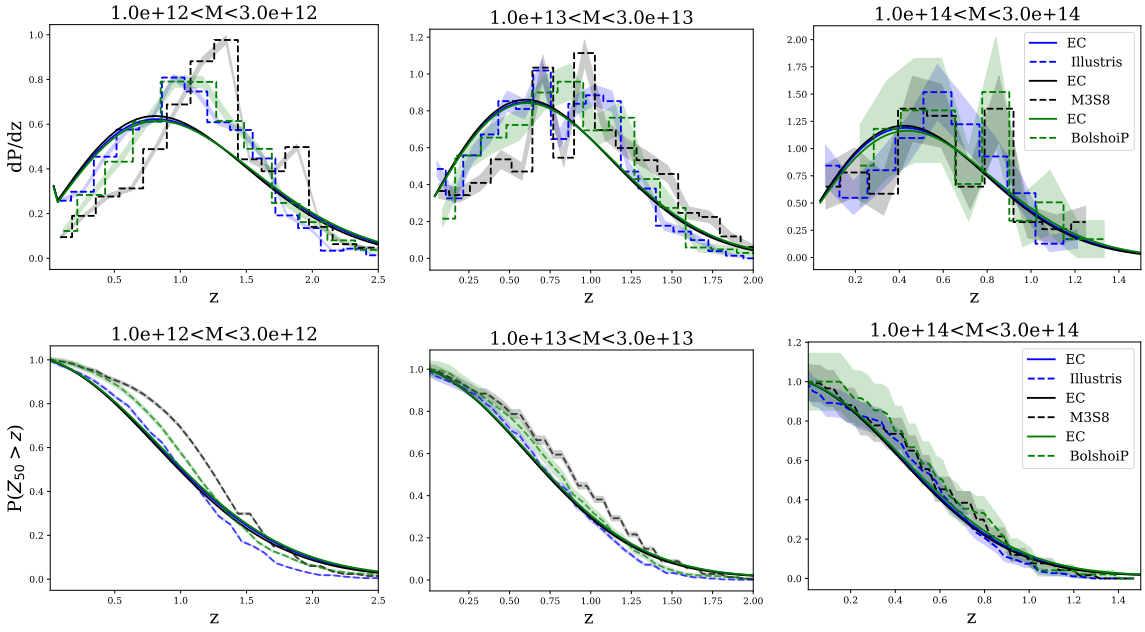

Figure 2.7: Differential (top) and cumulative (bottom) distributions of z_{50} in our fiducial cosmology, for three different ranges of halo mass. Smooth curves show the (EC) analytic prediction, while dashed lines with shading show the numerical results and associated Poisson uncertainties.

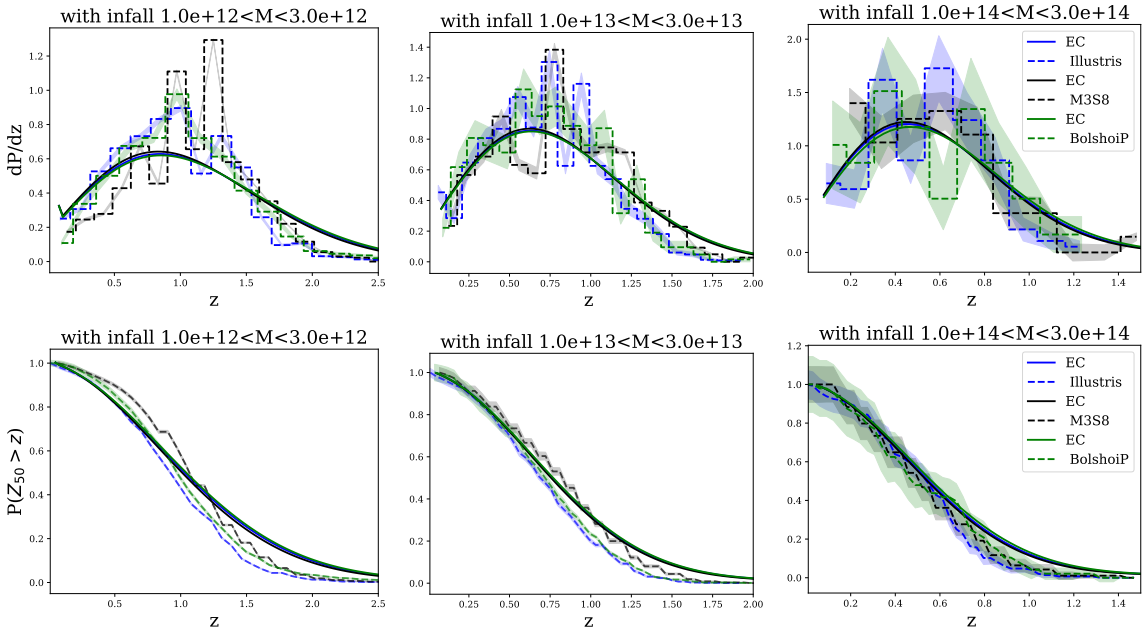

Figure 2.8: Differential and cumulative z_{50} distributions, as in Fig. 2.7, after taking into account the delay due to infall.

numerical and analytic results is greatest at low mass, so resolution may also play a role. Finally, even the revised EC version of EPS remains an approximate theory, so its predictions may be inaccurate at some level.

We summarize the comparison between numerical and analytic results in Fig. 2.9, which shows the median z_{50} of each of the simulations, together with the uncertainty (points with errorbars), compared to the analytical predictions from the EC model of Zhang et al. (2008). As discussed in Section 2.3.4, age is most sensitive to the amplitude of fluctuations σ_8, and depends only weakly on $\Omega_{\rm m}$. At high mass, the numerical results agree well with the analytic prediction, at least in terms of the median value of z_{50}. Since this mass range is the one relevant for cluster surveys, we conclude that our previous analytic estimates are reasonably valid, although the details of the halo age distribution require further study in future work.

2.5 Observational prospects

A number of ongoing and future surveys are expected to produce very large samples of galaxy clusters, with $O(10^5)$ significant detections, out to redshifts of $z = 1$ or higher (e.g. Pillepich et al., 2012; Sartoris et al., 2016; Abazajian et al., 2019b). Supposing such a sample were available, with age information based on one or more observational proxies, we can ask what sensitivity this dataset would have to the cosmological parameters, or equivalently how large a sample would be needed to provide significant improvement on parameter constraints.

To estimate age observationally, we need a structural proxy for age (as expressed, say, by the formation epoch z_{50}). There are several known examples of structural properties that correlate with z_{50} (Wong & Taylor, 2012), including concentration (Zhao et al., 2003; Wang et al., 2020), shape (as a product of major mergers -Drakos et al. 2019a), substructure (e.g. Gao et al., 2004; Taylor & Babul, 2005; Diemand et al., 2007), or overall degree of relaxation, as measured by a centre-of-mass offset (Macciò et al., 2007; Power et al., 2012). We will take concentration as an example here, as its age dependence is the best studied. We note that on some mass scales and at some redshifts, baryons may have an important effect on halo structure, and on concentration specifically. We will start by discussing concentration measurements ignoring these possible effects, but then consider them separately in Section 2.5.4 below.

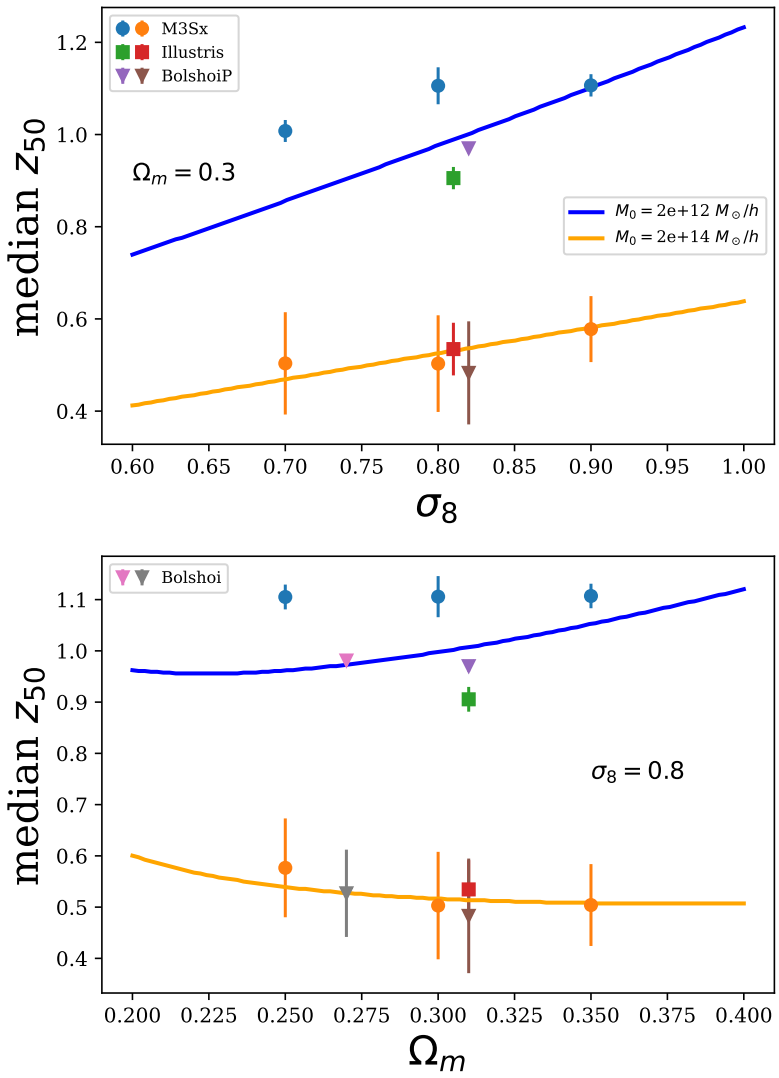

Figure 2.9: Median value of z_{50} for present-day haloes, as a function of σ_8 (top panel) and Ω_m (bottom panel). Smooth curves show the EC prediction; points with error bars show the numerical results and Poisson uncertainties, corrected for infall time. Each symbol represents a different set of simulations and different point colors represent different masses. Note that on the right panel two versions of Bolshoi are shown, one with Planck cosmology and one with WMAP cosmology. See Table 2.1.

2.5.1 Mean Concentration versus z_{50}

Since the discovery of the universal density profile (Navarro et al., 1997, NFW hereafter), the value of the concentration parameter $c = r_{\rm vir}/r_{\rm s}$ has been linked to the halo's formation history. In NFW, concentration depends on how early a critical fraction of the final mass was first assembled into (any number of) progenitors. Subsequent models (e.g. Bullock et al., 2001; Wechsler et al., 2002; Zhao et al., 2003, 2009; Ludlow et al., 2014; Correa et al., 2015a) related concentration instead to the growth history of a single main progenitor, as expressed by the mass accretion history (MAH). In the simplest picture (Wechsler et al., 2002; Zhao et al., 2003), $c \simeq c_0(a_0/a_r)$ where a_r is the scale factor at the end of the period of rapid growth in the MAH, and $c_0 \sim 3$–4 is the concentration of newly-formed systems at this time. In these models, there is, therefore, a direct correlation between c and z_{50}, or any similar estimate of the formation epoch $z_{\rm f}$ (Wong & Taylor, 2012).

All of these models focus on the relation between the average growth rate and the mean concentration of a sample of haloes of a given range of mass and redshift (although Ludlow et al. 2013 does consider the connection between individual MAHs and concentration values.) Major mergers can lead to large variations in concentration, however, depending on the net input of (orbital) energy (Drakos et al., 2019b). Most recently, Wang et al. (2020) have shown that the measured value of the concentration parameter oscillates during major mergers, and that these fluctuations may dominate the statistics of the average values measured for large ensembles. Clearly, the subject is complicated and requires further study; we will not consider it in further detail here, but will assume a correlation between c and z_{50}, that makes mean concentration measurements sensitive to mean age.

Fig. 2.10 shows this correlation in practice, as measured from our grid of MxSy simulations for different cosmological parameters (we have chosen a lower halo mass range, 1–$3{\times}10^{12}M_\odot$, to reduce Poisson noise in the figure). The basic pattern is similar for each set of cosmological parameters, and has been explored extensively in the literature (e.g. Wechsler et al., 2002; Zhao et al., 2003, 2009; Giocoli et al., 2012; Correa et al., 2015b; Ludlow et al., 2013), though interestingly, there is also a slight change in the mean relation over the range of parameters explored. In particular, the intercept of the linear regression relation between $\log c$ and $\log(1+z)$ increases monotonically, both with Ω_m, and with σ_8. This indicates that the concentration has an additional cosmological sensitivity, beyond its main dependence on formation history. Here too, there is clearly further complexity to explore in the concentration-mass-redshift relation; in future work, we will focus on understanding and calibrating the mean $c(z, M)$ and c–z_{50} relationships, and consider more generally the links between concentration, mass accretion history, and cosmology. For the purpose of our present calculations, we will assume a power-law correlation with a fiducial

scatter of 30%, which provides a reasonable fit to the results from all nine cosmologies.

2.5.2 Measured Concentration versus Mean Concentration

From the preceding results, the actual 3D concentration of an individual halo should scatter by $\sim 30\%$ for a given value of z_{50}. This actual concentration can be estimated in various ways, including weak lensing convergence, detailed modelling of the lensing potential in strong lensing systems, X-ray or SZ emission, or even the galaxy distribution within a group or cluster. Each of these techniques will add observational errors and biases; see, for instance, Groener et al. (2016), which provides a fairly recent review of individual concentration estimates in clusters. Generally, the observational errors are 0.1 or 0.2 dex, i.e. 25–60%, for each individual system. There are also systematic uncertainties, both identified and unidentified, associated with each method. In principle, future work with large samples, dedicated simulations, and comparison between observational modalities may help reduce these. Overall, we will assume typical errors of either 30% or 50% in going from an actual 3D concentration to an observational estimate.

Combining these errors with the intrinsic scatter in the c–z_{50} relation, we expect a net scatter of $\epsilon_{z_{50}} = 40$–60% in the relation between an observational estimate of concentration and the formation epoch z_{50}. This large uncertainty makes individual measurements relatively uninteresting; we can compare the situation to weak gravitational lensing, however, where shape measurements for individual galaxies are extremely noisy, but careful averaging extracts the mean value in an unbiased way.

2.5.3 z_{50} versus Cosmology

The remaining factor in our calculation is the connection between an estimate of the mean value of z_{50} and the values of the cosmological parameters. From Fig. 2.9, to achieve a nominal precision of 0.01 in σ_8, we need 0.55% precision in the estimate of $\langle z_{50} \rangle$. Assuming unbiased averaging over a sample of N clusters, $\sigma\langle z_{50} \rangle = \epsilon_z/\sqrt{N}$. Solving, we get $N = (0.40/0.0055)^2$–$(0.60/0.0055)^2 \sim$ 5,000–12,000. Thus, with low-precision but unbiased concentration measurements for $O(10{,}000)$ clusters, we could obtain constraints on the value of σ_8 that correspond to 1/10 or less of the current range of uncertainty in this parameter.

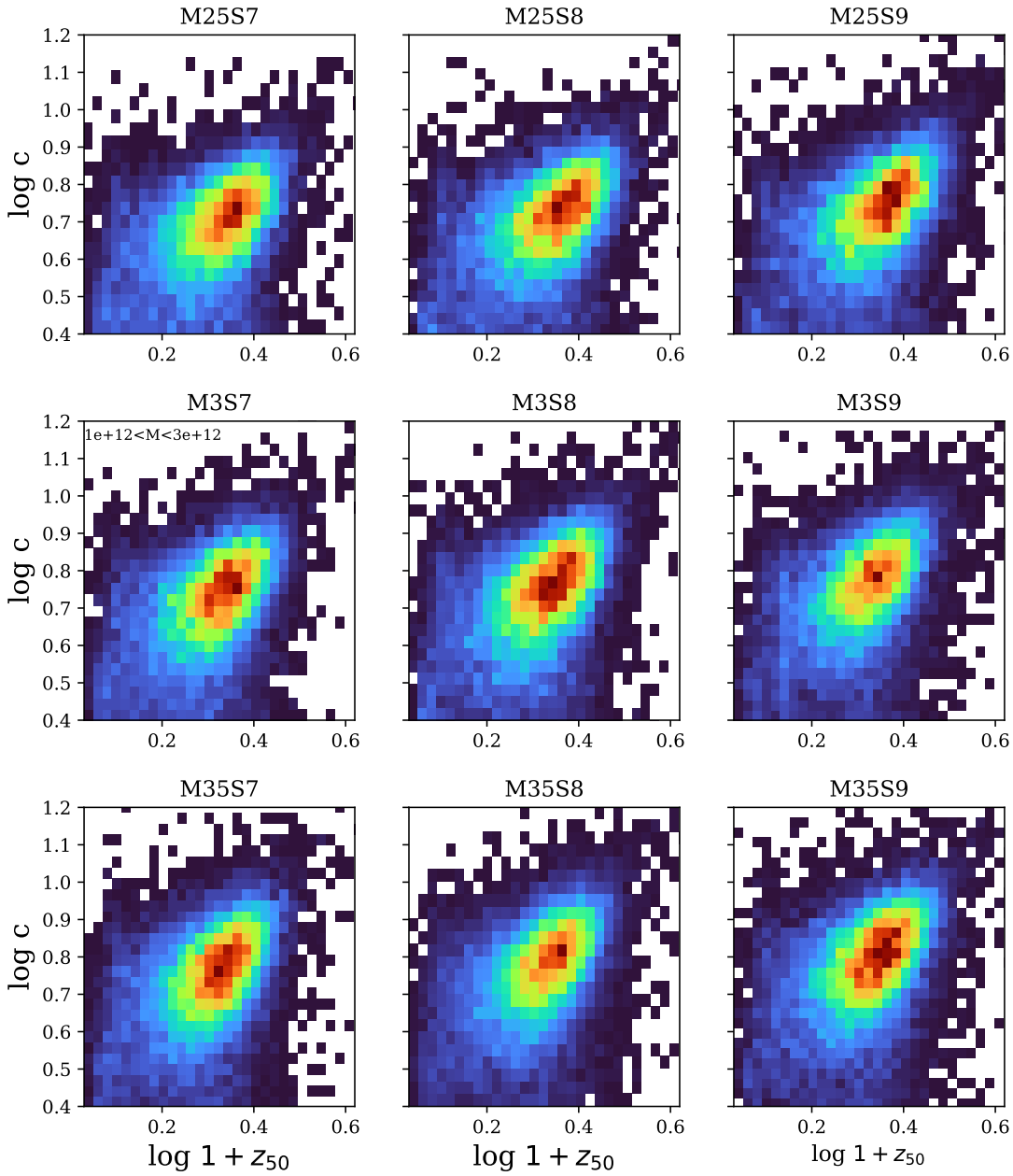

Figure 2.10: Concentration versus z_{50} for present-day haloes with masses between $10^{12} M_{\odot}$ and $3 \times 10^{12} M_{\odot}$, in the 9 MxSy simulations.

2.5.4 Baryonic Corrections

A major uncertainty in the preceding calculations is the net effect of baryons on cluster concentrations. Baryons may alter the halo density profile, increasing halo concentration via adiabatic contraction, or reducing it through outflows driven by stellar or AGN feedback. These effects are complex, and depend on mass, redshift, and radius within the halo. Overall, these processes can affect halo concentration and masses significantly (Debackere et al., 2021). Although simulations suggest that baryonic effects are largest in galaxy-scale haloes (e.g. Velliscig et al., 2014), they may still be significant when measuring concentration or internal structure in clusters – Debackere et al. (2021), for instance, find a 10% bias in estimates of the scale length r_s at masses of $5{\times}10^{14}h^{-1}M_{\odot}$.

Baryonic effects are not yet well enough understood to include in our predictions; in particular, their detailed dependence on the cosmological parameters is not yet known. We can point out a few possible avenues, however, to calibrating and correcting for their effects on structural measurements. First, simulations now model these effects with increasing accuracy, allowing the potential for calibration of any net bias in structural properties. Second, observations of nearby, well-studied systems allow verification of the simulations, independent of the samples used for cosmological tests. Third, as pointed out in Section 2.3.4, over some mass and redshift ranges, cluster age and abundance are predicted to vary almost identically with the cosmological parameters $\Omega_{\rm m}$ and σ_8. This should provide an independent test of any assumed concentration-mass relationships, at least for this range of mass and redshift, as the constraint in the $\Omega_{\rm m}$–σ_8 plane derived from age/concentration measurements must agree with the one derived from abundance, and since the contours for the two are parallel, there is little room for bias in one relative to the other.

A final, and basic, reason for optimism is the differential nature of structural tests, whether based on concentration or on other structural properties. These would depend on the relative distribution of structural properties, measured across a population of systems. A simple proxy for the mean age of the sample, for instance, might be the number of high-concentration systems, relative to low-concentration ones. Thus, to lowest order, a net shift in concentration for the whole population would largely cancel out, reducing the bias in the final results. At the same time, the *shape* of the measured concentration distribution for the whole sample would provide another test of the consistency of the method, and any biases or effects due to sample selection.

Overall, it is clear that the impact of baryonic effects on halo structure requires much more detailed study, to see whether and to what degree they would compromise structural tests of cosmology, for a given survey and methodology.

2.5.5 What is Achievable?

In summary, in previous sections, we have shown that with low-precision but unbiased concentration estimates for $O(10^4)$ clusters, one could obtain excellent constraints on σ_8, assuming the net effect of baryons is small, or can be corrected using simulations. While measuring concentration observationally is challenging, our method does not require particularly accurate measurements for individual systems – given the intrinsic scatter in the c–z_{50} relation, the errors in individual estimates need only be accurate at the $\sim$30% level.

We can consider this goal in more detail for a particular survey. The *Euclid* mission is a 1.2m space telescope, operating in the visible and near-infrared (NIR). As part of its wide survey, it will image 15,000 deg^2 of the sky in one optical and three NIR bands, detecting galaxies down to an AB magnitude of 24 or fainter. Photo-zs will be derived for these objects, using ancillary data from ground-based surveys such as UNIONS (Chambers et al., 2020). The *Euclid* wide survey should provide large catalogues of clusters detected photometrically (that is by clustering in projection and in photo-z space), and also as peaks in weak lensing maps. Based on the forecasts of (Sartoris et al., 2016), the photometric detections should include all clusters with masses $M \gtrsim 10^{14} M_\odot$ out to redshift $z \sim 2$. Below redshift z = 0.5, the wide survey is expected to detect 1.5 million clusters at 3σ or greater significance, and 200,000 clusters at 5σ significance. Extrapolating from these predictions, the number of 7σ detections is in excess of 20,000 (with half that number at redshifts $z < 0.7$). Admittedly, these objects may be slightly more massive than the example considered above in section 2.5.3 ($3.5 \times 10^{14} M_\odot/h$, versus $2 \times 10^{14} M_\odot/h$) but they are close enough that the slope of the $\langle z_{50} \rangle$–σ_8 relation should be similar.

The Roman Space Telescope mission is a 2.4m wide-field space telescope, with optical/NIR imaging and slitless spectroscopy capabilities. Its High Latitude Survey will image $\sim$2000 deg^2 of sky in four NIR bands, reaching depths 1–2 AB magnitudes deeper than Euclid Wide, as well as providing slitless spectroscopy of brighter targets. This deeper data over a smaller area should produce a cluster sample that extends to lower masses and higher redshifts, and thus provides an interesting counterpoint to Euclid data. In particular, we note that at high redshift, the complementarity of age and abundance is reduced (cf. the left-hand panels of Fig. 2.6). This could provide an important consistency check on age estimates, as discussed previously. Finally, a number of other forthcoming experiments expect to detect large numbers of clusters, including *eROSITA* (Pillepich et al., 2012) in the X-ray, CMB-S4 (Abazajian et al., 2019b) in the mm, and the ground-based UNIONS (Chambers et al., 2020), DESI (DESI Collaboration et al., 2016), and Rubin LSST (LSST Science Collaboration et al., 2009) surveys.

Overall, we conclude that multiple samples of $O(10^4)$ clusters with sufficient signal-

to-noise ratio (SNR) to allow structural measurements should become available in the near future. One could imagine using a large, uniform survey such as *Euclid* for the low-redshift sample selection, together with other complementary observations to make structural measurements on individual clusters. A high-redshift sample could then be used to test and calibrate age proxies, as mentioned above. It might also be possible to stack clusters to use a mean projected density profile, measured at high SNR, to derive constraints. We will consider these and other approaches in future work.

As in weak lensing studies, the challenge of averaging over large numbers of low SNR measurements will be in controlling for and reducing systematics. Beyond the baryonic effects discussed in the previous section, systematics related to basic structure formation could also affect sample selection (e.g. by preferentially highlighting or neglecting disturbed systems); they could bias individual mass estimates (although the slope of the concentration-mass relation is fairly shallow, so accurate masses are less important than in abundance studies); or they could bias concentration measurements, e.g. by biasing the sample selection towards objects with a particular 3D shape, or with a disturbed IGM (if the confirmation or structural measurements are performed in the X-ray). In lensing-based studies, false peaks and projections could be a particular problem, as these may look less regular and have lower concentrations, biasing the average. Environment can also have an impact on formation time through assembly bias, as simulations have shown that haloes form earlier in dense environments (Gao et al., 2005; Wechsler et al., 2006; Harker et al., 2006), so unbiased sampling of large volumes is important. Here again, there is much future work to be done considering the the possible biases for different observational modalities and survey strategies.

2.6 Conclusion

The enormous success of CMB analyses and large cosmological surveys over the last few decades has been driven, for the most part, by a robust and detailed understanding of structure formation in the linear regime. Cluster number counts are an important exception, but they only probe one limited aspect of non-linear structure formation. As cluster catalogues grow by several orders of magnitude in size over the next decade, it is worth considering what other cosmological information we might extract from them. Measurements of internal halo structure can, in principle, tell us about the rate of non-linear structure formation, and are worth considering as a next-generation cosmological test.

Previous work has established that as haloes grow through hierarchical merging, this process leaves structural signatures that can last for many dynamical times, that is, for

many Gyr at low redshift. As a result, structural measurements provide several different avenues to estimate cluster assembly times or "ages". In this chapter, we have shown that for typical clusters at $z < 1$, age varies almost orthogonally to abundance in the space of the cosmological parameters $\Omega_{\rm m}$ and σ_8. The same datasets that provide abundance constraints could be used to estimate mean values for structural parameters, and thus age, providing significantly tighter parameter constraints from a single set of observations.

Of course, given the accuracy of current constraints from the CMB, it may seem less interesting to invest further in other techniques. A survey of current results hints at tension between the different measurements, however, emphasizing the importance of redundant cosmological tests, over different ranges of redshift, mass and/or spatial scale. To resolve the deep mysteries of dark energy and dark matter, and to rule out yet-undiscovered variations on the current cosmological model, we need to test it as sensitively as possible, across as broad as possible a range of parameter space. In pursuit of this goal, our growing understanding of non-linear structure formation will open up many exciting possibilities for new tests and new tools.

Chapter 3

Halo Growth and Merger Rates as a Cosmological Test

3.1 Abstract

Dark matter haloes grow at a rate that depends on the value of the cosmological parameters σ_8 and $\Omega_{\rm m}$ through the initial power spectrum and the linear growth factor. While halo abundance is routinely used to constrain these parameters, through cluster abundance studies, the halo growth rate is not. In recent work, we proposed constraining the cosmological parameters using the overall dynamical "age" of clusters, expressed, for instance, by their half-mass assembly redshift z_{50}. Here we explore the prospects for using the instantaneous growth rate, as determined from the halo merger rate, the average growth rate over the last dynamical time, or from the fraction of systems with recent episodes of major growth. We show that the merger rate is mainly sensitive to the amplitude of fluctuations σ_8, while the rates of recent growth provide constraints in the $\Omega_{\rm m}$–σ_8 plane that are almost orthogonal to those provided by abundance studies. Data collected for forthcoming cluster abundance studies, or studies of the galaxy merger rate in current and future galaxy surveys, may thus provide additional constraints on the cosmological parameters complementary to those already derived from halo abundance.

3.2 Introduction

The standard Lambda Cold Dark Matter (Λ CDM) cosmological model provides an extremely effective framework for understanding and predicting cosmological observations. As the accuracy of observational constraints increases, however, several small discrepancies have begun to challenge the success of the ΛCDM model. In particular, in measurements of the amplitude of density perturbations, (commonly represented by the parameter σ_8, the rms of density fluctuations smoothed on scales of 8 Mpc/h), and the growth of perturbations (determined by the matter density parameter Ω_m), a tension has emerged between results based on the CMB power spectrum (Planck Collaboration et al., 2020; Aiola et al., 2020) and studies probing the late-time matter density field directly through weak gravitational lensing (e.g. Heymans et al., 2021a) or cluster abundance (e.g. Abdullah et al., 2020; Bocquet et al., 2019b).

This discrepancy, referred to as the "S_8 tension" in reference to the growth of structure parameter $S_8 \propto \sqrt{\Omega_m}\sigma_8$, has grown significantly in recent years as more precise weak-lensing studies have been released, in particular the Dark Energy Survey (DES Y3) (Abbott et al., 2022) and the Kilo Degree Survey (KiDS-1000) (Heymans et al., 2021b), and is about $\sim 3\sigma$ at the time of writing. Various solutions to the tension have been proposed, including systematic problems in the analysis (e.g. Sánchez, 2020), biased cluster mass estimates (e.g. Douspis et al., 2019; Debackere et al., 2021) or modifications to the standard cosmological model (e.g. Di Valentino et al., 2015; Böhringer & Chon, 2016; Planck Collaboration et al., 2016b; Heimersheim et al., 2020), but it is not clear that any of these fully resolve the problem. Given the persistent discrepancy, it is worth exploring other independent methods for estimating σ_8 and Ω_m, to investigate all possible origins for the tension.

In the standard cosmological model, dark matter becomes non-relativistic (cold) at early times, and structures grow hierarchically after the initial gravitational collapse of peaks in the density field (Jõeveer et al., 1978; Peebles, 1980; White & Frenk, 1991; Padmanabhan, 1993; Dodelson, 2003). Within this framework, the largest structures, galaxy clusters, are the last objects to assemble most of their mass, yet they also form around and thus probe the highest peaks in the initial density field. Consequently, present-day cluster abundance can be used to estimate the early abundance and subsequent growth of the density peaks, which in turn have a clear dependence on σ_8 and Ω_m (Press & Schechter, 1974). The cluster count method has been used extensively for several decades to constrain these parameters (Evrard, 1989; Henry & Arnaud, 1991; Lilje, 1992; Wang & Steinhardt, 1998; Abdullah et al., 2020).

This cosmological test uses remarkably little information about individual clusters, requiring only their observed redshift and mass proxy, as determined from observations in

the X-ray (e.g. Henry et al., 2009; Mantz et al., 2010; Böhringer et al., 2014), weak-lensing surveys (e.g. Kacprzak et al., 2016), optical galaxy surveys, or sub-mm imaging via the Sunyaev-Zel'dovich effect (e.g. de Haan et al., 2016; Planck Collaboration et al., 2020). Despite the wealth of information present in these observational data, structural features of clusters such as their concentration, substructure and shape have not been exploited for cosmological purposes. This is partly because of the difficulty of measuring, modelling and understanding these features, but also because the large data sets necessary to perform cosmological analyses using highly variable, complex properties such as cluster structure were previously unavailable. This situation is now changing rapidly, as forthcoming missions and surveys, including *Euclid*, DESI (DESI Collaboration et al., 2016), the *Vera C. Rubin Observatory* (LSST Science Collaboration et al., 2009), *eRosita* (Pillepich et al., 2012), the *Nancy Grace Roman Telescope*, or UNIONS (Chambers et al., 2020), are expected to provide data for very large samples of galaxy clusters. Furthermore, new approaches to the analysis of complex nonlinear data, such as those associated with machine learning, are becoming more common. These new data sets and new analysis tools make cosmological analyses with cluster structural properties a promising avenue to explore.

The idea of using the cosmological dependence of cluster formation histories to constrain cosmology is not new, but was discussed in the literature three decades ago (Richstone et al., 1992; Evrard et al., 1993; Mohr et al., 1995). These original tests leveraged the fact that the structural properties of galaxy clusters are related to how relaxed they are – their projected shape and non-axisymmetry – and to the state of the universe when they accreted their mass – concentration – (see Taylor, 2011, for a review). Subsequent work has examined structural properties, showing that they are generally consistent with expectations from LCDM (e.g. Oguri et al., 2010; Sereno et al., 2018), but has not used them to constrain cosmological parameters specifically.

In Amoura et al. (2021) (Chapter 2), we showed that for values of σ_8 varying between 0.75-0.85, the resulting median age of galaxy clusters, as expressed by the epoch z_{50} by which a system had accreted half its final mass, would vary by more than 10%. Combining accurate, unbiased measurements of structural parameters such as concentration for a large enough sample, such a difference could easily be distinguished in future cluster samples. More importantly, for clusters of mass $\sim 10^{14} M_{\odot}/h$ at low redshift, the constraints obtained this way are orthogonal to the typical banana-shaped constraints following contours of constant S_8.

While our previous work focussed on the overall 'age' of clusters, i.e. some average measure such as z_{50} defined over their whole accretion history, the instantaneous growth rate may sometimes be easier to determine from observations. This rate is reflected in halo merger rates, the mean increase in mass over some recent interval of time, or the fraction

of systems that have recently experienced a large increase in mass.

Tests of halo growth or the merger rate could in principle be applied on any mass scale traced by visible matter. The merger rate on galaxy scales has been studied extensively both observationally and in simulations, using various tracers of merger activity, including close pairs of galaxies, starbursts, and morphologically distorted galaxies (e.g. Lotz et al., 2011; Xu et al., 2012; Mundy et al., 2017). Since the machinery for estimating halo merger rates is the same independent of scale, we will also consider galaxy-scale growth and merger rates, although constraining these with observations involves several additional challenges, as discussed in Section 4.

The outline of the chapter is as follows. In Section 2, we use analytical models based on the Extended Press-Schechter (EPS) formalism to estimate how various measures of halo growth vary with Ω_m and σ_8. In Section 3 we compare these predictions to dark-matter-only N-Body simulations, and discuss the discrepancies between the two. In Section 4 we consider the prospects for measuring halo growth observationally, either directly on cluster scales, or indirectly on galaxy halo scales. We summarize our results and conclude in Section 5.

3.3 Cosmological Sensitivity of Halo Merger Rates

An analytic estimate for the halo merger rate was first derived by Lacey & Cole (1993), using the approach of Press & Schechter (1974) to create the so called 'Extended Press-Schechter' (EPS) formalism. Sheth & Tormen (2002) derived a major correction to Press-Schechter theory, accounting for ellipsoidal collapse; this was subsequently included in EPS theory, e.g. by Zhang et al. (2008). The halo merger rate has also been measured in N-body simulations, starting with Lacey & Cole (1994). Early work by Gottlöber et al. (2001), for instance, studied the dependence on environment, while Fakhouri & Ma (2008) and Fakhouri et al. (2010) used the Millenium simulations to obtain accurate global merger rates, providing a framework to count mergers and compare numerical results to EPS predictions, as well as a universal fitting formula. We will use these results as the basis for most of our calculations. (For an alternative approach, that counts the rate per progenitor instead of the rate per descendant halo, see Genel et al. 2009.)

Given the indirect connection between galaxies and haloes, the galaxy merger rate should behave slightly differently from the halo merger rate, as discussed further in Section 4. Stewart et al. (2009) used N-body simulations to estimate how observable indicators of galaxy mergers should scale with galaxy luminosity, stellar mass, merger mass ratio and

redshift. More recently, galaxy merger rates have been estimated from hydrodynamical simulations (e.g Rodriguez-Gomez et al., 2015). We expect these estimates to be more accurate than earlier, dark-matter only results, although they are typically only applicable to a single cosmology.

3.3.1 Merger Rate Definitions

We will follow the definitions of Fakhouri & Ma (2008) in describing the merger rate: first, the symmetric merger rate $B(M_1, M_2, z_D)dM_1dM_2$ is the average rate per unit volume per unit redshift, between progenitors of masses between M_1+dM and M_2+dM, where z_D is the redshift at which the descendant is identified. This rate has units of mergers/volume/(unit redshift). If we want to express the rate in terms of the descendant mass $M_0 = M_1+M_2$ and the merger mass ratio $\xi = M_2/M_1$ instead, we can use the function $B(M_0, \xi, z_D)dMd\xi$, also with units mergers/volume/(unit redshift). If we normalise the rate by the halo number density $n(M_0, z_D)$, we get the dimensionless quantity B/n, with units mergers/dz/$d\xi$:

$$B/n \equiv \frac{B(M_0, \xi, z_D)}{n(M_0, z_D)} \tag{3.1}$$

This quantity will be the basis for all the rates that we consider in this chapter. We can also integrate B/n in order to calculate the rate of all mergers between mass ratios x and X, at fixed descendant mass.

$$\frac{dN}{dz}(M_0, z_D, x, X) = \int_x^X \frac{B}{n}(M_0, z_D)d\xi \tag{3.2}$$

3.3.2 Analytical Models

The Extended Press-Schechter (EPS) formalism provides an analytic framework to estimate the merger or growth rates of dark matter haloes (Lacey & Cole, 1993), based on the spherical collapse model. The merger rate per halo, as a function of the descendant mass M_0 and the merger ratio $\xi = M_2/M_1$, is

$$\frac{B(M_0, \xi, z)}{n(M_0, z)} = \sqrt{\frac{2}{\pi}}\frac{d\delta_c}{dz}\frac{1}{\sigma(M_1)}\left|\frac{d\ln\sigma}{d\ln M_1}\right|\left[1 - \frac{S(M_0)}{S(M_1)}\right]^{-3/2}, \tag{3.3}$$

where M_1 is one of the progenitors, δ_c is the critical overdensity for collapse, and $S \equiv \sigma^2(M)$ is the variance of the linear density field smoothed at a scale corresponding to a mass M.

This model can be made more accurate by using an ellipsoidal collapse model instead (Sheth & Tormen, 1999, 2002). Zhang et al. (2008) provide an ellipsoidal collapse correction to the spherical collapse merger rate

$$\frac{B(M_0,\xi,z)}{n(M_0,z)} = \left.\frac{B(M_0,\xi,z)}{n(M_0,z)}\right|_{\rm sph} \times A_0 \exp\left(-\frac{A_1^2\tilde{S}}{2}\right)\left[1+A_2\tilde{S}^{3/2}\left(1+\frac{A_1\tilde{S}^{1/2}}{\Gamma(3/2)}\right)\right], \tag{3.4}$$

where $A_0 = 0.8661(1-0.133\nu_0^{-0.615})$, $A_1 = 0.308\nu_0^{-0.115}$, $A_2 = 0.0373\nu_0^{-0.115}$, $\nu_0 = \omega_0^2/S(M_0)$, $\tilde{S} = \Delta S/S(M_0)$, $\Delta S = S(M_1) - S(M_0)$, and $\omega \equiv \delta_c(z)$. The difference between spherical and ellipsoidal collapse-based merger rates is illustrated in Fig. 3.1. Throughout this chapter, our analytical predictions will all be based on Eqn. 3.4.

3.3.3 Cosmological Dependence of Merger Rate

Analytical models provide a practical way to estimate how merger and growth rates will vary with the cosmological parameters. The merger rate is sensitive to cosmology through the power spectrum, as reflected by the variance of the smoothed field of density perturbations $S(M|\Omega_m,\sigma_8)$, and the growth factor D, through the critical overdensity for collapse $\delta_c(z) = \delta_c/D(z|\Omega_m)$ where $\delta_c = 1.686\Omega_m^{0.0055}$. Details of how Ω_m and σ_8 influence the cluster number count and formation time through their effect on the matter power spectrum and linear perturbation growth rate, as well as the resulting banana-shaped constraints, are discussed in Amoura et al. (2021).

In Fig. 3.2, we show how the merger rate estimated from Eqn. 3.4 varies in the Ω_m–σ_8 parameter space, for group- (left-hand panels) and cluster-mass (right-hand panels) haloes, and for three different mass ratios (top to bottom), at $z = 0.3$. The colour scale shows the variation in the rate relative to a fiducial value calculated for $\Omega_m = 0.3$, $\sigma_8 = 0.8$:

$$\Delta B/B_{fid} = \frac{B/n(M,z,\xi|\Omega_m,\sigma_8) - B/n(M,z,\xi|0.3,0.8)}{B/n(M,z,\xi|0.3,0.8)}. \tag{3.5}$$

We see that the sensitivity to Ω_m and σ_8 is independent of merger mass ratio. The rate varies by about 20% to 30% over the range of σ_8 considered here, but depends only weakly on Ω_m, with almost no dependence at the group mass scale. Thus, the halo merger rate can in principle be used to measure σ_8 independently from Ω_m.

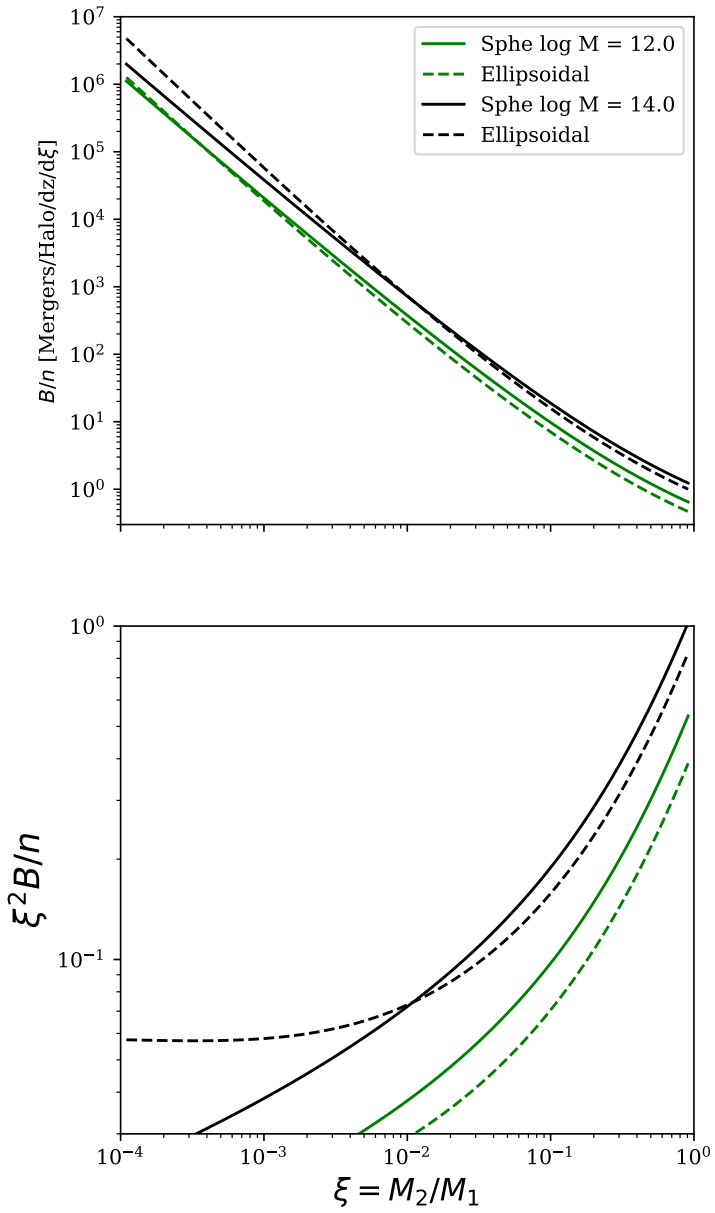

Figure 3.1: Halo merger rates predicted by the spherical collapse (Lacey & Cole, 1993) and ellipsoidal collapse (Zhang et al., 2008) models, for the two halo masses indicated, at z=0.1. The bottom panel shows the rate weighted by ξ^2 to highlight the differences.

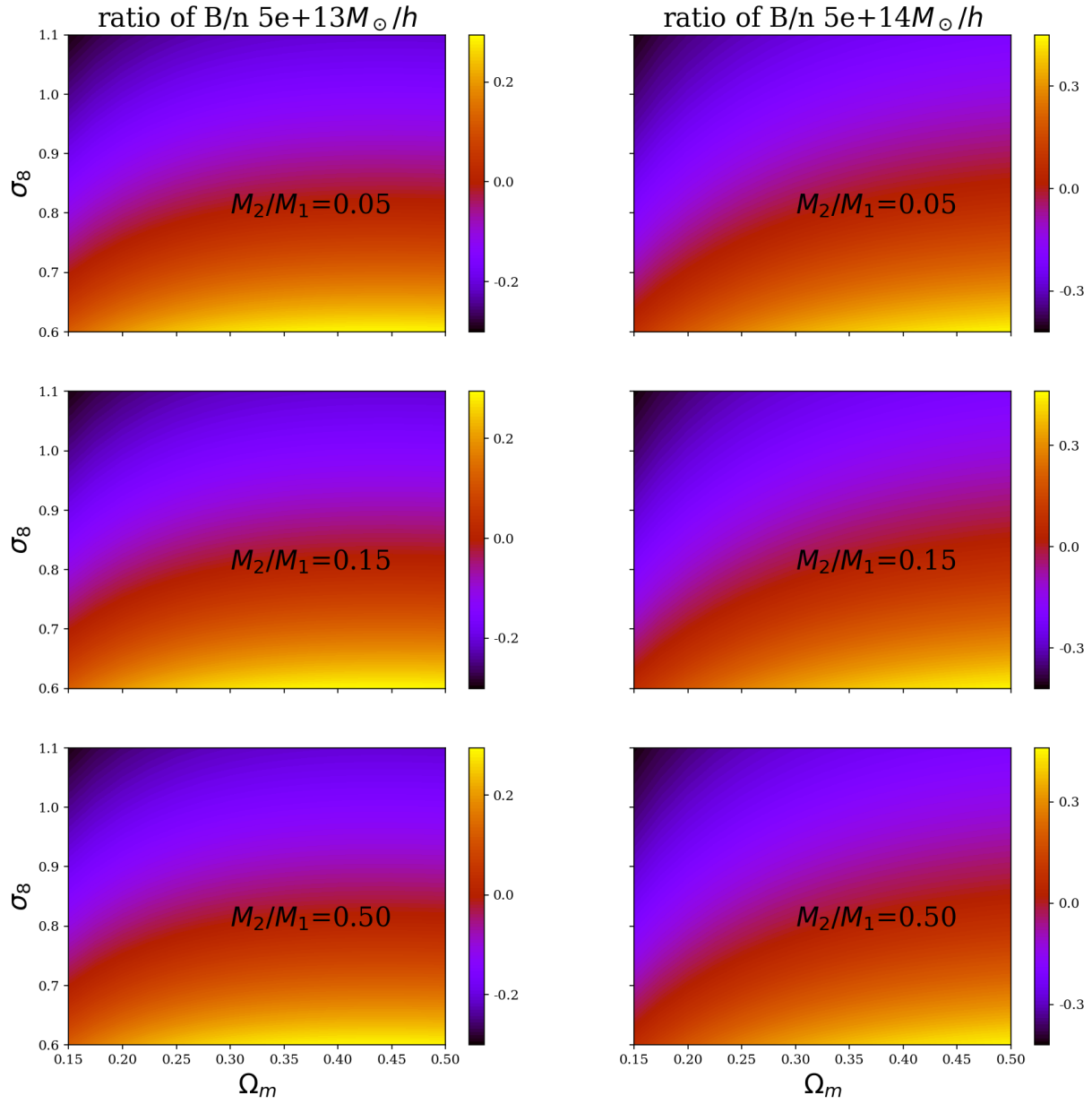

Figure 3.2: Variation in the merger rate per halo B/n at $z = 0.3$, as a function of Ω_m and σ_8, relative to a fiducial rate for $\Omega_m = 0.3$ and $\sigma_8 = 0.8$. The rate is calculated assuming the ellipsoidal collapse model (Eqn. 3.5).

3.3.4 Average Halo Growth

Material accreted onto a halo through mergers will settle into the main potential through tidal stripping and dynamical friction, over a timescale on the order of the dynamical time $t_{\rm dyn}$. Thus, in addition to the instantaneous merger rate, we also consider the the net increase in halo mass over this timescale. For a given final redshift z_0, we first calculate the redshift z_1 corresponding to one dynamical time in the past. The amount by which a halo grows over this redshift interval should then be

$$\Delta M = \int_{z_0}^{z_1} dz \int_0^1 B/n(M(z), z, \xi) \left[\frac{\xi}{1+\xi} M(z)\right] d\xi \,. \tag{3.6}$$

Since the merger rate varies slowly with mass, and the dynamical time is short enough that major mergers are rare, we can make the approximations $B/n(M(z), z, \xi) \sim B/n(M_0, z, \xi)$ and $(1/(1+\xi))M(z) \sim M_1$ and define the Average Halo Growth (AHG) as:

$$AHG(M_0, z_0) = \left\langle \frac{\Delta M}{M_1} \right\rangle \equiv \int_{z_0}^{z_1} dz \int_0^1 B/n(M_0, z, \xi) \xi d\xi \,, \tag{3.7}$$

that is, it is the increase in mass a halo experiences over one preceding dynamical time, relative to its initial mass, as a function of the final mass and redshift.

Fig. 3.3 shows the cosmological dependence of the AHG, for three different redshifts (top to bottom), and the same group and cluster masses as in Fig. 3.2. As expected, haloes tend to grow faster at these redshifts in low σ_8 and/or high Ω_m universes. The influence of Ω_M becomes weaker relative to σ_8 for lower masses, and for higher redshifts. We note that contours of constant AHG are almost orthogonal to those typical of cluster abundance or weak lensing constraints (cf. Chapter 2). While the amplitude of the variation depends on the mass and redshift, there is generally between 50% to 100% change in the AHG between cosmologies with $\sigma_8 = 0.7$ and those with $\sigma_8 = 0.9$.

3.3.5 Large-growth Systems

Another summary statistic with a close connection to observable phenomena is the fraction of haloes that experience a large increase in mass over a given period of time. As for the AHG, we choose the dynamical time $t_{\rm dyn}$ as the relevant timescale, and count the fraction of systems that have grown by more than 1/3 over this time. To estimate this fraction analytically, we make the approximation that the growth involves a single large merger

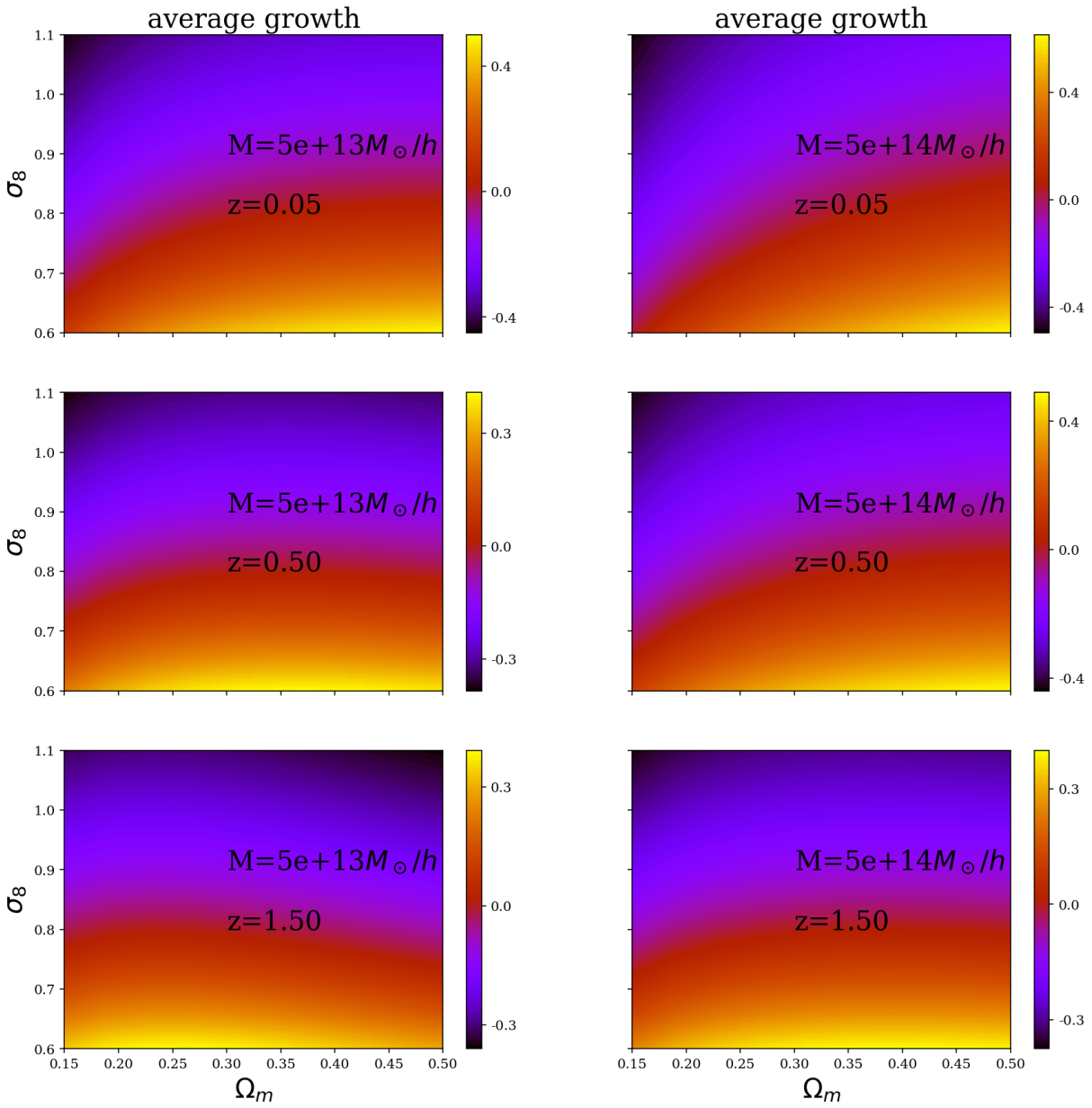

Figure 3.3: The average halo growth (AHG) over the last dynamical timescale $t_{\rm dyn}$, as a function of Ω_m and σ_8 relative to the value at $\Omega_m = 0.3$ and $\sigma_8 = 0.8$, for the masses and redshifts indicated.

with $\xi > 1/3$, such that

$$LGS(M_0, z_0) \equiv \int_{z_0}^{z_1} dz \int_{1/3}^{1} B/n(M_0, z_0, \xi) d\xi \,. \tag{3.8}$$

Fig. 3.4 shows how the LGS fraction depends on Ω_m and σ_8, for the same mass and redshift bins as Fig. 3.3. The cosmological dependence is almost identical to that of the AHG, such that both quantities could in principle provide cosmological tests of comparable sensitivity.

3.4 Comparison to simulations

As discussed in Chapter 2, the analytic models of the previous section are only approximate. To test their validity, we will also consider merger rates measured in several different N-body simulations.

3.4.1 Simulation Data

We use a set of dark-matter-only simulations to test the analytical predictions of the ellipsoidal collapse model. These include publicly available halo catalogues and merger trees, but also our own set of simulations run for different cosmologies, as follows :

1. The Illustris-TNG simulation (Nelson et al., 2019) which uses the SUBFIND halo finder and the SUBLINK merger tree algorithm (Rodriguez-Gomez et al., 2015).

2. The Bolshoi/BolshoiP simulation (Klypin et al., 2011), with a halo catalog generated with ROCKSTAR (Behroozi et al., 2013a) and merger trees generated with the CONSISTENT TREES algorithm (Behroozi et al., 2013b).

3. A set of 9 of our own simulations, introduced in Chapter 2. We will refer to these as MxSy, where x can be 25/3/35 for $\Omega_m = 0.25/0.3/0.35$ respectively, and y can be 7/8/9, for $\sigma_8 = 0.7/0.8/0.9$ respectively. These simulations were run with GADGET 2 (Springel, 2005), and the halo catalogue and merger trees were generated with the AMIGA HALO FINDER (AHF; Knollmann & Knebe, 2009).

Simulation parameters are summarized in Table 3.1.

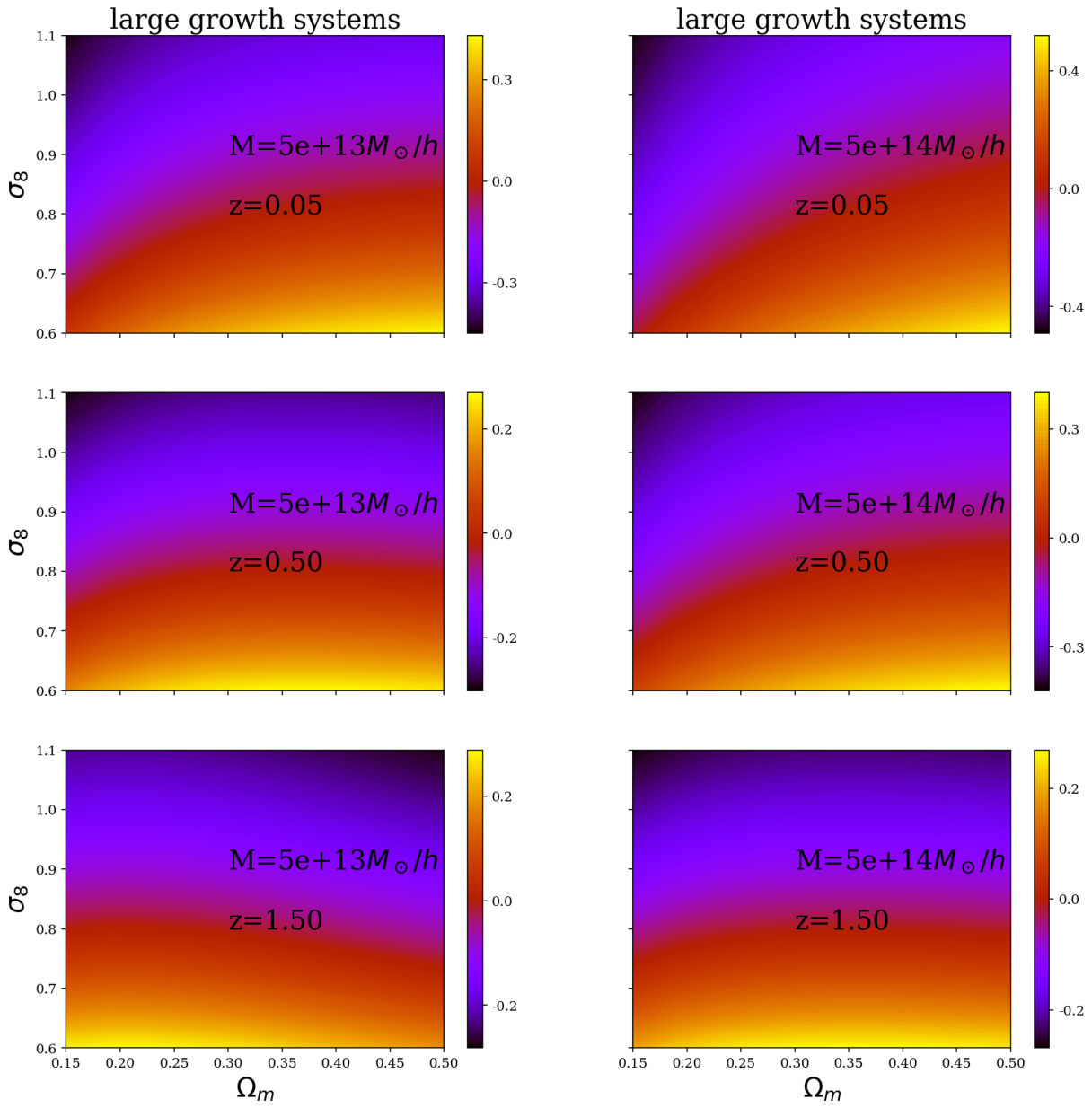

Figure 3.4: The fraction of large-growth systems (LGS) as a function of Ω_m and σ_8 relative to the value at $\Omega_m = 0.3$ and $\sigma_8 = 0.8$, for the same mass and redshift bins as in Fig. 3.3.

Name	$\Omega_{\rm m}$	σ_8	$m_{\rm part}$ $[M_\odot/h]$	$N_{\rm part}$	merger tree	$N_{\rm snap}$
TNG	0.31	0.81	3×10^9	625^3	Sublink	100
Bolshoi	0.27	0.82	1.35×10^8	2048^3	Cons. Trees	181
BolshoiP	0.31	0.82	1.55×10^8	2048^3	Cons. Trees	178
MxSy	0.25/0.3/0.35	0.7/0.8/0.9	4×10^9	512^3	AHF	44

Table 3.1: Summary of the simulations used and their main parameters, including the cosmological parameters, the particle mass, the total number of particles $N_{\rm part}$, the merger tree code, and the number of snapshots $N_{\rm snap}$ used to make the merger trees. The MxSy simulations are a set of 9 of our own simulations that span a range of different values of $\Omega_{\rm m}$ and σ_8.

3.4.2 Merger Rates

To estimate the merger rate per descendent halo, B/n, in the simulations, we count the individual mergers associated with a given descendent as follows. Going to the previous snapshot, we identify all $N_{\rm prog}$ progenitors of the descendent, and count a total of $N_{\rm prog} - 1$ mergers, each with the most massive progenitor (implying merger mass ratios $\xi < 1$). The exact definition of the progenitors varies, depending on the simulation and the merger tree algorithm. In addition, each simulation has a different snapshot frequency, which can affect the measured merger rate at high redshift. After some experimentation, for most of our tests we restricted ourselves to merger rate estimates from our own simulations, where the analysis is homogeneous.

Fig. 3.5 shows the numerical merger rate as a function of mass ratio ξ from each simulation, compared to the analytical predictions. Generally, the numerical results are in reasonable agreement with the analytical models, but include far fewer major mergers. This may be an artefact of our method for counting mergers. If the progenitors of a given descendent include one large halo and several smaller ones, we always count $N - 1$ minor mergers between the largest progenitor and each of the other progenitors. With a higher frequency snapshots, we might find that intervening major mergers had occurred between pairs of low-mass progenitors, before they merged with the largest progenitor. In addition, tidal stripping can cause systems to lose some mass even before they are recorded as merging. Either of these effects could explain the deficit of large mass-ratio mergers and the slight excess of low mass-ratio mergers.

In order to study the cosmological dependence of the merger rate more specifically, we calculated the number of mergers between z=0.05 and z=0.45, for mass ratios between $0.01 < \xi < 0.03$, $0.03 < \xi < 0.1$ and $0.1 < \xi < 0.3$. We avoided major mergers, given

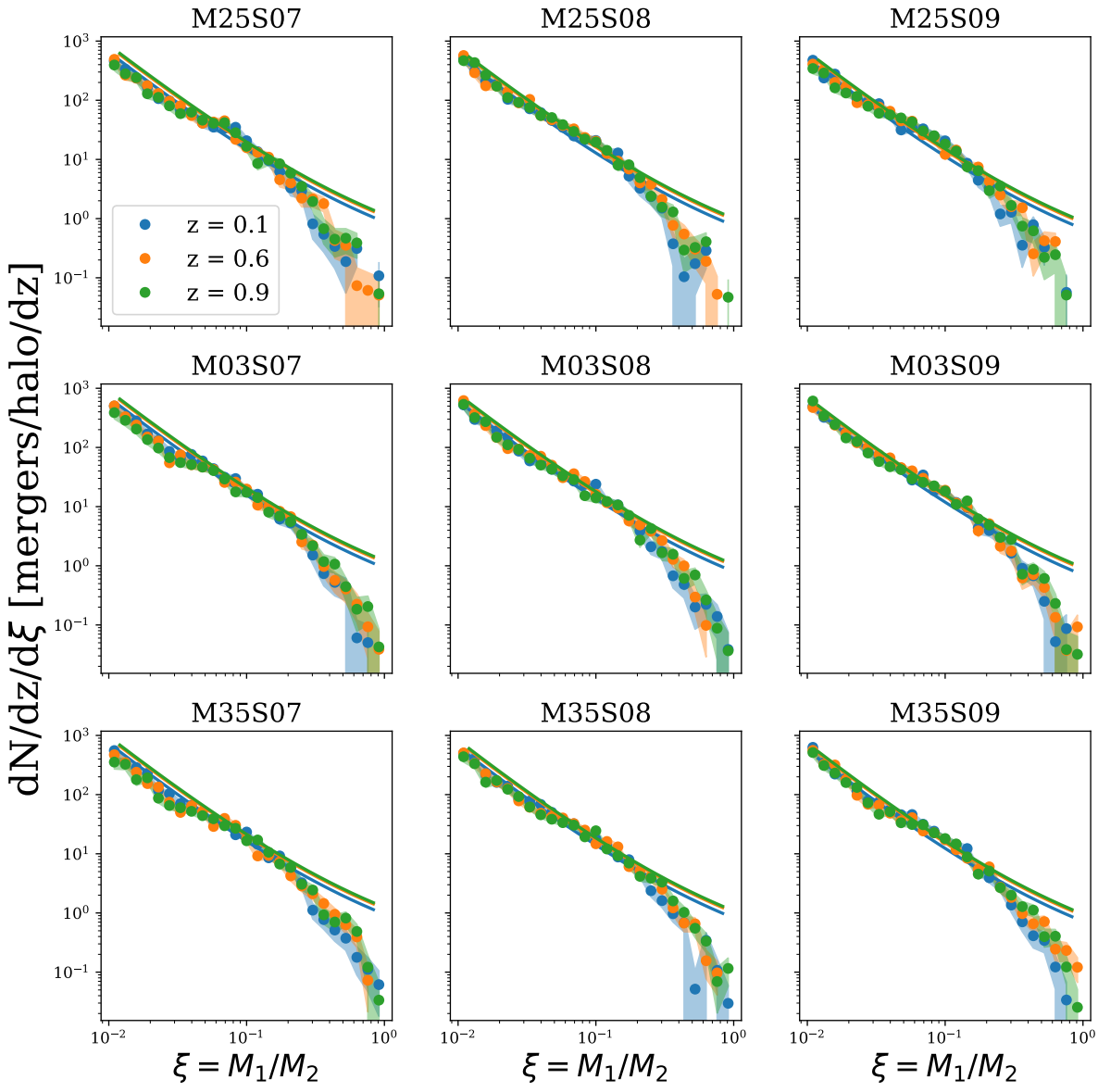

Figure 3.5: The merger rate per descendant halo as a function of mass ratio ξ for each of the MxSy simulations (points and shaded regions), compared to the EPS rate predicted by the ellipsoidal collapse model. Note the deficit of major mergers, relative to the analytic predictions.

the potential problems described above. We also restricted ourselves to lower redshifts, where the snapshot cadence is reasonably frequent relative to the dynamical time. We considered all haloes with $M > 10^{13} M_{\odot}$, first to avoid mass resolution effects at low mass, and second to have enough statistics given the first constraint. Binning all masses together is reasonable, given that the merger rate is only weakly sensitive to mass, going as $\sim M^{0.13}$ (Fakhouri et al., 2010; Genel et al., 2009).

Fig. 3.6 shows these merger rates, as a function of Ω_m at fixed σ_8 (top panels), and as a function of σ_8 at fixed Ω_m (bottom panels), compared to the analytical predictions. Both numerical and analytic results show the same general behaviour. The simulations contain more minor mergers ($0.01 < \xi < 0.03$) than predicted by theory, which may reflect the counting problems discussed above, but the dependence on cosmological parameters is similar between the numerical and analytic results.

3.4.3 Average Halo Growth

Next, we compare the average halo growth rate measured in simulations to the rate predicted by EPS theory. The comparison is particularly interesting, since the simulations and halo finders have finite resolution, and will always miss a component of the merger history below their resolution limit. For this test, we consider results from all the simulations listed in Table 1, to highlight the differences between them. For every halo in each simulation, we measure the mass growth over one dynamical time $\left(M(z - z_{dyn}) - M(z)\right)/M(z)$ and average this quantity in each mass bin. We then calculate the same quantity in analytical models by integrating the instantaneous merger rate over the same redshift range. The resulting rates are shown in Fig. 3.7.

Bolshoi simulations agree with very well at all redshifts, while Illustris shows a flattening at high mass. The set of MxSy simulations agree well at low z and less at high z. Most simulations have lower growth rates at high redshift than predicted by theory. On the other hand, all the numerical results agree with the analytic predictions in the general mass and redshift dependence in the growth rate, often differing by a single overall shift in normalization. We speculate that halo-finding algorithms may be at the origin of this discrepancy between different simulations, and between simulations and analytical predictions, as discussed in the literature (Knebe et al., 2011; Avila et al., 2014; Hopkins et al., 2010). Cases of haloes losing mass, flyby events and other numerical artifacts introduced by the different ways haloes are defined, detected and linked in different halo-finder and merger tree algorithms can cause an artificial increase in the average growth. Even after accounting for and removing the most spectacular events, where haloes appear to gain

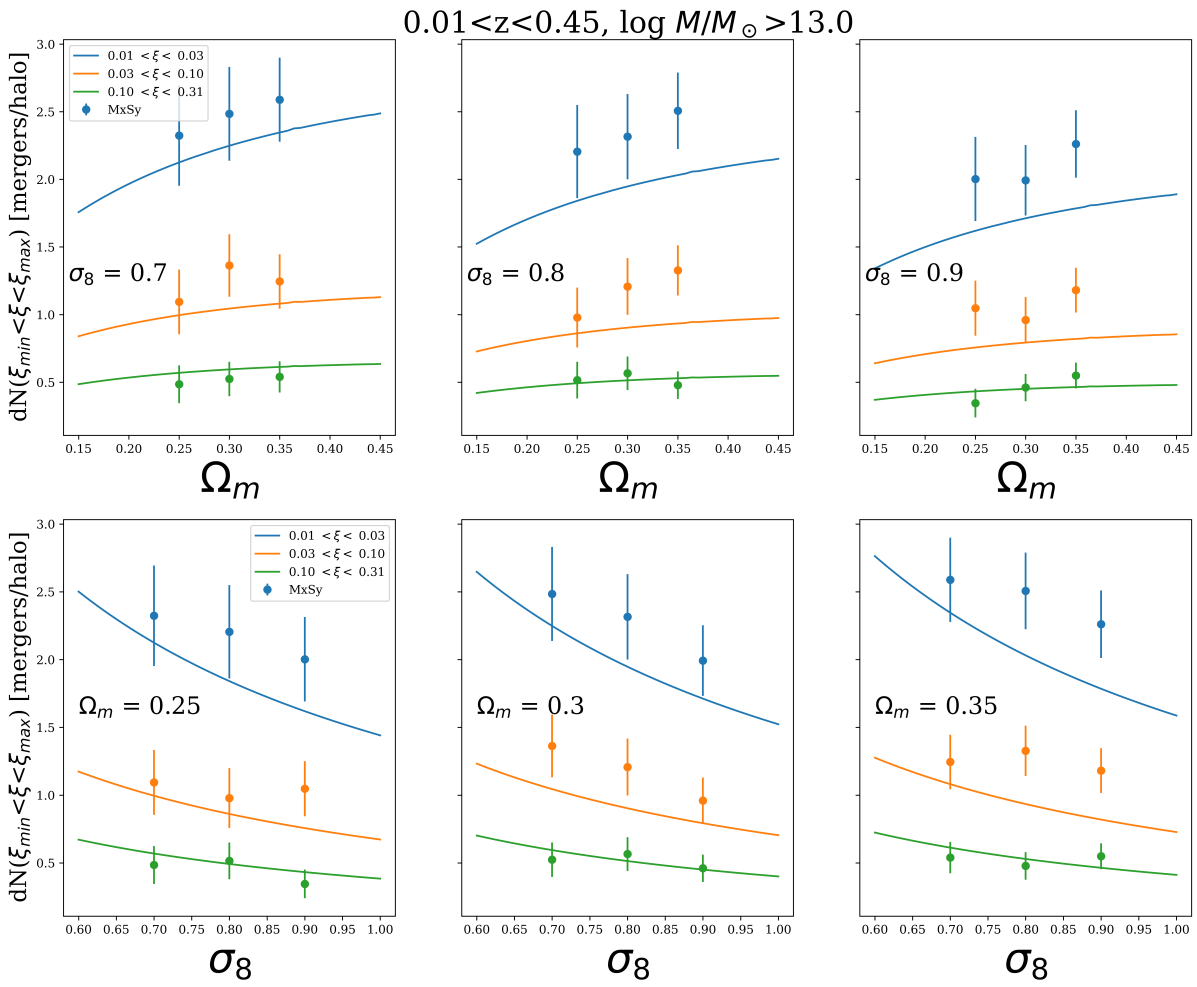

Figure 3.6: Cosmological dependence of the merger rate for various mass ratios. The points with errorbars indicate rates estimated from the MxSy simulations, while the smooth curves show the analytic predictions.

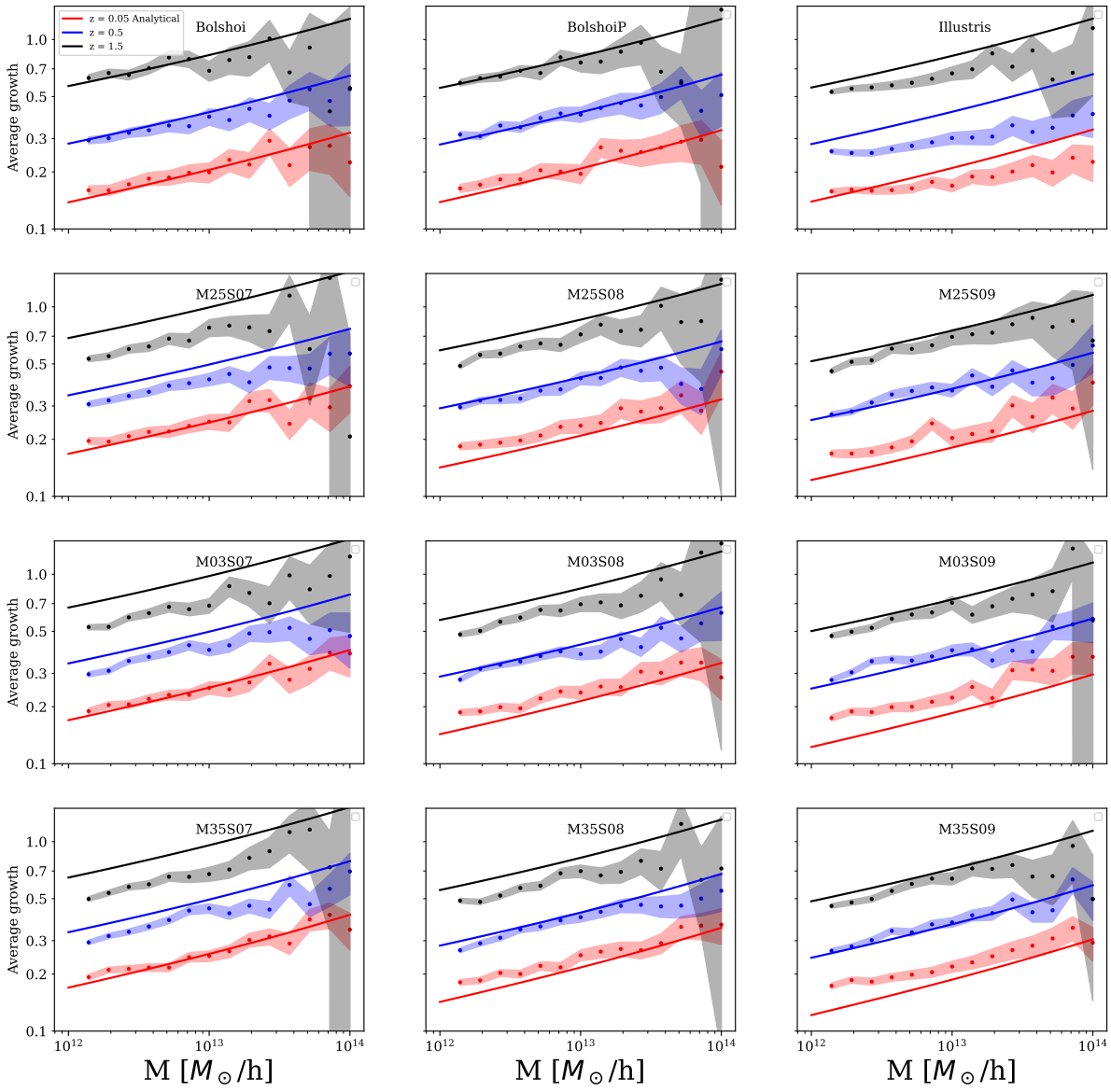

Figure 3.7: The average growth of haloes since the last dynamical timescale t_{dyn}, measured in the simulations indicated. The shaded areas represent Poissonian errors. Solid lines show the EPS predictions for comparison.

several times their mass between consecutive snapshots, the average growth at low redshift remains larger than analytical predictions.

We now consider the cosmological dependence of the AHG. Given the differences between different simulations and analysis tools shown above, we will restrict ourselves to our own MxSy simulations, which represent a homogeneous set. To simplify the comparison between simulations, we fit the simulation results with a power-law

$$AHG(M) = A(M/M_0)^{\alpha} + p\,, \tag{3.9}$$

where the parameters A, M_0 and α are fixed in each panel, and the normalisation p varies with σ_8. For z=1, we find that a broken power law is a better fit. This fit is meant to reproduce the overall mass dependence of the AHG specifically for our set of simulations. We show the fits and discuss them further in Appendix C.

In Fig. 3.8, we compare the dependence of the AHG on σ_8, for different values of Ω_m, and for different masses and redshifts. As we have seen already in Fig. 3.7, the halo growth from simulations is significantly lower at high redshift than the analytic prediction. Accounting for this redshift-dependent offset, the numerical results show the predicted drop in growth with increasing σ_8, but seem less sensitive to σ_8 than expected, particularly at high redshift. The numerical values in high-σ_8 (low growth) cosmologies exceed the analytic predictions. The origin of this discrepancy is not immediately clear. One possibility is that because of the relatively large spacing between snapshots in the MxSy simulations, the true growth rate is *over*-estimated in cosmologies where it is intrinsically low. We will investigate this possibility in future work.

3.4.4 LGS Fraction

In Fig. 3.9 we show the LGS fractions measured in the simulations, compared to the analytic predictions. Unlike the AHG, all simulations are in agreement with the analytic predictions at z=0.05; at higher redshift, the mass dependence seems slightly flatter than predicted. Note that Illustris is the outlier again, finding considerably more large growth systems at high redshift.

As for the AHG, we fit the power law in Eqn. 3.9, for each of the MxSy simulations, and show the variation of the normalisation at different values of σ_8 in Fig. 3.10. The range of variation of the LGS fraction with σ_8 is closer to the analytical predictions at low redshift, but still slightly smaller at higher redshift.

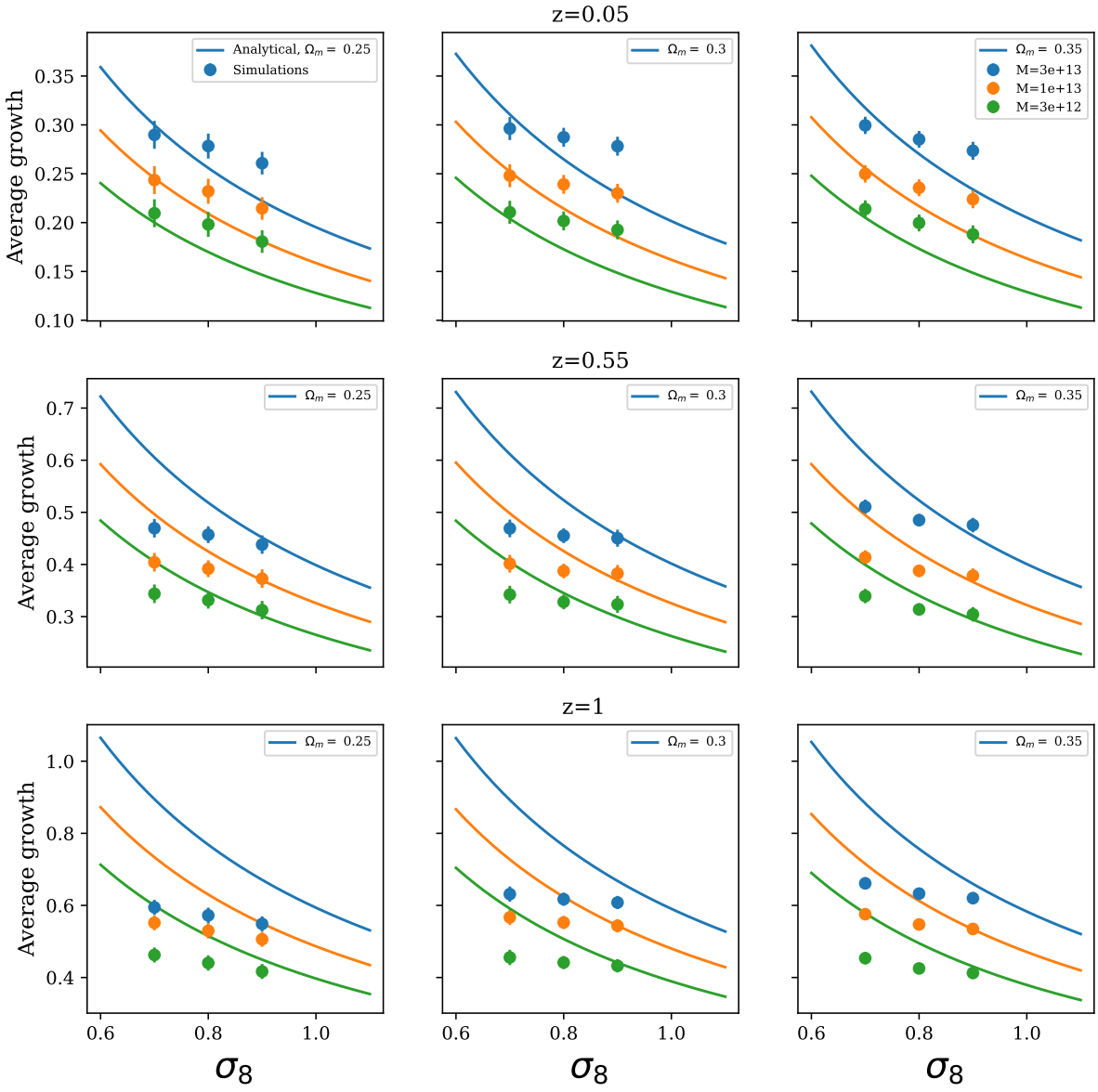

Figure 3.8: Average halo growth as a function of σ_8, for the redshifts, masses, and values of Ω_m indicated. The points show the results measured in the MxSy simulations, while the curves show the EPS prediction.

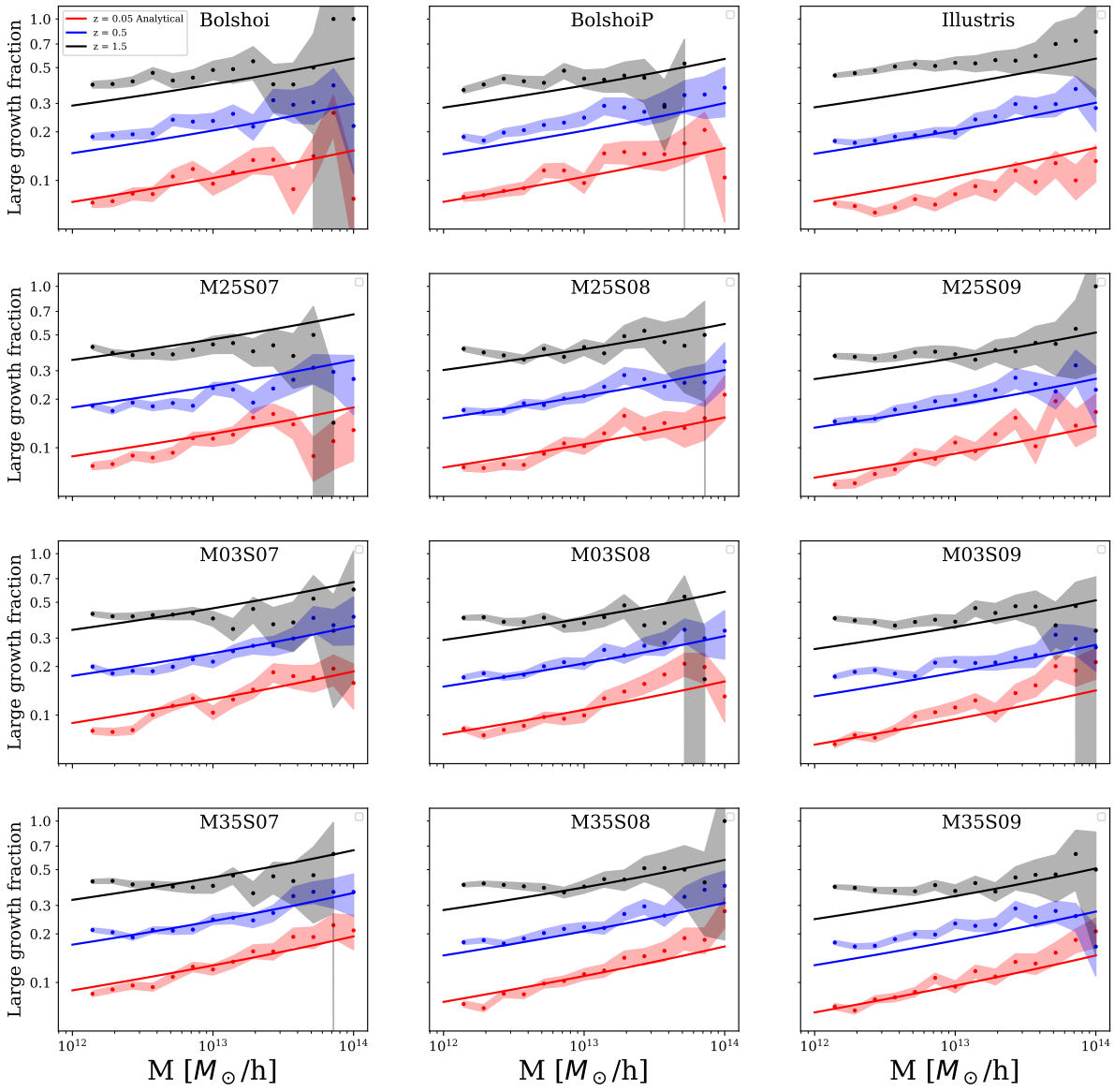

Figure 3.9: Fraction of haloes that experienced a large growth ($> 1/3$) since the last dynamical timescale t_{dyn}, as a function of mass (points with shaded error regions). Solid lines show the EPS predictions.

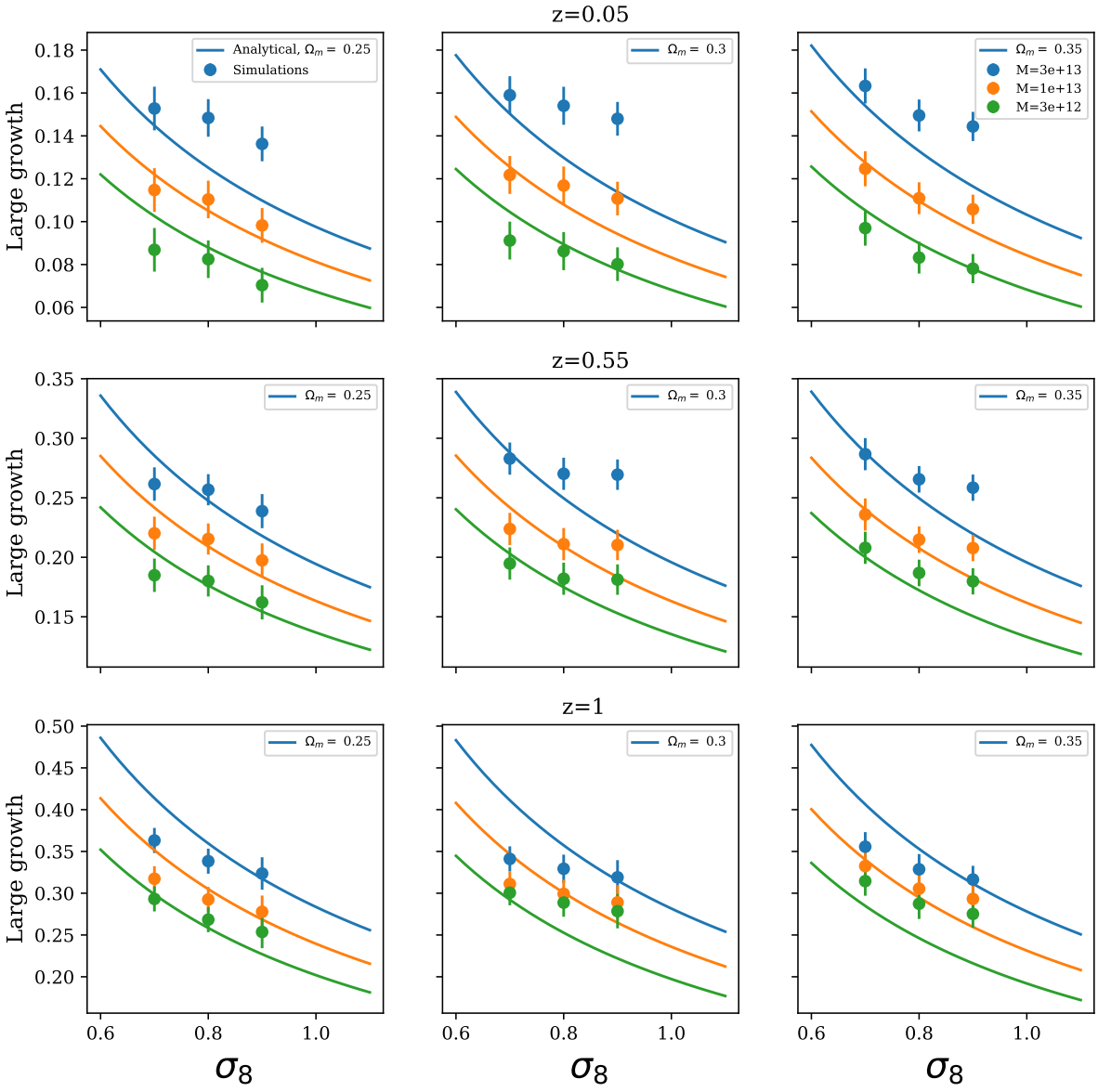

Figure 3.10: Comparison of the σ_8 dependence of the fraction of haloes with large growth between simulations and analytical models. The cosmological trend is similar, with a lower amplitude.

3.4.5 Simulations vs. Analytic Predictions: Summary

Comparing numerical and analytic results, we find broad agreement, but also some discrepancies. Unfortunately, without further detailed work, it is not clear which if either is the most accurate, although we suspect at least some of the discrepancies are related to the snapshot cadence and merger tree algorithms used to analyse the simulations. The predicted and measured cosmological dependence are in closest agreement for the halo merger rate and the LGS fraction. Focussing on these quantities, we infer that we could differentiate between values of σ_8 between 0.7 and 0.9 if we could measure either one with a precision of better than 10% while avoiding any observational systematics. In Section 3.5 below, we will consider whether this goal is realistically achievable.

3.5 Observational prospects

The results of the previous section suggest that $\sim$10% precision would be required in merger or growth rate measurements, in order to provide useful cosmological constraints. We will now examine whether this precision could be reached in practice. We consider tests on two scales, either the galaxy cluster scale, or the scale of individual galaxy haloes.

3.5.1 Measuring Merger and Growth Rates on Cluster Scales

While galaxy clusters are relatively rare, in the near future multiple missions and surveys including *Euclid* in the optical and IR (Sartoris et al., 2016), *eROSITA* (Pillepich et al., 2012) in the X-ray, CMB-S4 (Abazajian et al., 2019b) in the mm, and the ground-based UNIONS (Chambers et al., 2020), DESI (DESI Collaboration et al., 2016), and Rubin LSST (LSST Science Collaboration et al., 2009) surveys should produce mass-limited samples of $O(10^4)$ clusters with sufficient signal-to-noise ratio (SNR) to allow structural measurements. Clusters with sufficient SNR to detect major mergers should number in the hundreds or thousands. Furthermore, these will typically be low redshift, massive systems where complementary information from many modalities is available, including galaxies with measured redshifts, weak and/or strong lensing mass models, X-ray surface brightness maps, and SZ maps in the sub-mm.

To measure the instantaneous halo merger rate for clusters would require identifying all infalling groups at or near the virial radius. This could be challenging due to projection effects and/or limited galaxy redshift information. Furthermore, one would need to estimate total masses for the infalling systems, with errors in the mean mass for a sample not

exceeding 10%. A realistic survey of $O(400)$ massive clusters with weak lensing mass maps might identify infalling systems in, say the range $\xi = 0.1$-0.2, with 50-70% completeness, over a narrow redshift range where the average number of mergers is one per cluster. The Poisson uncertainty in the merger rate would then be $1/\sqrt{400} = 5\%$; the uncertainty in the mean mass of the infalling systems would be $200\%/\sqrt{400} = 10\%$, while the uncertainty in the completeness might be $\sim$20%. We conclude that while the first two sources of uncertainty are close to the goal of 10% errors, the uncertainty in the completeness would be too large to obtain useful cosmological constraints.

Alternately, one could consider measuring the LGS fraction. Systems that have recently accreted a third or more of their material would be easier to identify, via kinematic substructure, offsets between the gaseous, stellar and dark components (e.g. Clowe et al., 2006; Mann & Ebeling, 2012; Zenteno et al., 2020), or overall X-ray morphology (e.g. Yuan et al., 2022). Assuming these features can be detected regardless of projection effects, we may assume approximately 100% completeness in the LGS sample. Assuming a LGS fraction 20-30% for massive clusters at low redshift, a sample of 400 might produce 100 LGS systems, resulting in Poisson errors with the required uncertainty of 10%. On the other hand, distinguishing between degrees of relaxation (e.g. systems that had experienced large growth within the past 1.0 dynamical times, versus 2.0 or 0.5 dynamical times) might be more challenging, and would require extensive calibration with simulations.

Overall, we conclude that measuring the halo merger rate or growth rate on cluster scales seems challenging, but not impossible. At a minimum, future cluster samples should provide a consistency test for parameters derived from other methods.

3.5.2 Measuring Merger and Growth Rates on Galaxy Scales

Given that halo merger and growth rates depend only weakly on halo mass, and galaxy haloes are far more abundant, it is worth considering tests based on this smaller mass scale. Galaxy merger rates have been studied extensively, both theoretically, either through semi-analytical/semi-empirical models (e.g. Stewart et al., 2009; Huško et al., 2022) or hydrodynamical simulations (e.g. Rodriguez-Gomez et al., 2015; Pfister et al., 2020; Contreras-Santos et al., 2022), and observationally (e.g. Lotz et al., 2011; Xu et al., 2012; Mundy et al., 2017). Two important complications arise in relating galaxy merger rates to galaxy halo merger rates; first, the delay between the two, and second the relation between halo mass and stellar mass. We consider each of these in turn.

Delay Time Due to Infall

Mergers between galaxy haloes, as defined in most analyses of numerical simulations, occur around the virial radius. Assuming both haloes contain visible galaxies, halo mergers then lead to galaxy mergers, after some delay for infall to the centre of the main halo. Galaxy mergers are identified observationally using features—close pairs, tidal distortion and debris, or triggered starbursts—that trace the first and second pericentric passages. We note that the delay due to infall means that 'merging' galaxies identified via these features at one redshift actually trace the halo merger rate at a higher redshift.

Fig. 3.11 shows the observed redshift at which the first (top curve) and second (bottom curve) pericentric passages occur for a given halo merger redshift z_{hm}. (The dotted line shows a 1-1 correspondence for reference.) To calculate these, we have assumed that the orbital properties of the satellite and the potential of the main system are conserved, and that pericentric passages occur around 1/8 and 9/8 of the radial orbital period at the virial radius (Taylor & Babul, 2004), which corresponds to 0.1 and 0.9 times the period of a circular orbit at the virial radius, $P_{\rm vir}$, or 0.06 and 0.57 times the Hubble timescale $H(z)^{-1}$ at the redshift z_{hm}.

From this figure, we see that while the first pericentric passage occurs at only slightly lower redshift than the initial merger, the second pericentric passage occurs significantly later, and is only observable for halo mergers at $z_{\rm hm} > 1.5$. These calculations assume conservation of the orbit and the potential over 1 or more radial periods; the reality in major mergers is more complicated, and a significant fraction of orbits may get scattered in these cases (de Luna et al. in prep.).

The Impact of the Stellar-to-Halo Mass Relation

In this chapter, we have considered the growth and merger rates for haloes. In contrast, observational studies of galaxy-scale mergers measure these rates as a function of luminosity or stellar mass. The stellar-to-halo mass relation (SHMR) is fairly well constrained from a variety of observations (e.g. Behroozi et al., 2019, and references therein), and has the form of a broken power-law that changes slope abruptly on group scales. Since the halo merger rate is close to a single power-law in halo mass ratio ξ, the shape of the SHMR should produce a kink in the merger rate measured as a function of *stellar* mass ratio.

To illustrate this effect, we approximate the halo merger rate shown in Fig. 3.1 as a power law $B/n \sim R_0 \xi^{-1.66} M_{h,1}^{0.13}$. We then use the SHMR of (Behroozi et al., 2019) to convert halo masses and mass intervals to stellar masses and mass intervals. Fig. 3.12

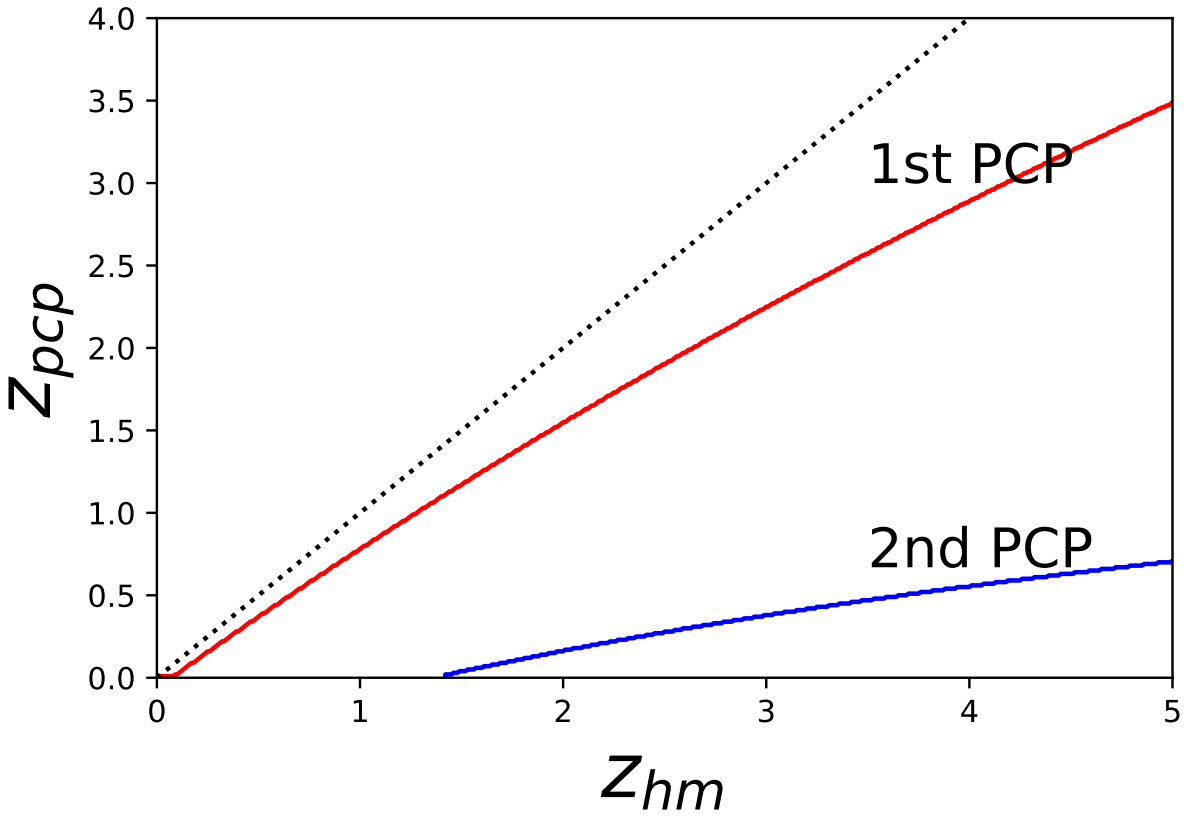

Figure 3.11: Redshift at which first (top curve) and second (bottom curve) pericentric passages would occur, assuming a halo merger at z_{hm} and conservation of orbital properties. The dotted line shows a 1-1 correspondence for reference.

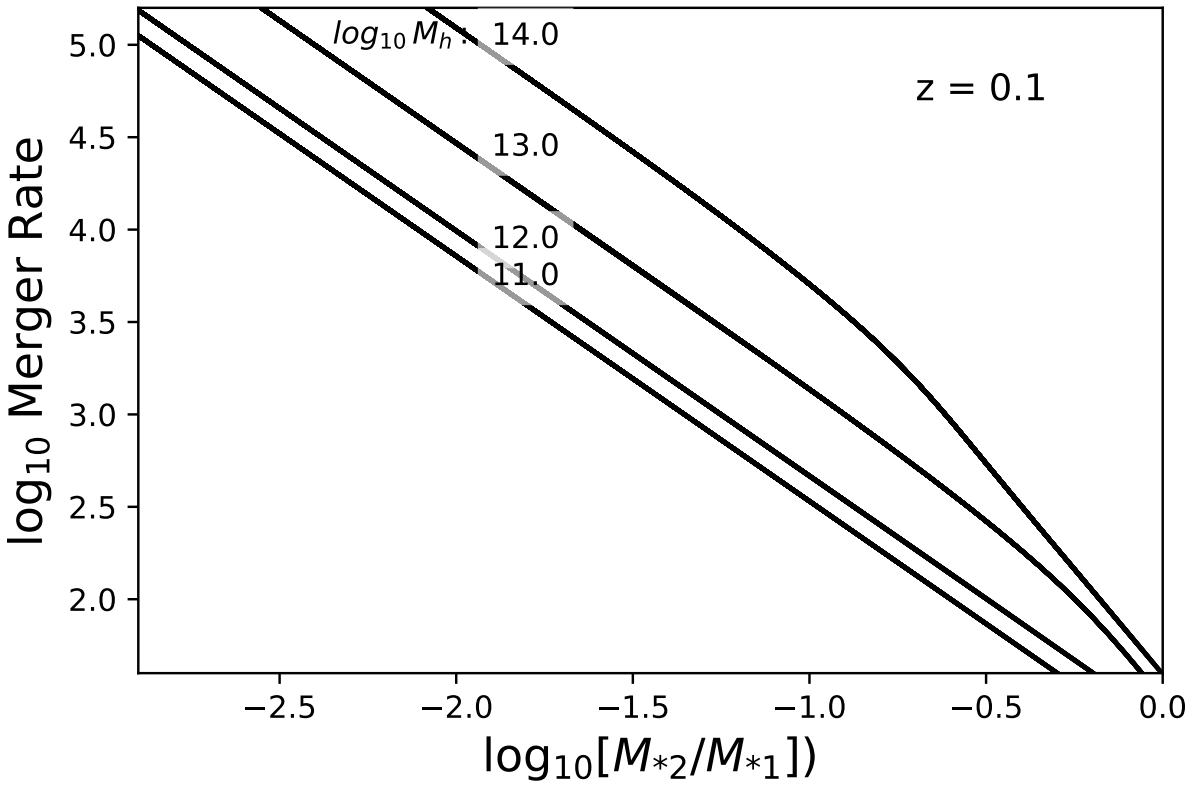

Figure 3.12: Merger rate onto central galaxies in haloes of the mass indicated, as a function of stellar mass ratio, at z=0.1. Note the feature in the merger rate on group and cluster scales.

shows how the merger rate is expected to vary with stellar mass ratio, for galaxies merging into systems with various primary halo masses. While the merger rate onto galaxy-mass haloes retains a simple power-law form, on group and cluster scales, the kink in the SHMR appears as a change in the slope of the merger rate in stellar mass unites. This feature might be observable when recording the rate of group or galaxy-scale accretion onto clusters.

Uncertainties in the Galaxy Merger Rate

Having taken into consideration the complications discussed above in relating the galaxy merger rate to the halo merger rate, there remains the question of how accurately the latter can be determined. Galaxy mergers can be detected by either looking at objects

that are very likely to merge, such as close pairs, or objects which exhibit recent evidence for merger activity, such as tidal features.

Recent measurements of the galaxy close pair fraction in particular show that the scatter between different studies is significantly reduced if selection criteria are closely matched. Mundy et al. (2017), Mantha et al. (2018) and Duncan et al. (2019), for instance, find similar trends in the merger rate as a function of redshift and stellar mass ratios ξ, with a scatter that is about a factor of 2–3. The combined sample also matches theoretical predictions from the Illustris hydrodynamical simulations (Rodriguez-Gomez et al., 2015) at about this level. While this precision may improve with future work, including machine-learning (ML) approaches to identifying merging systems (e.g. Goulding et al., 2018; Ackermann et al., 2018; Bottrell et al., 2019; Martin et al., 2020), the current uncertainty significantly exceeds our target accuracy of 10%. We conclude that galaxy-scale mergers, although abundant and intrinsically interesting for the study of galaxy evolution, are unlikely to produce useful cosmological constraints.

3.6 Summary and Conclusions

Tensions between current cosmological results at high and low redshift, as well as the flood of data on low-redshift clusters and galaxies expected from forthcoming surveys, encourage us to consider new methods for constraining cosmological parameters, based on non-linear structure formation and halo properties. In recent work, we found that measurements of the overall dynamical age of clusters via structural proxies such as concentration might provide quite sensitive constraints on the parameters Ω_m and σ_8. This is in part because, over a reasonable range of halo mass and redshfit, the degeneracy direction for age is almost orthogonal to the direction for abundance, and thus age constraints are very complimentary to abundance constraints.

In this chapter, we have considered instead the instantaneous growth rate of haloes, as determined either from the halo merger rate, or through measures of overall accretion within the preceding dynamical time. Estimating these rates analytically, we find that halo merger rates, average growth rates, and the fraction of systems with significant recent growth (the LGS fraction) should all have slightly different dependence on the cosmological parameters, but should also be complimentary to abundance-based constraints.

Measuring merger and growth rates in a number of different numerical simulations, we find trends similar to the analytic predictions, but do not confirm all of these exactly. Further work is needed here, to understand how mass resolution, snapshot cadence and

the merger tree algorithm affect the results. Assuming the analytic predictions are correct, however, a measurement of the halo merger rate or the LGS fraction with an accuracy of $\sim$10% would be required to distinguish between cosmologies with $\sigma_8 = 0.7$ and $\sigma_8 = 0.9$.

Finally, we have considered several different paths to obtaining accurate measurements of the merger or growth rates observationally. On cluster scales, counting individual mergers may result in large uncertainties related to completeness, so a target of 10% seems optimistic. On the other hand, a measurement of the LGS fraction seems more feasible, since clusters with recent episodes of significant growth should be easy to identify. Galaxy mergers provide a completely different path to determining the merger rate. There are several complications here, however, including the offset between halo merger times and galaxy merger times, and scatter in the relation between halo mass and stellar mass. Given current uncertainties in the galaxy merger rate, our target accuracy seems unrealistic on galaxy scales, although the galaxy merger rate remains extremely interesting for other reasons.

Considering these results together with those of Chapter 2, we conclude that structural studies of galaxy clusters provide several promising avenues for constraining cosmological parameters. The distribution of cluster concentration parameters or projected shapes, or the prevalence of disturbed clusters, should all provide tests of the cosmological model complimentary to those already in use. We will continue to explore this possibility in future work.

Chapter 4

What structural properties tell us about halo age

4.1 Abstract

We have seen in Chapters 2 and 3 that halo age and growth rate can be used to constrain cosmology, and because halo age is expected to correlate with structural properties, we can use the observational measurement of structural properties to constrain cosmology directly. In this chapter, I investigate the relationship between the formation history of dark matter haloes and their structural properties. Using a set of high-resolution simulations, I track the history of haloes across 100 snapshots to measure growth history quantities and structural properties. Most haloes' mass accretion histories (MAHs) differ along one dimension, which is associated with the halo general "age". I also identify different indicators of halo age and assess their connection to halo structure. I investigate which age indicator correlates with the internal structure of haloes, finding that the redshift at which haloes amassed half their mass (z_{50}) and when they had their last major merger ($z_{\rm mm30}$) have the most significant correlation with structural properties, specifically concentration and the offset between the centre-of-mass and the density peak. Combining different structural probes can efficiently reduce the scatter in z_{50} predictions. This work highlights the potential of using the relationship between halo age and structure to constrain cosmological models.

4.2 Introduction

Because dark matter haloes grow hierarchically through merger and smooth accretion, we expect this process to leave imprints on their structures. For example, since the central parts of haloes are likely to be accreted earlier, they are expected to be denser as a reflection of the density of the Universe when those parts were accreted; hence older haloes are predicted to have larger concentrations than younger ones (Jing, 2000; Bullock et al., 2001; Wechsler et al., 2002; Gao et al., 2005; Jing et al., 2007; Correa et al., 2015c). Other structural parameters have also been shown to depend on halo age or merger history, including shape (Drakos et al., 2019a), substructure (Gao et al., 2004; Diemand et al., 2007) and centre-of-mass offset (Macciò et al., 2007; Power et al., 2012).

While the link between the halo structural parameters and age has been established and extensively studied, there is no real consensus on how to best describe halo age and how many quantities are necessary to describe the formation history fully. Wong & Taylor (2012) attempted to answer this question by using a non-parametric Principal Component Analysis (PCA) on the mass accretion history (MAH) and found that more than half of the total variance between haloes was happening along one axis and 80 per cent across two, suggesting that the MAH is a two-dimensional quantity. In this chapter, I attempt the same type of analysis with a larger sample of haloes at higher masses and with several other indicators to study the nature of the halo age and dynamical state in more detail.

Other halo properties have been subject to similar analyses by Jeeson-Daniel et al. (2011); Skibba & Macciò (2011) and Wong & Taylor (2012) and found that much of the variance is accounted for by one axis along concentration and mass and a second along the relaxedness of the halo. They also found a strong correlation between age and concentration.

Building upon these analyses, one can look at which structural properties correlate with age and whether they can inform about different populations of haloes with different ages. More specifically, I want to try to answer how to predict halo age with structural properties reliably and accurately. This would allow us to link observable structural properties to Ω_m and σ_8, since we have seen in Chapters 2 and 3 that age and growth rate are sensitive to cosmology.

Because we can measure halo structural properties such as shape and concentration from observations, successfully linking those properties to halo or cluster age would allow for constraints on Ω_m and σ_8 with cluster structure observations given the results shown in Chapter 2. In practice, we might eventually be able to go directly from structure to cosmological constraints using simulations and mock observations to calibrate the theoretical

relationship inferred from age.

The outline of this chapter is as follows. In Section 4.3, I describe the simulations used and different halo age and structural properties that I will study and give their definitions. In Section 4.4, I attempt to find the best way to quantify the halo formation history while losing the least amount of information and then see which structural properties correlate with these parameters in Section 4.5. Finally, I summarize my findings and conclude in Section 4.6.

4.3 Halo properties

4.3.1 Simulations

To study the link between halo age parameters and structural properties, I have run a set of more than 20 high-resolution DM-only N-body simulations with $N = 1024^3$ particles, each with different values of Ω_m and σ_8 and all other parameters (cosmological and simulation) fixed. They have large enough box sizes ($L = 500$Mpc/h) to have enough cluster- and group-sized haloes and enough snapshots ($N_{\rm snap} = 119$), to track the history of DM haloes up to $z = 5$. All simulations have been run using the publicly available code GADGET-4[1] (Springel et al., 2021). They all have a softening length $l_s = 2.5$kpc, a particle mass of $m_p = 9.69 \times 10^9 M_\odot$, a Hubble parameter $H_0 = 70$km/s/Mpc and initial spectral index $n_s = 0.965$. Most of this chapter's analysis was done with one of the simulations with $\Omega_m = 0.3$ and $\sigma_8 = 0.85$. I have repeated the analysis with simulations with different Ω_m and σ_8 and found similar results.

I have run the Amiga Halo Finder (AHF) (Gill et al., 2004; Knollmann & Knebe, 2009) on each simulation and snapshot, together with the AHF merger tree algorithm. Each halo has a minimum of 100 particles, which sets the minimum mass of haloes to $M > 9.69 \times 10^{11} M_\odot$. This limit is set to have reliable structural property measurements, but it is arguably large and will introduce a loss of information about low-mass haloes' early growth because they will quickly reach the minimum halo mass and either stay there or disappear. This is a compromise between the desire to have a large enough sample of group and cluster mass haloes, limiting poor structural parameter fits and being able to track a halo's history far enough into the past.

Given a merger tree algorithm and halo mass definition, in our case, $M_{\rm halo} \equiv M_{\rm 200c}$ where $M_{\rm 200c}$ is the mass inside a sphere defined by the radius where the density drops

[1] https://wwwmpa.mpa-garching.mpg.de/gadget4/

below 200 times the critical density ρ_c, we can track the mass history of any given halo $M(z)$. We define a halo's mass accretion history (MAH) as $\mathcal{M}(z) \equiv M(z)/M(z=0)$. Note that because of the nature of Amiga Halo Finder and merger tree algorithm, the MAH is not monotonic; a halo can, and will often, undergo mass loss.

4.3.2 Halo age

After they collapse into gravitationally bound structures, haloes grow their masses in two main ways, slow smooth accretion and rapid growth through mergers. One can get information about the accretion phases and merger events by analysing the halo's mass accretion history $\mathcal{M}(z)$. Still, it is not obvious how to define a halo's "age" or formation time. Different phases of halo growth could, in principle, be more critical than others or encompass more age information. To explore a wide range of age parameters, I chose to use three categories of halo age indicators :

- redshift at which the halo had amassed a fraction x of its z_0 mass, where x can be 90%, 75%, 50%, 10% and 1%. These will be referred to as z_{90}, z_{75}, z_{50}, z_{10}, z_1 respectively.
- last redshift at which the halo has grown by a given x fraction where x here can be 30%, 25%, 20% and 10%, referred to as z_{mm30}, z_{mm25}, z_{mm20}, z_{mm10}. The growth here is measured with respect to the initial mass, specifically, where $\frac{M(z_0)-M(z_i)}{M(z_i)} > x$ and $z_i > z_0$. The maximum initial redshift z_i considered is taken to be early enough for mergers to be fully accounted for in the host halo's mass; I found that five snapshots at low z are usually enough.
- fraction of z=0 mass a halo had at a given redshift z. This is, in practice, a sampling of the MAH $\mathcal{M}(z)$ in a few redshifts: z=0.1, 0.3, 1, 2, 3.
- the parameters γ and β from the McBride et al. (2009) fitting function : $M(z) = M_0(1+z)^{\beta}e^{-\gamma z}$. Since they are degenerate, I use the logarithmic growth rate at z=0, $-\frac{d\mathcal{M}}{dz}|_{z=0} = \gamma - \beta$ as the age indicator.

4.3.3 Structural properties

Dark matter haloes are complex objects which exhibit a variety of features, many of which result from the more-or-less recent growth and merger history of the halo. I looked at

a selection of structural properties to study how they link to the different age quantities described in 4.3.2, some of which were measured directly by AHF:

- cNFW: concentration parameter $c = r_{vir}/r_s$ where r_s is the scale radius of a fitted Navarro-Frenk-White (NFW) density profile.
- b: modelling the halo as an ellipsoid, ratio of its second to first principal axes.
- c: ratio of the third to first principal axes.
- λ: spin parameter as defined by Bullock et al. (2001).
- com_offset: offset, in Mpc, between the centre-of-mass of the halo and the halo centre defined as the peak of the density contours. Note that normalising by the halo's virial radius does not change correlations with age.
- Elongation: $E \equiv (b^2 + c^2)/2$.
- Triaxiality: $T \equiv \frac{1-b^2}{1-c^2}$.
- $V_{\rm rat}$: the ratio of the maximum circular velocity and the virial velocity $V_{\rm rat} = \frac{V_{\rm max}}{V_{\rm vir}}$.

Projected quantities were also either fitted or calculated for more direct comparison with observables:

- Conc: 2D NFW concentration where I used a maximum likelihood fit using particle positions directly, without binning. This typically gives more robust and accurate fits.
- χ^2_ρ: the χ^2 of the 2D density profile fit calculated as

$$\chi^2 \equiv \frac{1}{M} \sum_i \frac{(\rho_{NFW}(r_i) - \rho(r_i))^2}{\sigma_i^2}$$

  where σ_i is taken to be Poissonian errors.
- χ^2_M: same as above, for the mass profile.
- axis ratio: the halo's minor to major axis ratio; modelled as an ellipse. I calculate the axis ratio by finding the eigenvalues and eigenvectors of the reduced inertia tensor, then rotate the particle coordinates towards the eigenvector axes and iterate the process until convergence.

- 2D com offset: the offset (in Mpc) between the centre-of-mass and centre of the halo in a projected 2D plane.

These projected quantities can all be measured through observations, as well as $V_{\rm rat}$. The 3D real space quantities, such as Triaxiality and Elongation, are included to study their behaviour with halo age in general and understand aspects of halo dynamical state on a more fundamental level.

4.4 Quantifying the halo history

4.4.1 What information is contained in the mass accretion history ?

A halo's mass accretion history contains all the information about its growth journey. It is worth questioning how many aspects of the MAH differentiate haloes from each other. In the simplest case, all MAHs are self-similar and determined by the halo mass, or it could be a more complex function characterized by growth at different epochs. To assess the dimensionality of the MAH, we can use a Principal Component Analysis (PCA) (Pearson, 1901).

PCA is a change of basis in variable space to a new orthogonal basis where each of the new axes, called principal components (PCs), is in a direction that captures the most variance in the dataset. It is a standard dimensionality reduction method since it gives information about the relevant variables that hold the most information. It can be used to know the variables that capture the same information, the irrelevant ones, and those that are independent.

In our case, PCA is useful to assess the intrinsic dimensionality of the MAH; in other words, how many parameters are needed to fully describe a halo's MAH and what combinations of epochs can describe these parameters.

I used a selection of 38,247 haloes with masses $M > 10^{13} M_{\odot}/h$, and found that the MAH is very close to a one-dimensional quantity. As shown in Fig. 4.1, 60% of the total variance among halo MAHs occurs along a unique axis in the MAH space, similar to what Wong & Taylor (2012) found where the first PC accounted for 58.4% of the total variance. Other axes have minor contributions to the total variance at around 10%, which differs from Wong & Taylor (2012) which found that the second axis accounted for 25% of the total. More interestingly, the first axis is a global linear combination of the values of the MAH at

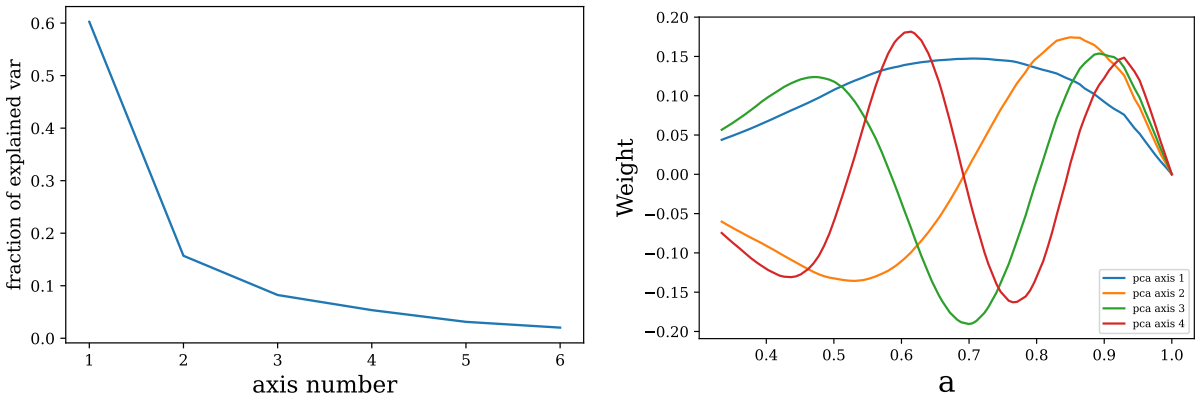

Figure 4.1: Left: the fraction of the total explained variance along each principal component. Note that more than 60% of the total variance is along the first axis, significantly more than all others. Most of the MAH information is contained along that axis. Right: weights of each of the first four PCA axes. The first one is a weighted sum of the MAH at all snapshots, with slightly more importance given to epochs that are not too recent or too early, corresponding to $0.25 < z < 0.65$. Other principal axes show Fourier-like decomposition, distinguishing between haloes by the number of fast growth periods they had and whether those periods happened at early or late times.

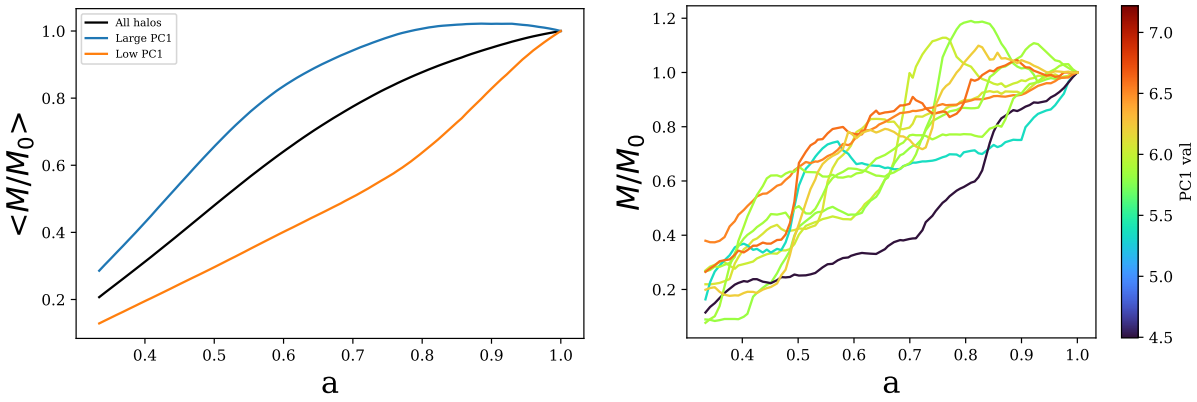

Figure 4.2: Left: the mean of the MAHs for all haloes (black), haloes with PC1 values larger than half a standard deviation (blue) from the mean, and haloes with PC1 values lower than half a standard deviation (orange) from the mean. We can see that large PC1 haloes have grown early, stopped, and lost mass in some instances, while low PC1 haloes have grown recently. Right: mass accretion histories from a random selection of haloes colour coded by the value of their PC1.

all snapshots, with slightly more emphasis on the halo's middle age between $0.6 < a < 0.8$ or $0.25 < z < 0.65$. The first principal component is the closest to a halo "age" measure one can have. Other principal components show Fourier-like behaviour, capturing different merger/growth frequencies, which might be linked to different halo environments.

This is further illustrated when we look at the average of the MAHs for haloes having low and large PC1 values as shown in Fig. 4.2, it is clear that large PC1 haloes have not had any recent growth, and most their mass was amassed before a=0.7, while haloes with low PC1 had most their growth recently. This behaviour is almost identical to the one found in Wong & Taylor (2012), showing that the first PC in both cases are the same physical quantity. This shows that PC1 is an excellent "age" quantifier and that the halo states at different times convey equivalent amounts of information, making a weighted sum of all of them the best way to know when a halo had most of its growth. In practice, this is just because old haloes have had larger masses for longer, so a sum of their relative mass throughout their history is going to be mechanically larger than young haloes that had lower relative masses throughout their history.

Other subsequent principal components separate haloes according to their growth phase at z=0 for each subsequent growth frequency, as shown in Fig. 4.3. They differentiate whether the most recent halo growth phase was slow or fast in cases where haloes have

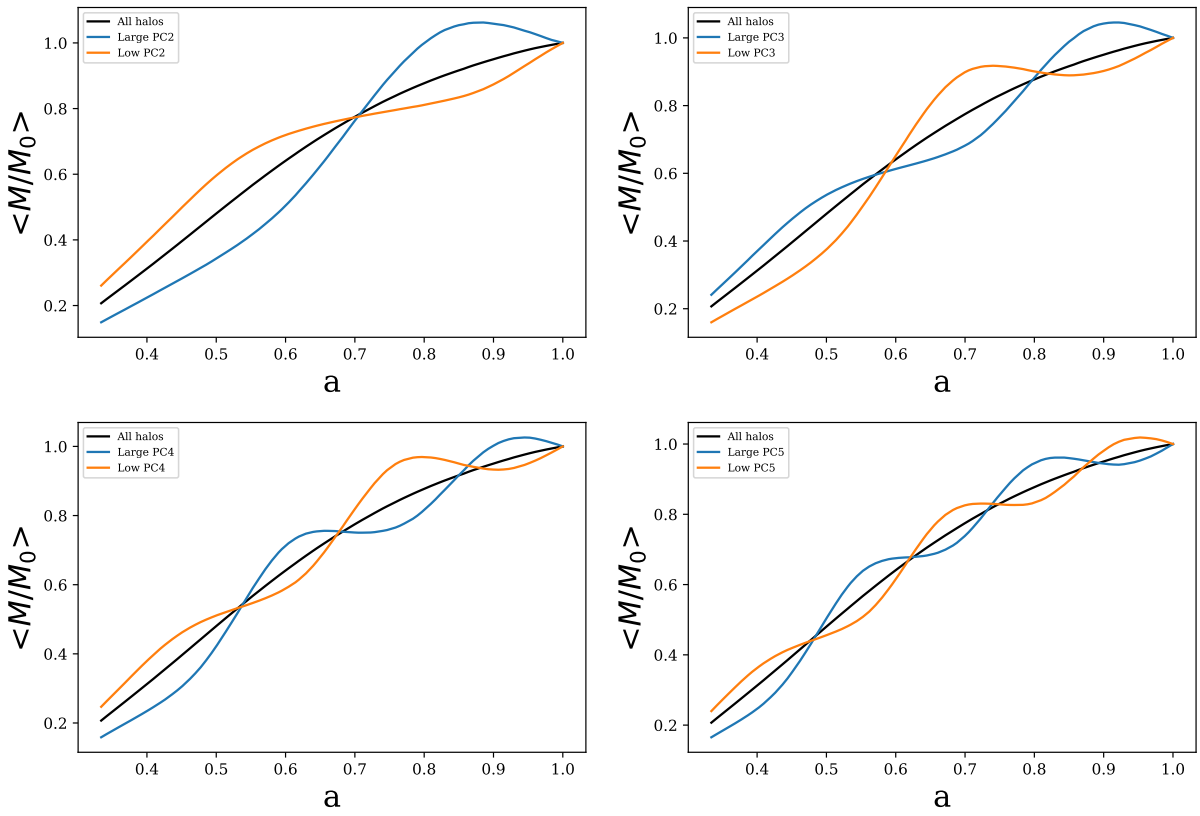

Figure 4.3: Difference of the mean MAH between haloes with lower principal component values and haloes with larger ones for each PC2, PC3, PC4 and PC5. Each PC contains information about halo growth at different frequencies.

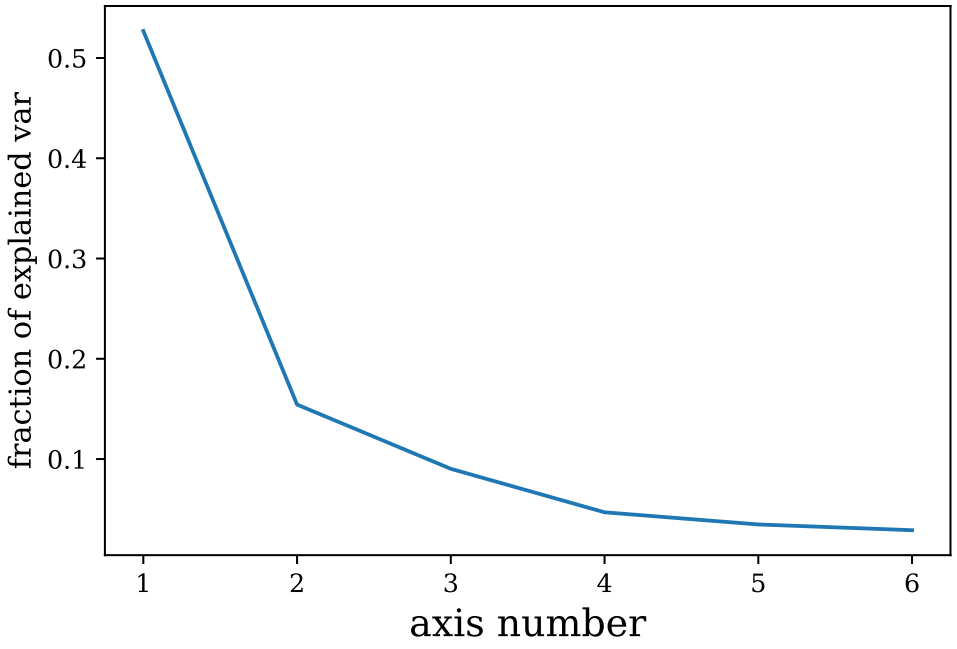

Figure 4.4: The fraction of variance that occurs along each of the first six axes of a PCA performed on the halo age indicators defined in Section 4.3.2. The results are similar to Fig 4.1 where the PCA was performed on the MAH, indicating that the same type of information is carried through the collection of age indicators and the MAH.

had alternating fast and slow growth phases.

4.4.2 What are the best indicators of halo "age"?

A variety of quantities can be used as a halo "age". In Chapter 2, I chose z_{50}, but one can argue for the use of any of the quantities described in Section 4.3.2 as an age indicator. We can try to perform the same type of analysis as in Section 4.4.1 with all quantities described in Section 4.3.2 to check which combination of indicators contains the most information about halo history.

First, we can perform a PCA with all age indicators defined in Section 4.3.2 which are the redshifts where the halo had reached a fraction x of its mass z_x, the redshift of the last major increase in mass z_{mmx}, the fraction of current halo mass at different redshifts and $\gamma - \beta$ the logarithmic growth rate at $z = 0$. With the addition of the first principal

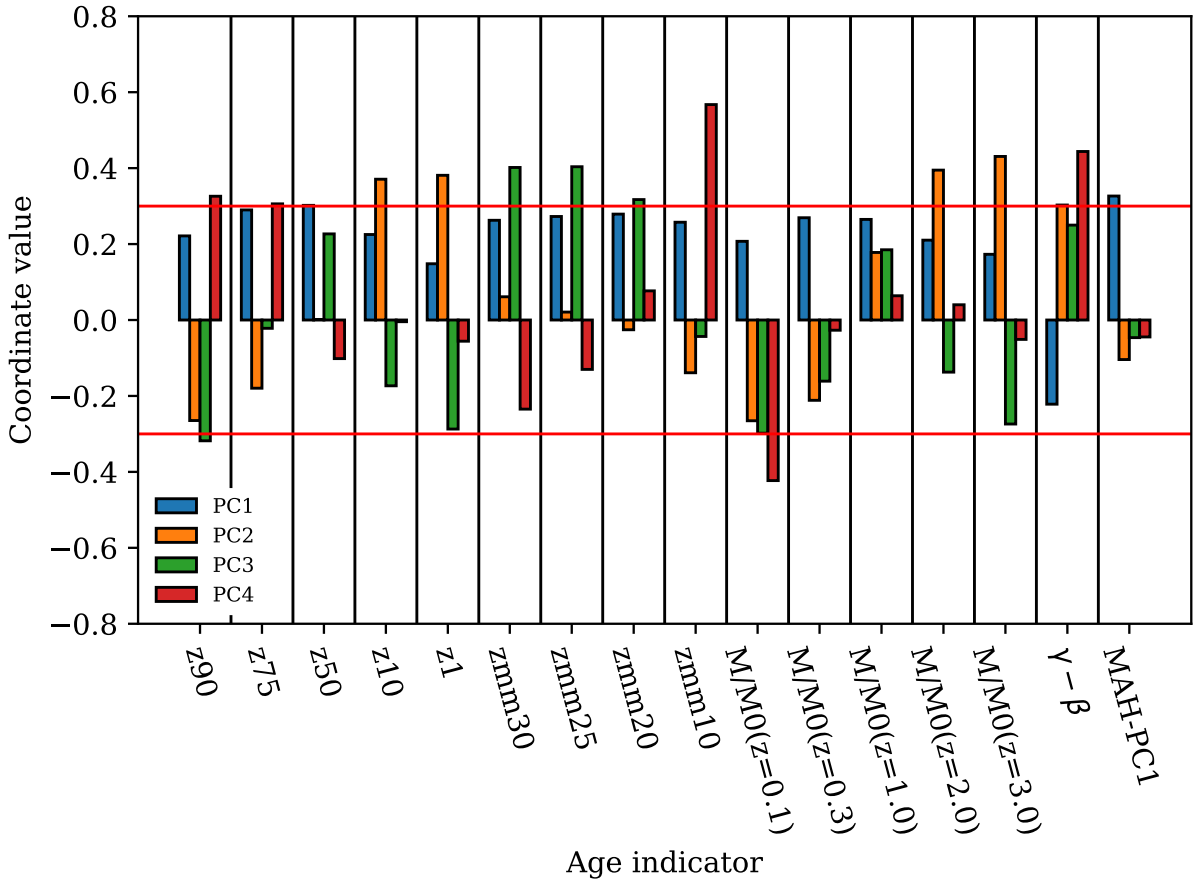

Figure 4.5: The linear decomposition of each of the first four PCs in terms of different age indicators. The red line is a threshold at 0.3 used to indicate which property is significant to different PCs. The first PC, responsible for more than 50% of the variance, is correlated with all indicators. Other axes capture early or late growth information.

component of the MAH found in Section 4.4.1 (MAH-PC1). It is clear that a lot of these convey very similar information but the Principal Component Analysis will naturally show which ones do and to what extent they do it.

The fraction of the explained variance is shown in Fig. 4.4. Almost half the total variance happens along one axis; the second and third principal component account for a significant part of the variance, while other subsequent axes are less important.

The weights of each property on the first four PCs are shown in the right panel of Fig. 4.5. Most age indicators are correlated with the first PC, which suggests that they are almost equivalent indicators of age; notably, the first principal component of the MAH (MAH-PC1) has the largest weight in PC1, confirming that a weighted average of a halo MAH is a good formation history indicator. The second PC correlates with early-time indicators and anti-correlates with late-time indicators; it carries information about how much the halo has grown in the early stage of its formation history. The third PC is correlated with the redshift of the last major merger and is broadly a measure of the very recent state of the halo, as shown by the fact that it anti-correlates with z_{90} and $\mathcal{M}$ at z=0.1. Finally, the fourth PC is also linked to very recent history but mostly captures slow growth and minor merger information exclusively near z=0.

Given that many of the variables used above are redundant, we can try to do the same analysis with a subset of them: z_{50}, z_1, z_{mm30}, $M/M_0(z=0.1))$, $M/M_0(z=3))$, $\gamma-\beta$ and MAH-PC1.

We see from Fig. 4.6 that the first PC captures the same amount of variance and represents the same first axis as previously. The second and third PCs capture more variance, while subsequent axes are less important.

From Fig. 4.7, we observe that the first PC is roughly the same as previously, with a sign difference, which does not change fundamentally the information conveyed through the axis. Notably, z_{50}, z_{mm30}, and the MAH-PC1 stand out as good overall formation time estimates. The second PC captures information about the very early-time formation history, while the third PC captures the very recent history.

The MAHs of haloes with different values of PC1 are shown in Fig. 4.8; it is a visual confirmation that haloes with high PC1 tend to reach half their mass at lower z, that they had a 1–3 major merger more recently.

Combining two age indicators is usually enough to describe a halo's formation history without losing too much information. The most important is a global age indicating when most of the halo's growth happened, typically when it went from 10% to 90% of its current mass. z_{50} is a good indicator for this period, so is z_{mm30}; although, z_{75}, z_{mm25}, z_{mm20} will

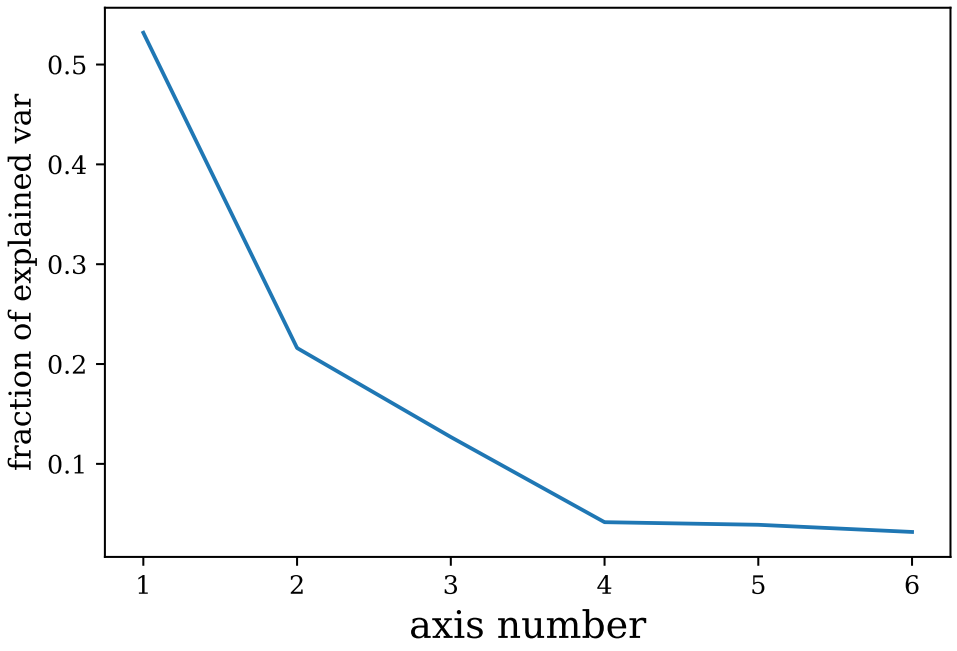

Figure 4.6: Same as in Fig. 4.4 with a subset of the indicators.

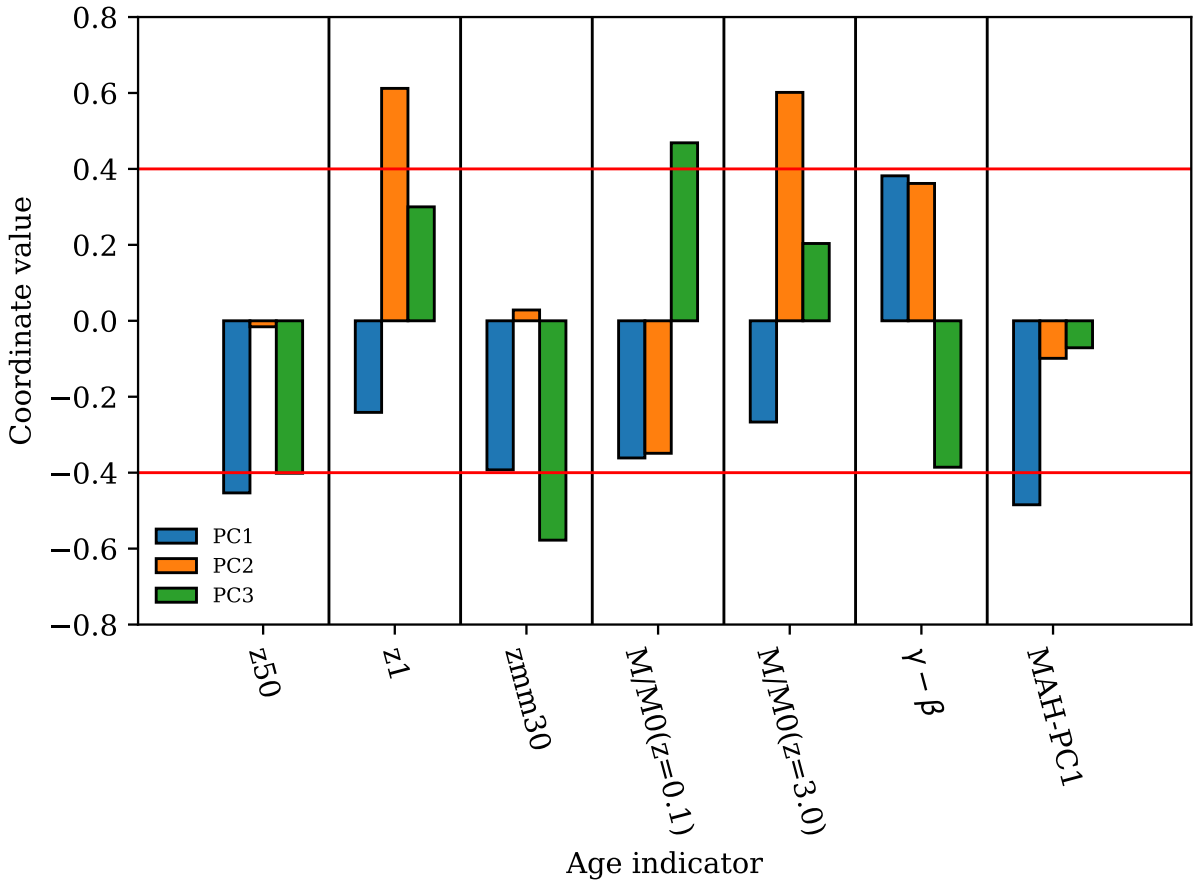

Figure 4.7: Same as in Fig. 4.5 with a subset of the indicators for the first three PCs.

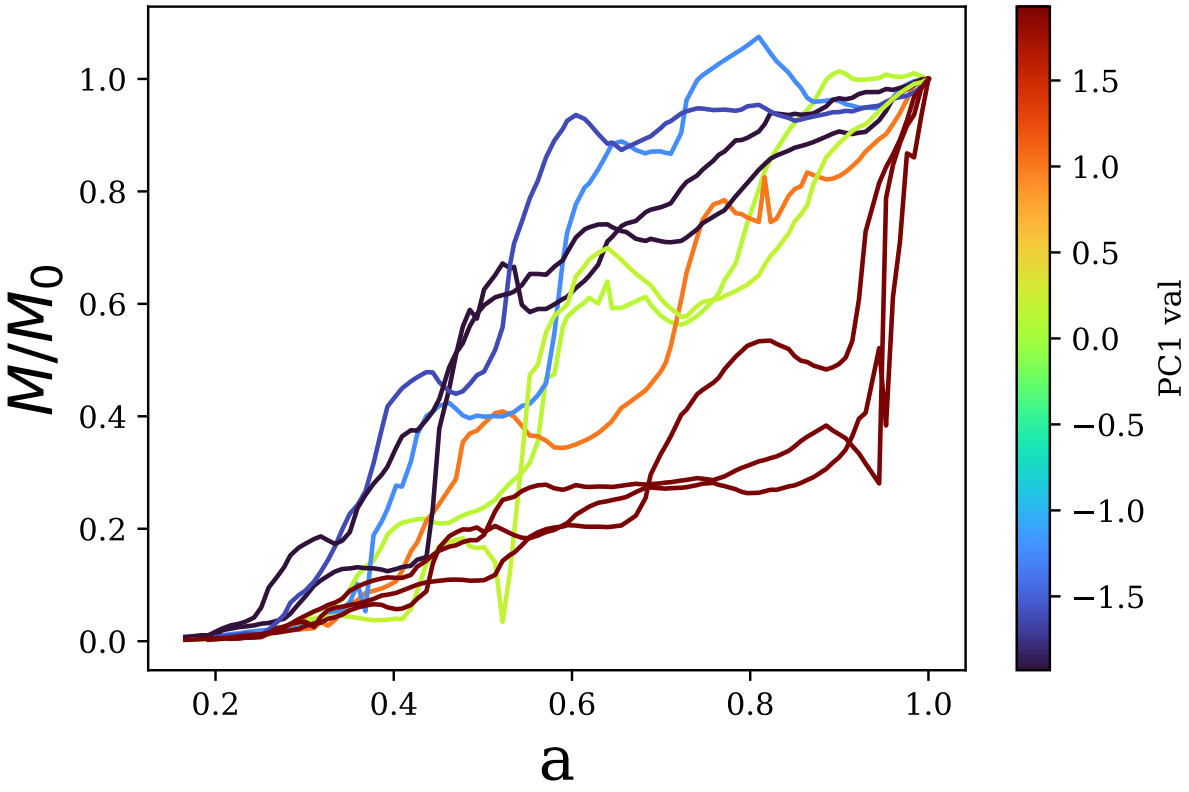

Figure 4.8: The MAH of a random selection of haloes colour-coded by the value of the first PC of the subset of age indicators.

carry the same information. The second complementary formation history indicator is the growth rate at very early or late times. We have seen that the redshift at which the halo has amassed any of its ten first per cent of mass, e.g. z_1 or z_{10}, provides information about the early stages of the formation history and can segregate haloes with similar "global age" by their early time growth. Additional information can be provided by indicators of the growth rate in very late stages, such as $\gamma - \beta$ fit parameters.

4.5 The dependence of structural properties on age

4.5.1 Structure and age correlations

After determining the main indicators of the haloes' formation history, we can try to see how they correlate with different aspects of their structure. I chose to look at four age indicators: z_{50}, z_{mm30}, z_{10} and $\gamma - \beta$; the first carries general age information, but major mergers might have a different, more obvious, structural impact. z_{10} carries the early formation information; I use z_{10} instead of z_1 or other early-time growth indicators to have larger halo samples and not be restricted by mass resolution effects. $\gamma - \beta$ is the logarithmic growth rate at z=0.

We can see how different structural properties correlate with the dynamical state indicators in Figs. 4.9, 4.10 and 4.11 which show the density plots, smoothed using a 2D Gaussian Kernel Density Estimator (KDE).

A first significant observation is that remarkably few structural properties correlate with the logarithmic growth rate at z=0: $\gamma - \beta$ except for the centre-of-mass offset both in 3D and in projection. This suggests that recent slow growth has little impact on the structure of the halo.

We can also note that z_{50} and z_{10} correlate with halo structure in very similar ways. How fast a halo gets its first 10% or 50% mass will have the same impact, although some correlations are more pronounced with z_{50} such as with concentration, the centre-of-mass offset and the χ^2 of the mass profile. This is consistent with our findings of Section 4.4 that the halo "age", which I associated with the first principal component, can be represented by z_f at any fraction of the mass f and they would carry equivalent amounts of information. The fact that z_{50} and z_{10} have similar impacts on structure also shows that the age-independent growth rate at early times that we associated with the second principal component in Section 4.4.2 is not correlated with halo structure.

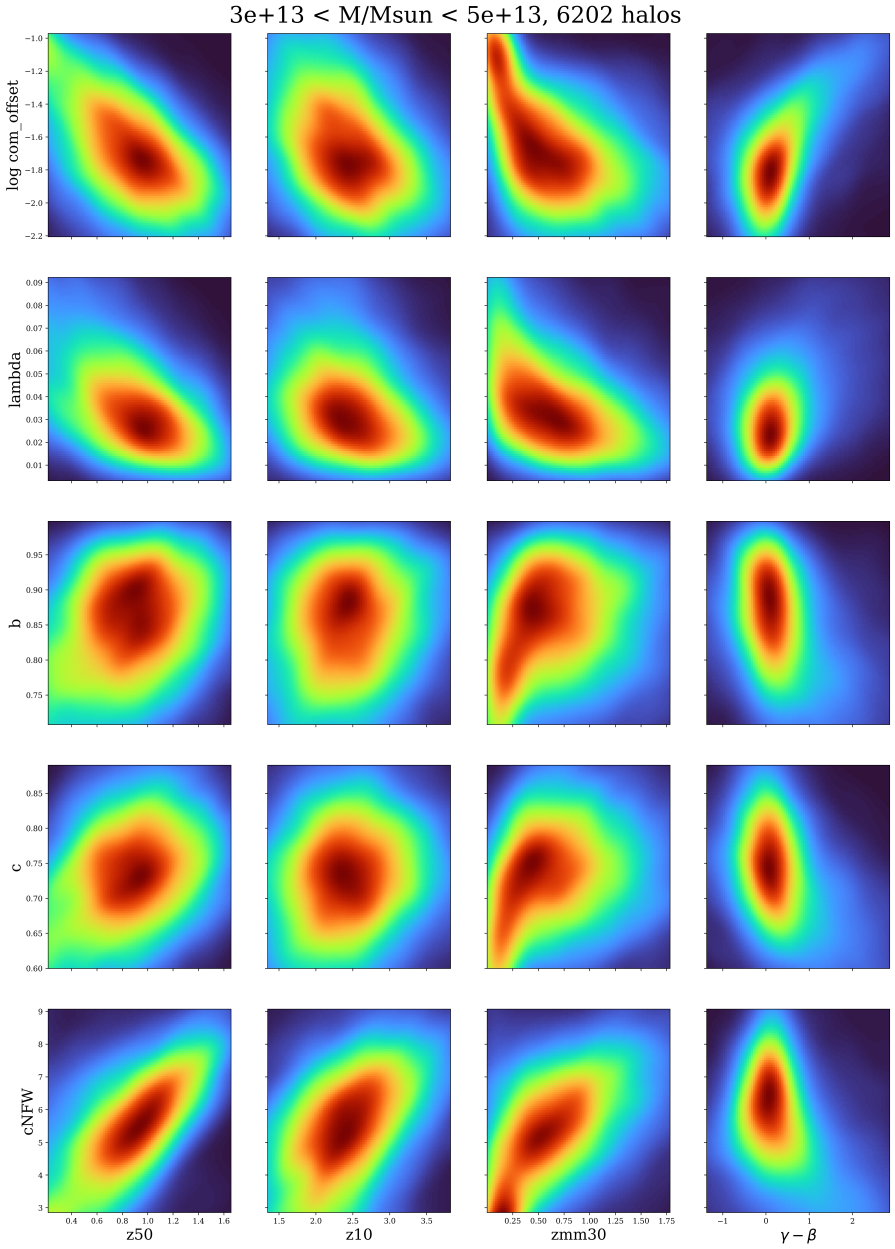

Figure 4.9: Density plots of the centre of mass offset, spin, second to first axis ratio, third to first axis ratio, and concentration taken from the Amiga Halo Finder (AHF) as a function of different dynamical state indicators: z_{50}, z_{10}, the redshift of last major merger z_{mm30}, and parameters from the McBride et al. (2009) fitting formula $\gamma - \beta$. All properties are described in Sections 4.3.3 and 4.3.2. The colour scale is linear.

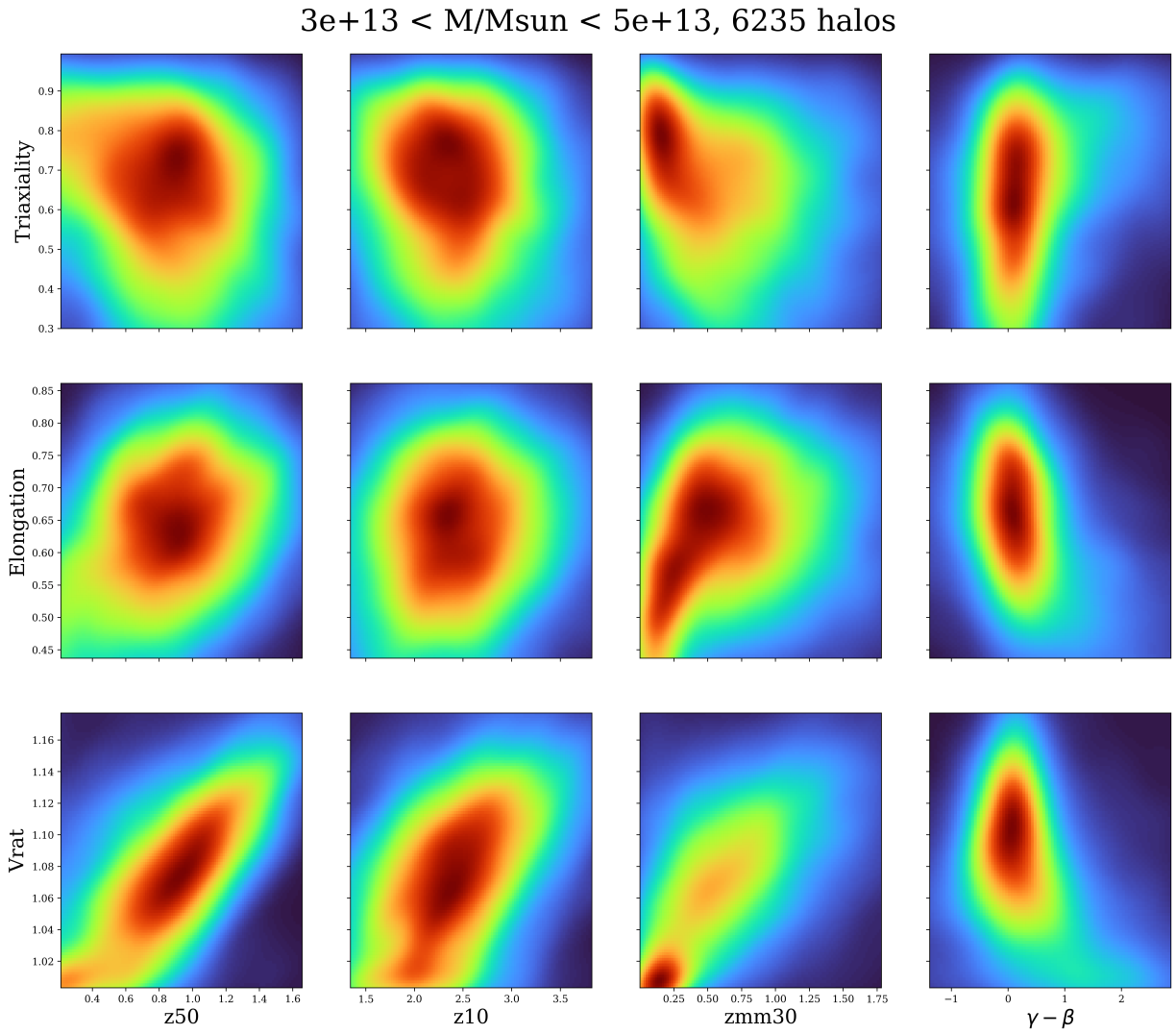

Figure 4.10: Density plots of the Triaxiality, Elongation, and ratio of $V_{\rm max}/V_{\rm vir}$ calculated from properties taken from Amiga Halo Finder (AHF) as a function of different dynamical state indicators: z_{50}, z_{10}, the redshift of last major merger z_{mm30}, and parameters from the McBride et al. (2009) fitting formula $\gamma - \beta$. All properties are described in Sections 4.3.3 and 4.3.2. The colour scale is linear.

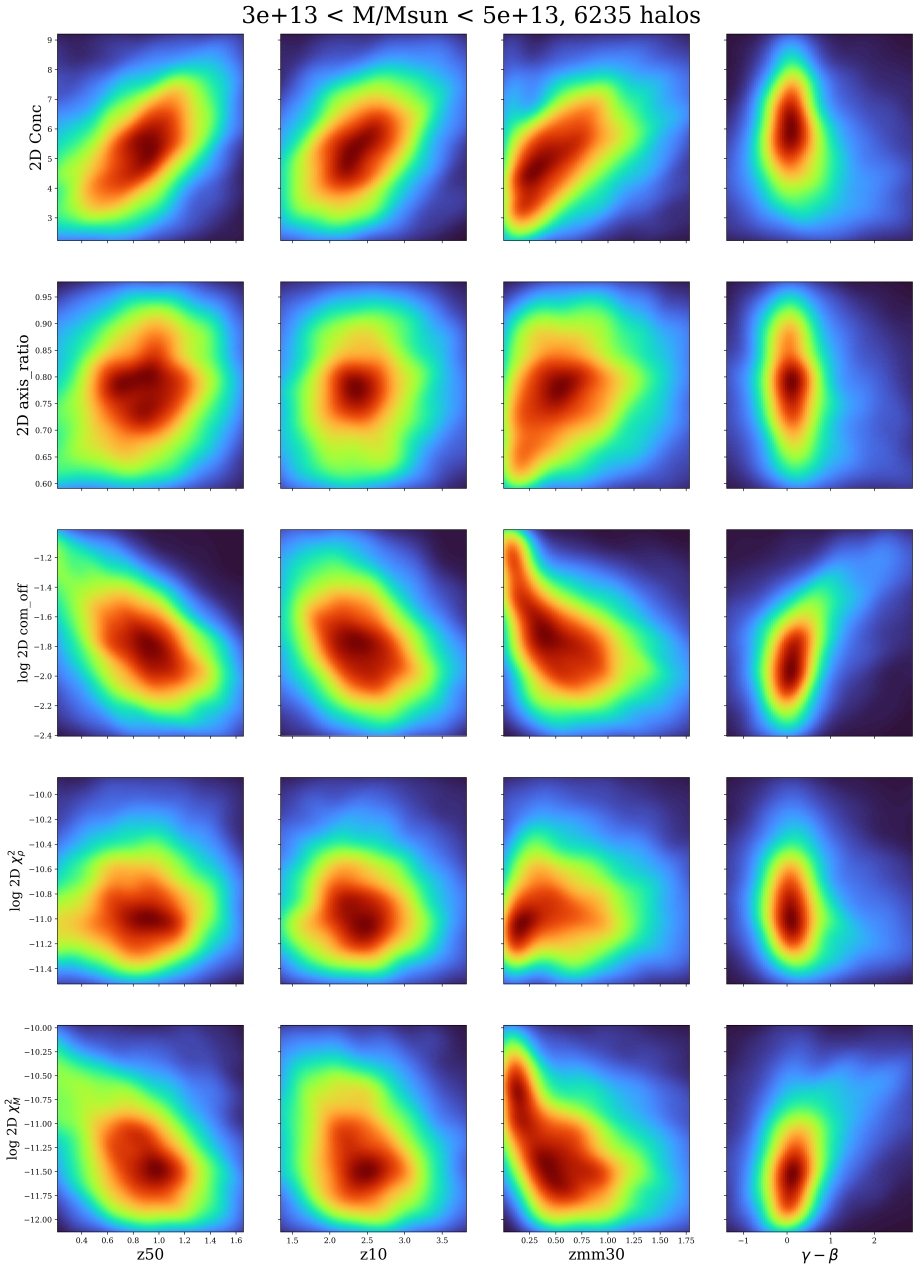

Figure 4.11: Density plots of the projected 2D concentration, axis ratio, centre of mass offset, χ^2 of the density profile and thee mass profile as a function of different dynamical state indicators: z_{50}, z_{10}, the redshift of last major merger z_{mm30}, and parameters from the McBride et al. (2009) fitting formula $\gamma - \beta$. All properties are described in Sections 4.3.3 and 4.3.2. The colour scale is linear.

One interesting feature in almost all correlations between the redshift of the last major merger z_{mm30} is the appearance of two distinct halo populations, one of which has had a major merger very recently $z_{mm30} < 0.25$. It can be seen clearly when looking at the centre-of-mass offsets (both in 3D and in projection) as well as the 3D concentration cNFW, Elongation and $V_{rat} = V_{max}/V_{vir}$. It is also present for most other properties. This low z_{mm30} halo population is almost the only one having low values of axis ratio $b < 0.8$ and $c < 0.7$; they have large centre-of-mass offsets have much lower concentrations and typically density profiles that are poorly fitted by an NFW profile. The population of haloes which had a recent major merger is the unrelaxed population and has been typically been distinguished from the relaxed haloes by looking at their centre of mass offset, virial ratio, shape and amount of substructure (Neto et al., 2007; Macciò et al., 2008; Meneghetti et al., 2014). This is similar to our findings that centre of mass offset and shape clearly distinguish between relaxed and unrelaxed objects.

Some structural properties show weak correlations with age. The halo spin λ does not seem to vary much with the halo age parameters. Similarly, halo shape indicators are weakly correlated with age parameters, except for z_{mm30}. Thus, halo shape is an excellent probe for recent mergers, as elongated haloes appear to be exclusively the result of a recent major merger.

The best halo age predictors are measures of concentration; either of the NFW concentration or V_{max}/V_{vir} show linear trends with z_{50} with limited scatter. The offset between the centre-of-mass and the density peak is also a reasonable age estimate but might be more challenging to observe.

4.5.2 Combining structural properties

Each of the different relations seen in Figs. 4.9, 4.10 and 4.11 show some level of scatter, but we do not need to restrict ourselves to one age predictor. We might expect that part of the scatter seen in one relation can be reduced by looking at another structural property, particularly if they capture different aspects of halo structure.

in Fig. 4.12. we can look at how z_{50} changes in a parameter space defined by two structural properties, specifically looking at the projected quantities that could be observable; here haloes with different z_{50} are distributed in the concentration versus axis ratio, centre-of-mass offset and χ^2_M planes at different masses.

We saw in Section 4.5.1 that the halo axis ratio, by itself, was not a good predictor of z_{50}; however, it does help when combined with concentration. In particular, for low-concentration haloes, those that are highly non-spherical are almost all very young ones.

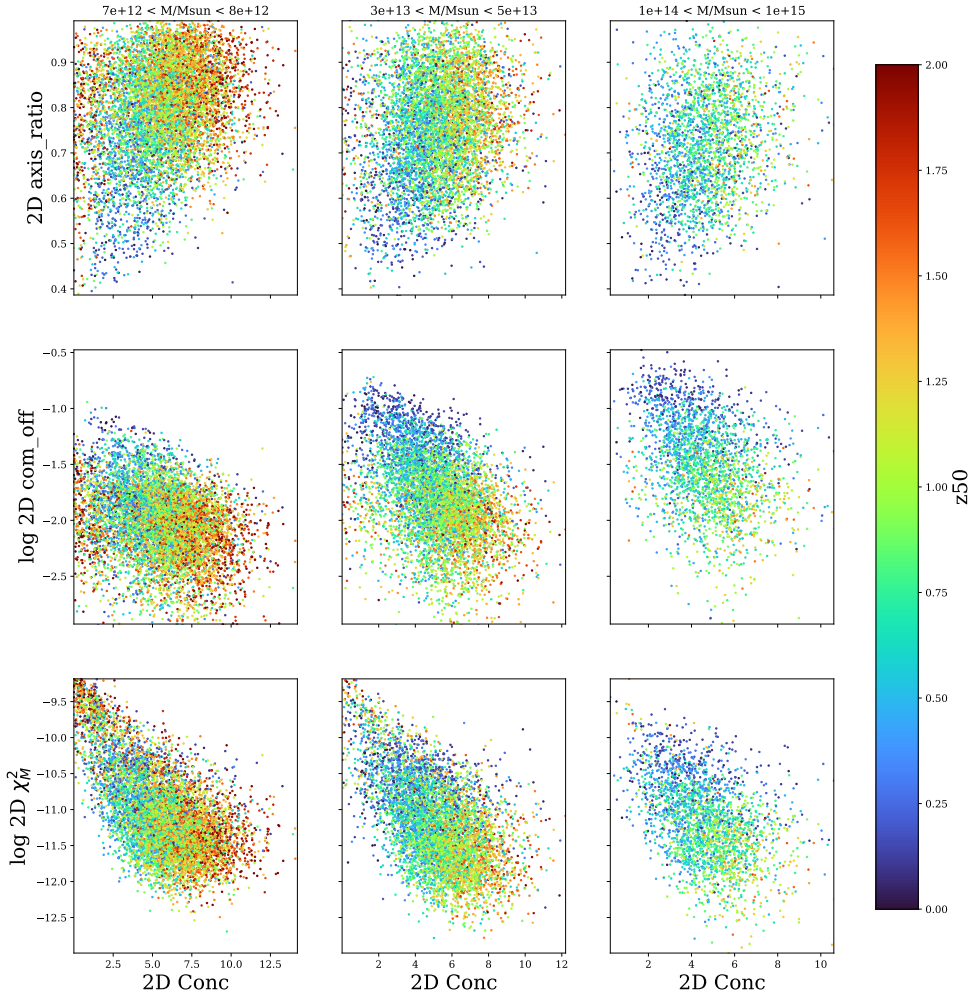

Figure 4.12: Scatter plot of concentration versus axes ratio, centre-of-mass offset and the χ^2 of the mass profile, colour-coded by z_{50} for each halo.

Although, the best combination of predictors is concentration and centre of the mass offset where there is an apparent gradient, and both concentration and centre-of-mass offset help reduce the scatter of each individual property. The scatter in z_{50} values at low and medium concentrations is significantly reduced by looking at the centre-of-mass offset, and similarly for medium values centre-of-mass offset. Another way to reduce the scatter in the concentration-z_{50} relation is to look at χ^2_M, particularly at higher masses; at any given concentration, younger haloes will have a higher χ^2_M than older ones.

In practice, several approaches can be used to estimate halo age by combining structural properties. A simple fit of the data with $z_{50} = f(\text{conc}, \text{com\_offset})$ would provide accurate enough z_{50} measurements. This is illustrated in Fig. 4.13 where I show how differentiating between haloes with a centre-of-mass offset criteria changes the $\langle z_{50} \rangle$ =f(concentration) relation, allowing more accurate estimates. A more ambitious approach is to train a machine learning regression model to predict z_{50} and z_{mm30} using all observable structural properties. This is the subject of ongoing work, and has the advantage of being able to use a larger part of the information in the data.

Note that at low mass, there is a population of haloes with very low concentration and large z_{50}. These are poorly fitted concentrations, likely due to resolution effects and insufficient particles for a good fit. This is confirmed by the fact that they have large χ^2 and only appear at low masses.

4.6 Summary and conclusion

The formation and growth of dark matter haloes in the standard ΛCDM cosmology is a complex and non-linear process, that gives rise to collapsed structures with different formation histories. Since the halo formation history varies with cosmology, one might expect that cosmology will also influence the structural properties of haloes. And samples of clusters are expected to grow substantially over the next decade, there is the potential to use this relationship between cosmology and halo internal structure to constrain cosmology.

To understand halo formation history and dynamical state and how it influences its structure, I ran a set of high-resolution simulations with a large enough box size (500Mpc/h) to have a good sample of group and cluster size haloes. I tracked the assembly of haloes across 100 snapshots to measure different growth history quantities, together with structural properties calculated at z=0.

I started by studying the dimensionality of the accretion history of dark matter haloes. To do that, I performed a Principal Component Analysis (PCA) and found that much of

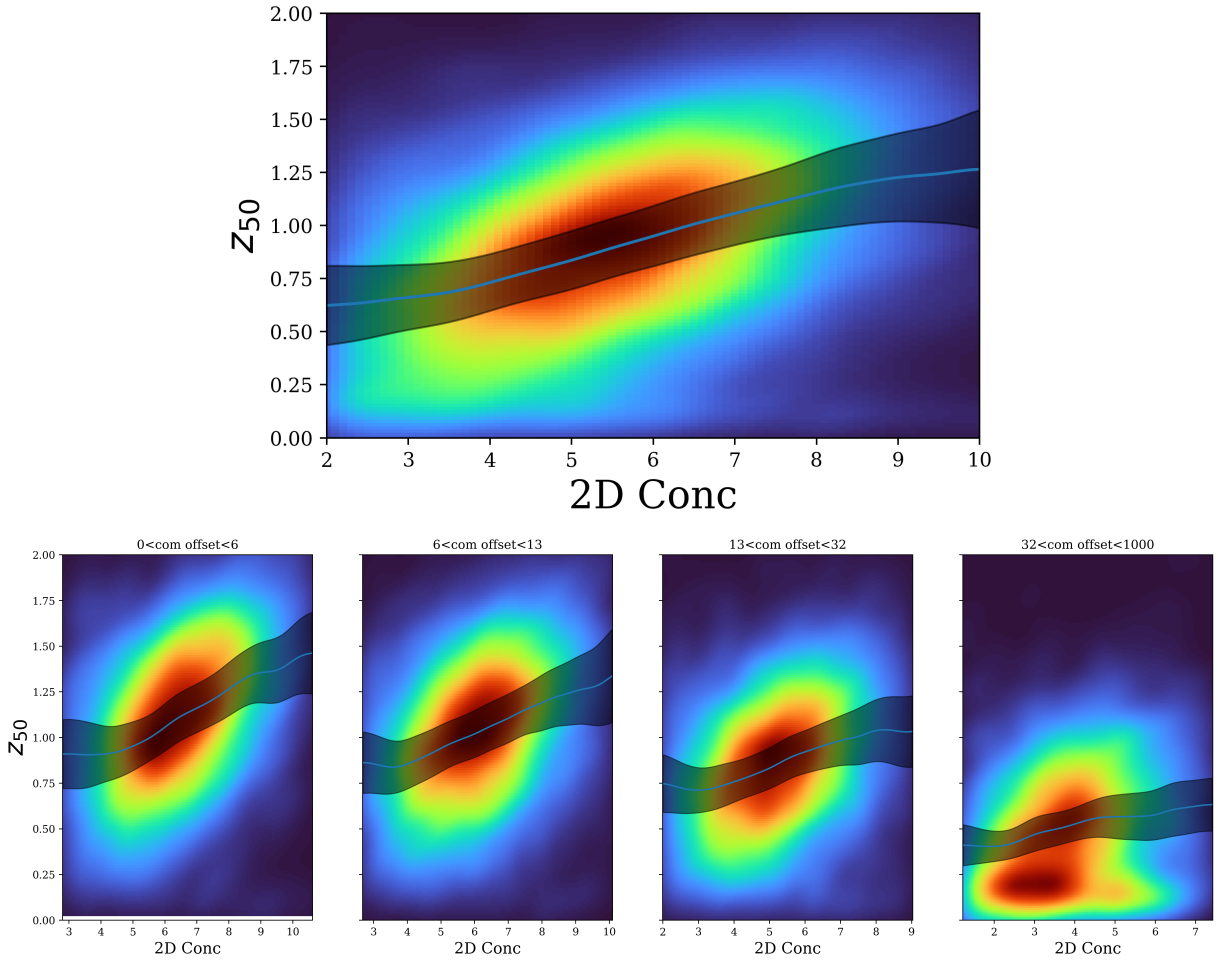

Figure 4.13: Top: density contour plot of z_{50} as a function of projected concentration for halos $M > 10^{13} M_{\odot}/h$. Bottom: the same figure, segregated by the projected centre-of-mass offset values in kpc/h. Solid lines show the average of z_{50} while the shaded region is a one standard deviation region.

the variation in halo growth histories is one-dimensional; that is, it can be described by a single parameter, which can be seen as their "age" and quantified by a weighted sum of the halo mass at all points in its history. The next step was to assess whether and which age indicators provide different information about a halo's formation history. I performed a PCA on various quantities typically used to measure haloes' age and dynamical state. I found that haloes' history differs in one major and two minor ways. The first one is a general age, well described by quantities like the redshift at which a halo has amassed a reasonable fraction of its mass, typically between 20% and 80% of its mass. The second and third way they differ is through their growth rate at very early and very late times.

Once we have quantities that capture different aspects of a halo's formation history, we can look at how these quantities influence the internal structure of haloes. I found that the redshift at which haloes amassed half their mass —z_{50}— and when they had their last major mergers —z_{mm30}— affect most structural properties, specifically, concentration and the offset between the centre-of-mass and the density peak show clear linear trends with z_{50}. Other quantities, such as the halo's elongation or its spin, also show correlations, but they tend to be weaker and with significant scatter.

While looking at correlations between z_{mm30} and concentration, shape and centre-of-mass offset, the population of haloes that had recent major mergers ($z_{mm30} < 0.3$) are clearly segregated from the rest of haloes. This is expected from the initial phase of a merger as they alter shape (Drakos et al., 2019a), have lower concentration because the merger results in more mass at large radii as it falls in, and a large centre-of-mass offset (Drakos et al., 2019b). As the merging halo reaches the first pericentre and starts a damped oscillation behaviour, concentration and centre-of-mass offset will alternate between higher and lower values depending on the phase of the merger as shown by Wang et al. (2020).

Combining different structural probes allows us to estimate z_{50} more accurately, particularly with concentration and either the centre-of-mass offset, the axis ratio or the χ^2 of the mass profile. For example, populations of group-sized haloes with low concentration and large centre-of-mass offsets are almost exclusively young, while the opposite is also true.

Given that we have already established in Chapter 2 that z_{50} varies by almost a factor of two in the Ω_m–σ_8 plane, and that with only concentration as a predictor, we would need around 10 000 clusters to have competitive constraints on σ_8, we expect a combination of concentration, shape and centre-of-mass offset to allow to achieve the same results with far fewer clusters.

Chapter 5

Conclusion

5.1 Summary of the Thesis

In the work conducted during my doctoral thesis, I explored new avenues to exploit cluster data for cosmological purposes. My focus was towards using the structural properties of galaxy clusters as tools to constrain Ω_m and σ_8. I did that by first studying how cluster age and growth rate were influenced by Ω_m and σ_8 and then showed how we could measure age through cluster internal structure.

The Cosmological Dependence of Cluster Age

Galaxy clusters are the result of the evolution and gravitational collapse of rare, high-amplitude density peaks. Their abundance is a degenerate combination of the amplitude of the initial power spectrum—parameterised by σ_8—which controls the number of high peaks, and the amount by which they have grown which is sensitive to Ω_m. Their formation epoch, however, does not have the same dependence. Higher peaks will result in earlier formation, whereas a higher present-day growth factor will result in clusters assembling later. Cluster age is, therefore, an interesting property with unique cosmological dependence.

Analytical models are a valuable tool for studying the cosmological dependence of cluster age. These models allow inexpensive ways to span a wide range of parameter space while maintaining a relatively accurate prediction of structure formation.

Extended Press-Schechter theory (EPS) predicts that the median time when haloes form varies by 100% over the range of the Ω_m-σ_8 plane allowed by other observational

constraints. Interestingly, the sensitivity in this plane is almost orthogonal to typical structure growth constraints from weak lensing or cluster counts at low redshift.

Although age predictions from simulations vary, depending on the halo finders and merger tree algorithms used, they all agree with the general sensitivity of the halo age as a function of Ω_m-σ_8. This suggests that the cosmological dependence of age is present, roughly of the amplitude predicted by EPS theory.

It is worth noting that cluster age is not directly observable. However, it correlates with observable structural quantities such as concentration. If we assume a net scatter of 40–60% in the relation between age and measured concentration and use concentration as a test, we would need an unbiased sample of 10,000 clusters to obtain impactful constraints on σ_8 about 10% of the current 2–σ contours. Overall, the use of analytical models and simulations has provided valuable insights into the cosmological dependence of cluster age and has the potential to provide further constraints on important cosmological parameters such as σ_8. This has two major advantages: first, the constraints break the Ω_m–σ_8 degeneracy, and second they make use of information available "for free" in cluster survey data.

The Cosmological Dependence of Growth Rate

The growth rate is a differential analogue of the formation time. Analytical models, specifically EPS, also provide useful predictions of its dependence on Ω_m and σ_8. We looked at three different growth rate quantities: the merger rate per halo, the average growth in a dynamical timescale, and the fraction of haloes that experienced a major merger in the last dynamical timescale.

Analytical models predict fairly similar cosmological dependence for all three quantities, mostly orthogonal to cluster count constraints but with varying sensitivity. At low redshift, the merger rate varies by 40% in the Ω_m–σ_8 plane, while the analytical models predict a variation of up to 80% at low redshift and cluster mass for both the average growth and fraction of systems with large growth.

Despite the fact that the best time resolution of instantaneous growth rates in simulations can only be as good as the inter-snapshot time, and that different halo finders and merger tree algorithms will give different results, we found that the growth rates measured in simulations have similar cosmological sensitivity similar to that predicted in analytical models. The amplitude of the variation in simulations tends to be slightly smaller than the analytical predictions. In order to provide constraints that would distinguish between

values of $\sigma_8 = 0.7$ and $\sigma_8 = 0.9$, we need to measure the merger rate on either cluster or galaxy scales with a precision of 10% or better.

While challenging, measuring merger and growth rates at cluster scales is not impossible but will likely require large samples and accurate simulations for calibration. On the other hand, on galaxy scales, statistical analysis of merger rates has been performed theoretically and observationally. The difficulty, in this case, lies in linking the halo and galaxy merger rates. Recent measurements of galaxy merger rates show a scatter of about a factor of 2–3, which significantly exceeds the targeted 10% uncertainty.

This means that, although galaxy mergers are abundant, it is unlikely that we will be able to use them as a cosmological constraint in the near future, as long as the scatter in merger rate observational estimates is not significantly reduced.

Measuring the Age of Clusters

While cosmological dependence of cluster age and growth rate provides new constraints and possibilities, it is only useful if we can have accurate enough measurements of them through observations. Although it has been established that the formation history of clusters and dark matter haloes leaves imprints in their observed internal structure, a detailed study of the interplay between the two is necessary to understand how we can probe halo age with structures.

While studying the history and structure of haloes in this thesis, several findings emerged. Firstly, the halo mass accretion history can be represented as a largely one-dimensional quantity that describes the halo's relative mass throughout its past. This history varies primarily based on the halo's global age, the growth rate at early times, and the growth rate at recent times. However, age is the primary factor that distinguishes between different halo histories.

Secondly, any property describing the halo's state near the middle of its growth indicates the halo's general age. These quantities, such as z_{50}, show the clearest correlations with halo concentration and the offset between the halo's centre-of-mass and density peak.

Finally, combining multiple structural probes helps to reduce scatter in halo age estimates significantly. Notably, using concentration in combination with either shape or centre-of-mass offset can help to distinguish between young and old halo populations.

Summary

Galaxy clusters are massive structures that can be used to constrain cosmological parameters due to their abundance, which requires accurately estimating their mass through observational proxies. However, this method neglects other properties of clusters that could contain cosmological information and induces a degeneracy in estimates of Ω_m and σ_8. Various statistics of the dynamical state of galaxy clusters can be used to differentiate between values of Ω_m and σ_8. These include cluster formation time, merger rate, and growth rate, which vary with cosmology differently than cluster abundance. Simulations predict that a number of potentially observable structural properties can predict cluster dynamical states, such as the halo concentration and the offset between the centre-of-mass and the centre of the halo. These findings suggest that other clusters' properties should also be considered when constraining cosmological parameters using galaxy clusters.

5.2 Challenges and Future Work

This thesis' aim is to show that using future samples to constrain cosmology with cluster structure is a promising avenue that is worth exploring and has the potential to open a new door in the cosmological exploitation of the highly non-linear scales of structure formation. Any concrete plan aimed at this end is subject to challenges that will require further work and resources to overcome; these come in two main categories:

Linking simulations to observables

Most of the work in this thesis has been done using either analytical models or N-body simulations. Relating structural properties in simulations to specific observables is not always trivial. Halo mass, for example, is usually difficult to measure and requires either weak lensing, which can induce errors (e.g. Becker & Kravtsov, 2011), or another mass proxy, such as cluster richness (e.g. Capasso et al., 2019), which involves calibrating the richness-mass relationship. Consequently, if one wants to measure the density profile of the dark matter in the cluster, its shape, its centre-of-mass, the peak of its matter density, they would need to link them with observables accurately.

In practice, this task is made easier if we have access to multi-probe studies of clusters, X-ray tracks the ICM gas, but only the central parts of the cluster, but combined with SZ detection of the ICM, we can reach outer parts of the cluster. In contrast, optical surveys detect the galactic and collisionless parts of the cluster. In all three all cases, there is

still a need to match observables to the target structural properties and take into account projection effects.

One way to solve this problem is through mock observations, simulations which are supposed to replicate the anticipated observations from a given survey (e.g. Contarini et al., 2022). Good quality mocks will be crucial in accurately mapping structural properties from simulations to observables and, even more important, in assessing the accuracy of those measurements.

Baryonic effects

The internal structure of dark matter haloes can be affected by various effects originating from the gaseous and stellar components. Gas cooling may typically produce a denser central core and a more concentrated halo because of the adiabatic contraction of the gas and the subsequent additional gravitational effect of the more centrally concentrated gas on the dark matter (Schaller et al., 2015). Conversely, stellar and Active Galactic Nuclei (AGN) feedback may lead to more extended and less massive haloes because of gas ejection (Velliscig et al., 2014; Lee et al., 2018; Castro et al., 2020). This may also alter the shape of dark matter haloes because of the gas carrying away angular momentum, making the halo more spherical (Lee et al., 2018).

The main worry in the context of this thesis is that these effects induce more scatter, not accounted for in DM-only simulations, making the cosmological sensitivity of cluster structure weaker. For example, If baryons make haloes rounder, detecting very young haloes by looking at the highly non-spherical population, as I suggested in Section 4.5, can become more challenging. Moreover, since baryonic effects depend on the properties of their host galaxies, galaxy formation is a very cosmology-sensitive process (Mo et al., 2010; van Daalen et al., 2011), the way baryons affect halo structure can also, in principle, be cosmology-dependent and might alter the predicted cosmological dependence of halo structure.

More work is required to properly account for the impact of baryons on the relationship between cosmology and halo internal structure. Good-quality hydrodynamical simulations with cosmological scale box sizes are needed. Still, they are computationally costly and are dependent on the underlying subgrid prescriptions of stellar and AGN feedback, which happens to be the most likely source of disruption of halo internal structure. Current and future hydrodynamical are trying to address these shortcomings, FIRE simulations (Hopkins et al., 2018, 2023) study baryonic effects with a very good resolution ($10^4 M_\odot$) while

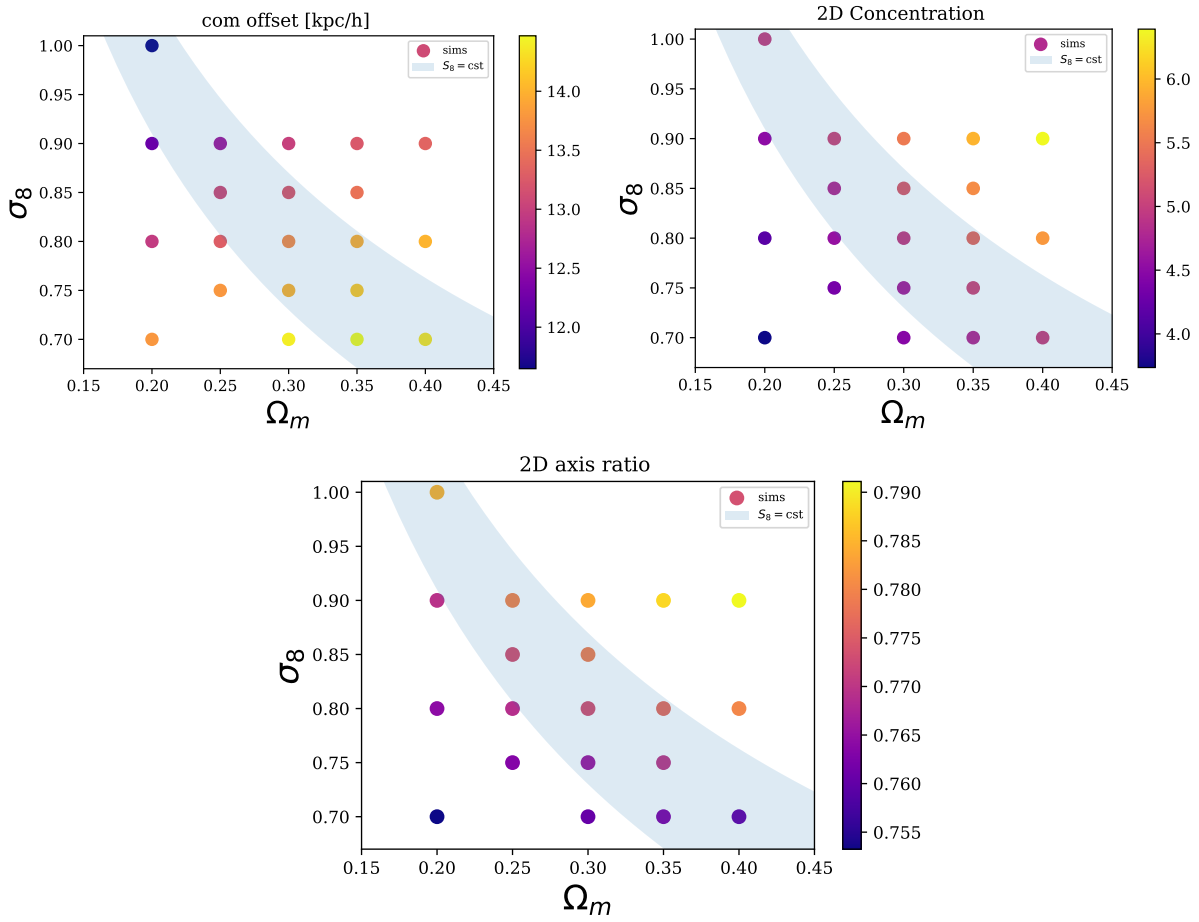

Figure 5.1: Change in the median value of the projected offset between the centre-of-mass and density peak (top left), concentration (top right) and axis ratio (bottom) for group-mass haloes calculated in simulations with different values of Ω_m and σ_8. The grey shaded area shows a contour of constant S_8.

the MilleniumTNG project aims at bridging the gap between cosmological-size simulations (500Mpc/h) and hydrodynamical effects (Pakmor et al., 2022).

What is observable ?

I show in Fig. 5.1 how three different structural properties which are, in practice, directly observable vary across the Ω_m–σ_8 plane. These are the offset between the centre-of-mass and density peak, concentration and axis ratio.

The centre-of-mass offset is of particular interest because, for one, it varies along the contour of constant S_8 as shown in Fig. 5.1 and second, because there are multiple ways to measure it. Using X-ray, lensing or combining probes. Although the relative change is about 20%, the absolute value is relatively small 11kpc–15kpc. As a test study, we can imagine a rough measurement of the offset between the X-ray peak and the Brightest Central Galaxy (BCG) of the cluster up to 100kpc accuracy. If we do this for 10,000 clusters, assuming only statistical errors, then we would get to an accuracy of 1 kpc, sufficient to distinguish between values of (Ω_m–σ_8) of (0.35–0.75) and (0.25–0.9).

Cluster and dark matter halo shape can be measured in various ways, most commonly through gravitational lensing (e.g. Robison et al., 2022) or by looking at the distribution of satellite galaxies (e.g. Shin et al., 2018). In either case, one can expect between 5% and 10% accuracy with a sample of O(10^4) clusters (Shin et al., 2018; Gonzalez et al., 2022). With a larger sample of O(10^5), and better signal-to-noise on individual detections, we could reach an accuracy of a few percent that would be required to have useful constraints according to the bottom panel of Fig. 5.1.

Concentration can also be estimated through lensing and galaxy number density profiles, as well as with X-ray and SZ (Groener et al., 2016). Provided we can accurately calibrate simulation predictions, according to the top right panel of Fig. 5.1, if we want to measure σ_8 at the 5% level, we would need roughly a 1% measurement of concentration. Assuming a conservative 60%-80% error on individual concentration estimates, we would "only" need about 4000-7000 clusters for our target. This is a very promising avenue, given we expect an order of magnitude more detections.

Next steps

The immediate next step for this work is to combine the findings of Chapters 2, 3 and 4 to go directly from cluster structure to Ω_m and σ_8 predictions. A few examples of this are

illustrated in Fig. 5.1, where I show how the projected centre-of-mass offset, concentration and axis ratio vary for different cosmologies. These come for the set of high-resolution N-body simulations that I have run as described in section 4.3.1. We can clearly see that this has the potential to break the Ω_m-σ_8 that follows contours of constant structure growth parameter S_8. Once we have a good sampling of Ω_m-σ_8 plane, multiple ways can be used to "fill" the plane and be able to associate structural parameters to any value of Ω_m and σ_8:

- **Emulators** In an ideal scenario, we could run thousands of simulations to effectively cover the relevant cosmological parameter space if we had access to unlimited computing power. However, the reality is that both N-body and hydrodynamical simulations are computationally intensive and can be a strain on computing resources. Furthermore, using these simulations to determine the differences between nearby regions in cosmological parameter space may not be necessary. Recent advancements in Machine Learning-based inference methods (e.g. Jamieson et al., 2022) provide a promising solution to this issue by offering a cost-effective way to generate simulation outputs at different cosmologies. These methods leverage a sample to learn and generate predictions, allowing for the simulation of a wider range of cosmological parameter space. Although the accuracy of these methods needs to be assessed further, the field is rapidly evolving and has great potential for use in cosmological predictions.

- **Interpolation** One noteworthy aspect of the behaviour displayed in Fig. 5.1 is the smooth and monotonic variation of the centre-of-mass offset across the plane. This is an encouraging indication of the quality of a simple interpolation of the function com_offset = f(Ω_m, σ_8) and for predicting other desired structural properties. Recently Chapman et al. (2022, 2023) have used Machine Learning-based interpolation methods to model the clustering of galaxies in the non-linear regime for different values of a set of five cosmological parameters. However, simpler and more effective interpolation methods can also be used since we are only focused on Ω_m and σ_8.

- **Machine Learning** Both methods described above make use of one quantity at a time for the cosmological inference; however, it is expected that a prediction from a combination of structural properties used together will typically be better than a combination of predictions from individual properties. This has been illustrated in section 4.5.2, where we saw that a combination of properties could split halo populations and significantly reduce scatter. Therefore, it is useful to think of ways of using all structural properties available to infer values of Ω_m and σ_8. Since we

already have a good sampling of the relevant parameter space, we could train a machine learning algorithm to predict the most likely combination of Ω_m–σ_8 given a set of median values of structural parameters. The data could be split into bins of mass and redshift to increase the amount of data used to predict cosmology if we expect to have observations across multiple ranges of masses and redshifts.

References

Abazajian K., et al., 2019a, arXiv e-prints, p. arXiv:1907.04473

Abazajian K., et al., 2019b, arXiv e-prints, p. arXiv:1907.04473

Abbott T. M. C., et al., 2018, Phys. Rev. D, 98, 043526

Abbott T. M. C., et al., 2020, Phys. Rev. D, 102, 023509

Abbott T. M. C., et al., 2022, Phys. Rev. D, 105, 023520

Abdalla E., et al., 2022, Journal of High Energy Astrophysics, 34, 49

Abdullah M. H., Wilson G., Klypin A., Old L., Praton E., Ali G. B., 2020, ApJS, 246, 2

Abellan G. F., Murgia R., Poulin V., Lavalle J., 2020, arXiv e-prints, p. arXiv:2008.09615

Ackermann S., Schawinski K., Zhang C., Weigel A. K., Turp M. D., 2018, MNRAS, 479, 415

Adams C., Blake C., 2020, MNRAS, 494, 3275

Aiola S., et al., 2020, J. Cosmology Astropart. Phys., 2020, 047

Akeson R., et al., 2019, arXiv e-prints, p. arXiv:1902.05569

Alam S., et al., 2017, MNRAS, 470, 2617

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

Allgood B., Flores R. A., Primack J. R., Kravtsov A. V., Wechsler R. H., Faltenbacher A., Bullock J. S., 2006, MNRAS, 367, 1781

Amoura Y., Drakos N. E., Berrouet A., Taylor J. E., 2021, MNRAS, 508, 100

Asgari M., et al., 2021, A&A, 645, A104

Avila S., et al., 2014, MNRAS, 441, 3488

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, AJ, 304, 15

Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291

Battye R. A., Charnock T., Moss A., 2015, Phys. Rev. D, 91, 103508

Becker M. R., Kravtsov A. V., 2011, The AJ, 740, 25

Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, ApJ, 762, 109

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

Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143

Berger M. J., Colella P., 1989, Journal of Computational Physics, 82, 64

Berger M. J., Oliger J., 1984, Journal of Computational Physics, 53, 484

Bett P., Eke V., Frenk C. S., Jenkins A., Helly J., Navarro J., 2007, MNRAS, 376, 215

Beutler F., et al., 2012, MNRAS, 423, 3430

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

Birkinshaw M., 1999, Phys. Rep., 310, 97

Blanchard A., Ilić S., 2021, arXiv e-prints, p. arXiv:2104.00756

Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, MNRAS, 456, 2361

Bocquet S., et al., 2019a, ApJ, 878, 55

Bocquet S., et al., 2019b, ApJ, 878, 55

Böhringer H., Chon G., 2016, Modern Physics Letters A, 31, 1640008

Böhringer H., Werner N., 2010, A&A Rev., 18, 127

Böhringer H., Chon G., Collins C. A., 2014, A&A, 570, A31

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, AJ, 379, 440

Boruah S. S., Hudson M. J., Lavaux G., 2020, MNRAS, 498, 2703

Bottrell C., et al., 2019, MNRAS, 490, 5390

Brainerd T. G., Blandford R. D., Smail I., 1996, ApJ, 466, 623

Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

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

Buote D. A., Humphrey P. J., 2012, MNRAS, 421, 1399

Butsky I., et al., 2016, MNRAS, 462, 663

Capasso R., et al., 2019, MNRAS, 486, 1594

Carlstrom J. E., Holder G. P., Reese E. D., 2002, ARA&A, 40, 643

Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499

Castro T., Borgani S., Dolag K., Marra V., Quartin M., Saro A., Sefusatti E., 2020, MNRAS, 500, 2316

Castro T., Borgani S., Dolag K., Marra V., Quartin M., Saro A., Sefusatti E., 2021, MNRAS, 500, 2316

Chambers K., Unions Team Including Pan-Starrs Team CFIS Team 2020, in American Astronomical Society Meeting Abstracts #235. p. 154.04

Chandrasekhar S., 1943, Reviews of Modern Physics, 15, 1

Chapman M. J., et al., 2022, MNRAS, 516, 617

Chapman M. J., Zhai Z., Percival W. J., 2023, arXiv e-prints, p. arXiv:2302.11621

Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109

Coe D., 2010, arXiv e-prints, p. arXiv:1005.0411

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

Conselice C. J., 2003, ApJS, 147, 1

Contarini S., et al., 2022, A&A, 667, A162

Contreras-Santos A., Knebe A., Cui W., Haggar R., Pearce F., Gray M., De Petris M., Yepes G., 2022, MNRAS, 515, 5375

Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015a, MNRAS, 450, 1514

Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015b, MNRAS, 452, 1217

Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015c, MNRAS, 452, 1217

Costanzi M., et al., 2021, Phys. Rev. D, 103, 043522

Cunha C., Huterer D., Doré O., 2010, Phys. Rev. D, 82, 023004

DESI Collaboration et al., 2016, arXiv e-prints, p. arXiv:1611.00036

Debackere S. N. B., Schaye J., Hoekstra H., 2021, arXiv e-prints, p. arXiv:2101.07800

Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, MNRAS, 456, 2486

Di Valentino E., Melchiorri A., Silk J., 2015, Phys. Rev. D, 92, 121302

Di Valentino E., Melchiorri A., Silk J., 2016, Phys. Rev. D, 93, 023513

Di Valentino E., et al., 2020a, arXiv e-prints, p. arXiv:2008.11285

Di Valentino E., Melchiorri A., Silk J., 2020b, Nature Astronomy, 4, 196

Di Valentino E., Melchiorri A., Mena O., Vagnozzi S., 2020c, Physics of the Dark Universe, 30, 100666

Di Valentino E., Melchiorri A., Mena O., Vagnozzi S., 2020d, Phys. Rev. D, 101, 063502

Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859

Diemer B., 2018, ApJS, 239, 35

Dodelson S., 2003, Modern Cosmology. Academic Press

Douspis M., Salvati L., Aghanim N., 2019, arXiv e-prints, p. arXiv:1901.05289

Drakos N. E., Taylor J. E., Berrouet A., Robotham A. S. G., Power C., 2019a, MNRAS, 487, 993

Drakos N. E., Taylor J. E., Berrouet A., Robotham A. S. G., Power C., 2019b, MNRAS, 487, 1008

Dubinski J., Carlberg R. G., 1991, AJ, 378, 496

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, MNRAS, 405, 2161

Duncan K., et al., 2019, ApJ, 876, 110

Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359

Dyson F. W., Eddington A. S., Davidson C., 1920, Philosophical Transactions of the Royal Society of London Series A, 220, 291

Einasto J., 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87

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

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

Euclid Collaboration et al., 2019, A&A, 627, A23

Evrard A. E., 1988, MNRAS, 235, 911

Evrard A. E., 1989, ApJ, 341, L71

Evrard A. E., 1990, ApJ, 363, 349

Evrard A. E., Mohr J. J., Fabricant D. G., Geller M. J., 1993, ApJ, 419, L9

Fakhouri O., Ma C.-P., 2008, MNRAS, 386, 577

Fakhouri O., Ma C.-P., 2009, MNRAS, 394, 1825

Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, MNRAS, 406, 2267

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

Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004, MNRAS, 355, 819

Gao L., Springel V., White S. D. M., 2005, MNRAS, 363, L66

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

Gautret L., Fort B., Mellier Y., 2000, Astronomy and Astrophysics, 353, 10

Genel S., Genzel R., Bouché N., Naab T., Sternberg A., 2009, ApJ, 701, 2002

Gill S. P. D., Knebe A., Gibson B. K., 2004, MNRAS, 351, 399

Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375

Giocoli C., Moreno J., Sheth R. K., Tormen G., 2007, MNRAS, 376, 977

Giocoli C., Tormen G., Sheth R. K., 2012, MNRAS, 422, 185

Gómez-Valent A., Solà Peracaula J., 2018, MNRAS, 478, 126

Gonzalez E. J., Hoffmann K., Gaztañaga E., García Lambas D. R., Fosalba P., Crocce M., Castander F. J., Makler M., 2022, MNRAS, 517, 4827

Gorski K., 1988, ApJ, 332, L7

Gottlöber S., Klypin A., Kravtsov A. V., 2001, ApJ, 546, 223

Goulding A. D., et al., 2018, PASJ, 70, S37

Groener A. M., Goldberg D. M., Sereno M., 2016, MNRAS, 455, 892

Gunn J. E., Gott J. Richard I., 1972, ApJ, 176, 1

Hahn C., et al., 2023, J. Cosmology Astropart. Phys., 2023, 010

Hamilton A. J. S., 2001, MNRAS, 322, 419

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

Hattori M., Kneib J., Makino N., 1999, Progress of Theoretical Physics Supplement, 133, 1

Heath D. J., 1977, MNRAS, 179, 351

Heimersheim S., Schöneberg N., Hooper D. C., Lesgourgues J., 2020, J. Cosmology Astropart. Phys., 2020, 016

Henry J. P., Arnaud K. A., 1991, ApJ, 372, 410

Henry J. P., Evrard A. E., Hoekstra H., Babul A., Mahdavi A., 2009, ApJ, 691, 1307

Heymans C., et al., 2013, MNRAS, 432, 2433

Heymans C., et al., 2020, arXiv e-prints, p. arXiv:2007.15632

Heymans C., et al., 2021a, A&A, 646, A140

Heymans C., et al., 2021b, A&A, 646, A140

Hikage C., et al., 2019, PASJ, 71, 43

Hildebrandt H., et al., 2017, MNRAS, 465, 1454

Hinshaw G., et al., 2013, ApJS, 208, 19

Hopkins P. F., et al., 2010, ApJ, 724, 915

Hopkins P. F., et al., 2018, MNRAS, 480, 800

Hopkins P. F., et al., 2023, MNRAS, 519, 3154

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

Huterer D., Shafer D. L., Scolnic D. M., Schmidt F., 2017, J. Cosmology Astropart. Phys., 2017, 015

Huško F., Lacey C. G., Baugh C. M., 2022, MNRAS, 509, 5918

Jõeveer M., Einasto J., Tago E., 1978, MNRAS, 185, 357

Jamieson D., Li Y., Alves de Oliveira R., Villaescusa-Navarro F., Ho S., Spergel D. N., 2022, arXiv e-prints, p. arXiv:2206.04594

Jeeson-Daniel A., Dalla Vecchia C., Haas M. R., Schaye J., 2011, MNRAS, 415, L69

Jenkins A., 2010, MNRAS, 403, 1859

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

Jimeno P., Broadhurst T., Lazkoz R., Angulo R., Diego J.-M., Umetsu K., Chu M.-c., 2017, MNRAS, 466, 2658

Jing Y. P., 2000, ApJ, 535, 30

Jing Y. P., Suto Y., 2002, AJ, 574, 538

Jing Y. P., Suto Y., Mo H. J., 2007, ApJ, 657, 664

Kacprzak T., et al., 2016, MNRAS, 463, 3653

Kazantzidis S., Kravtsov A. V., Zentner A. R., Allgood B., Nagai D., Moore B., 2004, AJ, 611, L73

Khintchine A., 1934, Mathematische Annalen, 109, 604

Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901

Kirby M., et al., 2019, arXiv e-prints, p. arXiv:1910.13548

Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340

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

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

Köhlinger F., Viola M., Valkenburg W., Joachimi B., Hoekstra H., Kuijken K., 2016, MNRAS, 456, 1508

LSST Dark Energy Science Collaboration (LSST DESC) et al., 2021, ApJS, 253, 31

LSST Science Collaboration et al., 2009, arXiv e-prints, p. arXiv:0912.0201

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

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

Lahav O., Liddle A. R., 2014, arXiv e-prints, p. arXiv:1401.1389

Leauthaud A., et al., 2017, MNRAS, 467, 3024

Lee B. E., Le Brun A. M. C., Haq M. E., Deering N. J., King L. J., Applegate D., McCarthy I. G., 2018, MNRAS, 479, 890

Lemaître G., 1927, Annales de la Société Scientifique de Bruxelles, 47, 49

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Lilje P. B., 1992, ApJ, 386, L33

Limousin M., Morandi A., Sereno M., Meneghetti M., Ettori S., Bartelmann M., Verdugo T., 2013, Space Sci. Rev., 177, 155

Lotz J. M., Primack J., Madau P., 2004, AJ, 128, 163

Lotz J. M., Jonsson P., Cox T. J., Croton D., Primack J. R., Somerville R. S., Stewart K., 2011, ApJ, 742, 103

Lucy L. B., 1977, AJ, 82, 1013

Ludlow A. D., et al., 2013, MNRAS, 432, 1103

Ludlow A. D., Navarro J. F., Angulo R. E., Boylan-Kolchin M., Springel V., Frenk C., White S. D. M., 2014, MNRAS, 441, 378

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

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

Mamon G. A., Łokas E. L., 2005, MNRAS, 363, 705

Mandelbaum R., 2018, ARA&A, 56, 393

Mann A. W., Ebeling H., 2012, MNRAS, 420, 2120

Mantha K. B., et al., 2018, MNRAS, 475, 1549

Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, MNRAS, 406, 1759

Mantz A. B., et al., 2015, MNRAS, 446, 2205

Martin G., Kaviraj S., Hocking A., Read S. C., Geach J. E., 2020, MNRAS, 491, 1408

McBride J., Fakhouri O., Ma C.-P., 2009, MNRAS, 398, 1858

Medezinski E., et al., 2017, in American Astronomical Society Meeting Abstracts #229. p. 125.08

Mellier Y., 1999, ARA&A, 37, 127

Meneghetti M., Bartelmann M., Jenkins A., Frenk C., 2007, MNRAS, 381, 171

Meneghetti M., et al., 2014, ApJ, 797, 34

Merloni A., et al., 2012, arXiv e-prints, p. arXiv:1209.3114

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution. Cambridge University Press

Mohr J. J., Evrard A. E., Fabricant D. G., Geller M. J., 1995, ApJ, 447, 8

Moore B., Governato F., Quinn T., Stadel J., Lake G., 1998, The AJ, 499, L5

Mundy C. J., Conselice C. J., Duncan K. J., Almaini O., Häußler B., Hartley W. G., 2017, MNRAS, 470, 3507

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

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

Navarro J. F., et al., 2004, MNRAS, 349, 1039

Navarro J. F., et al., 2010a, MNRAS, 402, 21

Navarro J. F., et al., 2010b, MNRAS, 402, 21

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

Nelson D., et al., 2019, Computational Astrophysics and Cosmology, 6, 2

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

Oguri M., Takada M., Okabe N., Smith G. P., 2010, MNRAS, 405, 2215

Okoli C., Afshordi N., 2016, MNRAS, 456, 3068

Padmanabhan T., 1993, Structure Formation in the Universe. Cambridge University Press

Pakmor R., et al., 2022, arXiv e-prints, p. arXiv:2210.10060

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

Parroni C., et al., 2017, ApJ, 848, 114

Pearson K., 1901, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 6, 559

Peebles P. J. E., 1980, The large-scale structure of the universe. Princeton University Press

Peter A. H. G., Rocha M., Bullock J. S., Kaplinghat M., 2013, MNRAS, 430, 105

Pfister H., Dotti M., Laigle C., Dubois Y., Volonteri M., 2020, MNRAS, 493, 922

Pillepich A., Porciani C., Reiprich T. H., 2012, MNRAS, 422, 44

Planck Collaboration et al., 2016a, A&A, 594, A14

Planck Collaboration et al., 2016b, A&A, 594, A24

Planck Collaboration et al., 2020, A&A, 641, A6

Plummer H. C., 1911, MNRAS, 71, 460

Poole G. B., Mutch S. J., Croton D. J., Wyithe S., 2017, MNRAS, 472, 3659

Poulin V., Boddy K. K., Bird S., Kamionkowski M., 2018, Phys. Rev. D, 97, 123504

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

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018

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

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

Richstone D., Loeb A., Turner E. L., 1992, ApJ, 393, 477

Riess A. G., et al., 2016, ApJ, 826, 56

Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, ApJ, 876, 85

Robison B., et al., 2022, arXiv e-prints, p. arXiv:2209.09088

Rodriguez-Gomez V., et al., 2015, MNRAS, 449, 49

Said K., Colless M., Magoulas C., Lucey J. R., Hudson M. J., 2020, MNRAS, 497, 1275

Sánchez A. G., 2020, Phys. Rev. D, 102, 123511

Sarazin C. L., 1986, Reviews of Modern Physics, 58, 1

Sarazin C. L., 2009, X-Ray Emission from Clusters of Galaxies. Cambridge University Press

Sartoris B., et al., 2016, MNRAS, 459, 1764

Schaller M., et al., 2015, MNRAS, 451, 1247

Schneider M. D., Frenk C. S., Cole S., 2012, J. Cosmology Astropart. Phys., 2012, 030

Sereno M., Umetsu K., Ettori S., Sayers J., Chiu I. N., Meneghetti M., Vega-Ferrero J., Zitrin A., 2018, ApJ, 860, L4

Sérsic J. L., 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

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

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

Shin T.-h., Clampitt J., Jain B., Bernstein G., Neil A., Rozo E., Rykoff E., 2018, MNRAS, 475, 2421

Skibba R. A., Macciò A. V., 2011, MNRAS, 416, 2388

Slipher V. M., 1915, Popular Astronomy, 23, 21

Snyder G. F., Rodriguez-Gomez V., Lotz J. M., Torrey P., Quirk A. C. N., Hernquist L., Vogelsberger M., Freeman P. E., 2019, MNRAS, 486, 3702

Sola J., 2018, arXiv e-prints, p. arXiv:1805.09810

Sola J., Gomez-Valent A., de Cruz Perez J., Moreno-Pulido C., 2020, arXiv e-prints, p. arXiv:2006.04273

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

Springel V., 2005, MNRAS, 364, 1105

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

Springel V., Pakmor R., Zier O., Reinecke M., 2021, MNRAS, 506, 2871

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

Stewart K. R., Bullock J. S., Barton E. J., Wechsler R. H., 2009, ApJ, 702, 1005

Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, 4, 173

Suto D., Peirani S., Dubois Y., Kitayama T., Nishimichi T., Sasaki S., Suto Y., 2017, PASJ, 69, 14

Tasitsiomi A., Kravtsov A. V., Gottlöber S., Klypin A. A., 2004, ApJ, 607, 125

Taylor J. E., 2011, Advances in Astronomy, 2011, 604898

Taylor J. E., Babul A., 2004, MNRAS, 348, 811

Taylor J. E., Babul A., 2005, MNRAS, 364, 515

Tenneti A., Mandelbaum R., Di Matteo T., Kiessling A., Khandai N., 2015, MNRAS, 453, 469

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

To C., et al., 2021, Phys. Rev. Lett., 126, 141301

Vega-Ferrero J., Yepes G., Gottlöber S., 2017, MNRAS, 467, 3226

Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Dalla Vecchia C., 2014, MNRAS, 442, 2641

Vera-Ciro C. A., Sales L. V., Helmi A., Frenk C. S., Navarro J. F., Springel V., Vogelsberger M., White S. D. M., 2011, MNRAS, 416, 1377

Verde L., Treu T., Riess A. G., 2019, Nature Astronomy, 3, 891

Walker M. G., Mateo M., Olszewski E. W., Peñarrubia J., Evans N. W., Gilmore G., 2009, AJ, 704, 1274

Walsh D., Carswell R. F., Weymann R. J., 1979, Nature, 279, 381

Wang L., Steinhardt P. J., 1998, ApJ, 508, 483

Wang K., Mao Y.-Y., Zentner A. R., Lange J. U., van den Bosch F. C., Wechsler R. H., 2020, MNRAS, 498, 4450

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

Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, The AJ, 568, 52

Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, ApJ, 652, 71

Weinberg D. H., Davé R., Katz N., Kollmeier J. A., 2003, in Holt S. H., Reynolds C. S., eds, American Institute of Physics Conference Series Vol. 666, The Emergence of Cosmic Structure. pp 157–169 (arXiv:astro-ph/0301186), doi:10.1063/1.1581786

White S. D. M., 1996, in Schaeffer R., Silk J., Spiro M., Zinn-Justin J., eds, Cosmology and Large Scale Structure. p. 349

White M., 2001, A&A, 367, 27

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

Wiener N., 1930, Acta mathematica, 55, 117

Wong A. W. C., Taylor J. E., 2012, ApJ, 757, 102

Xu C. K., Zhao Y., Scoville N., Capak P., Drory N., Gao Y., 2012, ApJ, 747, 85

Yuan Z. S., Han J. L., Wen Z. L., 2022, MNRAS, 513, 3013

Zel'dovich Y. B., 1970, A&A, 5, 84

Zenteno A., et al., 2020, MNRAS, 495, 705

Zhang J., Fakhouri O., Ma C.-P., 2008, MNRAS, 389, 1521

Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2003, The AJ, 597, L9

Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2009, The AJ, 707, 354

Zubeldia Í., Challinor A., 2019, MNRAS, 489, 401

de Haan T., et al., 2016, ApJ, 832, 95

eBOSS Collaboration et al., 2020, arXiv e-prints, p. arXiv:2007.08991

van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649

van den Bosch F. C., 2002, MNRAS, 331, 98

APPENDICES

Appendix A

On the origin of the cosmological dependence of the Halo Mass Function

A.1 Details of the Analytic Calculations

We use standard tools and techniques for the analytic calculations in Section 2.3. In particular, the growth factor $g(z)$ is calculated using the approximation given by Carroll et al. (1992), which is accurate to a few percent:

$$g(z) \approx \frac{5\Omega_m(z)}{2\left[\Omega_m^{4/7}(z) - \Omega_\Lambda(z) + (1+\Omega_m(z)/2)(1+\Omega_\Lambda(z)/70)\right]}\,. \tag{A.1}$$

The critical overdensity δ_c is the value a linearly-extrapolated density perturbation needs to reach to collapse and form a virialized object, and can be estimated using the spherical collapse model. Its present-day value varies weakly with Ω_m (Mo et al., 2010):

$$\delta_c = \frac{3}{5}\left(\frac{3\pi}{2}\right)^{2/3}\Omega_m^{0.0055} \approx 1.686\Omega_m^{0.0055}\,. \tag{A.2}$$

The linear matter power spectrum

$$P(k) = Ak^{n_s}T(k) \tag{A.3}$$

is computed using the approximation to the transfer function $T(k)$ given by Equation 16 in Eisenstein & Hu (1998). We take the primordial amplitude to be A=1 initially, and then adjust this value retroactively to set the correct value for σ_8. The index of the primordial spectrum is taken to be $n_s = 0.965$.

The variance of the density fluctuation field, σ^2, is computed numerically by convolving the power spectrum $P(k)$ with a top hat smoothing filter:

$$\sigma^2(R) = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) \widetilde{W}_R^2(k) dk \,, \tag{A.4}$$

where

$$\widetilde{W}_R(k) = 3 \frac{\sin(kR) - kR\cos(kR)}{(kR)^3} \tag{A.5}$$

We have compared our derived values of $P(k)$ and $\sigma(M)$ to values calculated using the Colossus python package (Diemer, 2018), and find good agreement.

A.2 Survey mass versus $\Omega_{\rm m}$

Halo abundance depends on the total mass of material sampled in a survey volume, M_V, and on the collapsed fraction f at that scale and redshift. The survey mass obviously depends on the mean matter density ρ_m and thus on Ω_m directly, but it also depends on Ω_m indirectly, through the volume sampled for a given solid angle and redshift range. The left panel of Fig. A.1 shows the total mass contained within a survey volume per unit redshift per unit solid angle at two different redshifts, as a function of the cosmological parameter $\Omega_{\rm m}$. While the volume element is a decreasing function of $\Omega_{\rm m}$ for flat ΛCDM cosmologies (since volume at a given redshift grows as Λ increases), the total mass enclosed increases overall, with the greatest $\Omega_{\rm m}$ sensitivity at low redshift.

In the end, however, the total mass M_V has relatively little influence on the overall shape of the abundance contours in the Ω_m–σ_8 plane, because its variation is of order a factor 2 or less, while the collapsed fraction varies over several orders of magnitude as Ω_m changes (right panel of Fig. A.1; see Appendix A.3 for a discussion of the parametric dependence of the collapsed fraction).

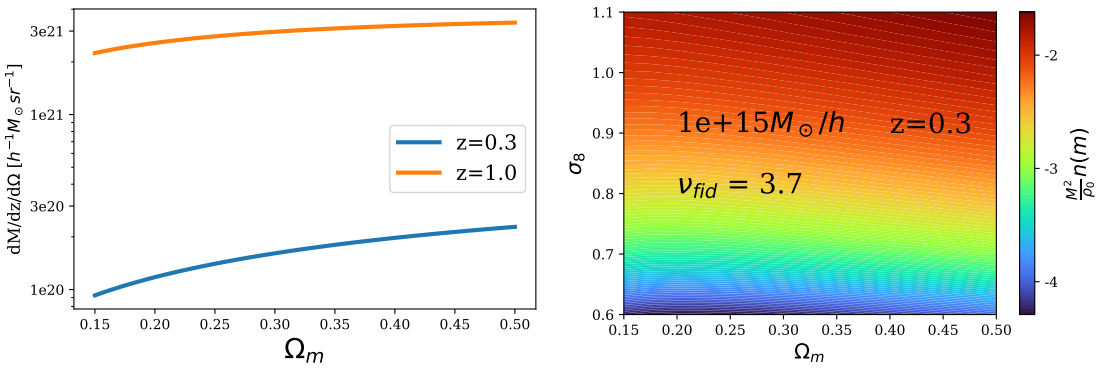

Figure A.1: Left : total mass M_V contained within a survey volume, per unit solid angle and per unit redshift interval, as a function of $\Omega_{\rm m}$, at redshifts 0.3 and 1 (note a flat ΛCDM cosmology is assumed). Right : Variation of the (EC) collapsed fraction $f_{\rm ST}$ in the $\Omega_{\rm m}$–σ_8 plane, for the particular choice of halo mass and redshift indicated. Note the similarity to the dependence of peak height (Fig. 2.1), although the colour scale here is now inverted, and logarithmic.

A.3 Peak Height and collapsed fraction versus $\Omega_{\rm m}$ and σ_8

We can write peak height as the product of three factors:

$$\nu = \frac{\delta_c(z)}{\sigma(M)} = \frac{\delta_c(z)}{\sigma_8 \Gamma(M)}\,, \tag{A.6}$$

where we have defined $\Gamma(M) \equiv \sigma(M)/\sigma_8$. The redshift evolution of the collapse threshold stems from the linear growth of fluctuations

$$\delta_c(z) = \delta_c(0)\frac{D(0)}{D(z)} = \delta_c(0)\frac{a_0 g(0)}{a g(z)}\,, \tag{A.7}$$

where D is the fluctuation amplitude, g is the linear growth factor, and a is the scale factor. Combining these,

$$\nu = \frac{(1+z)}{\sigma_8}\frac{1}{\Gamma(M)}\frac{\delta_c(0)}{g(z)/g(0)}\,. \tag{A.8}$$

Thus, while peak height scales simply as $1/\sigma_8$, the dependence on $\Omega_{\rm m}$ is through both the shape of the matter power spectrum $\Gamma(M)$ and the relative growth factor $g(z)/g(0)$. We consider each of these in turn.

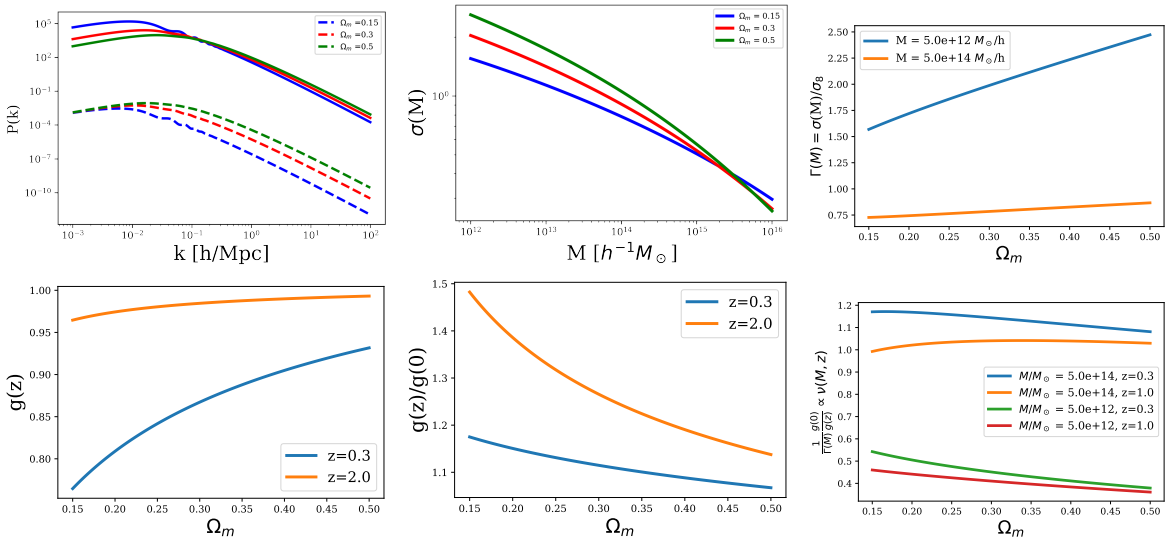

Figure A.2: Upper left: Variation of the power-spectrum $P(k)$ with $\Omega_{\rm m}$ at fixed primordial amplitude (dashed lines) and for fixed σ_8 (solid lines). The upper middle and upper right panels show how $\sigma(M)$ and $\Gamma(M)$ depend on $\Omega_{\rm m}$. The lower left and lower middle panels show how the growth factor and the normalized growth factor vary with $\Omega_{\rm m}$. The lower right panel shows how, as a consequence of these dependencies, the peak height ν varies with $\Omega_{\rm m}$ at fixed σ_8, following Eq. A.12.

$\Gamma(M)$

The function $\Gamma(M)$ describes the shape of the amplitude of fluctuations as a function of mass, $\sigma(M)$, independent of its normalization σ_8. This shape will depend on the value of the matter density parameter $\Omega_{\rm m}$. More specifically, matter-radiation equality occurs sooner in cosmologies with larger matter densities. Growth is suppressed by radiation on all scales below the horizon scale at recombination, but these scales are smaller and spend less time inside the horizon when $\Omega_{\rm m}$ is larger. Thus, for a fixed amplitude A of the primordial power spectrum, small-scale power at recombination will increase with $\Omega_{\rm m}$ (dashed curves in the top left panel of Fig. A.2).

Fixing σ_8, the amplitude of fluctuations at scales of $8h^{-1}$Mpc, reduces some of the difference in power (solid curves in the top left panel of Fig. A.2), but the shape of $\sigma(M)$ remains steeper in cosmologies with larger values of $\Omega_{\rm m}$ (top middle panel of Fig. A.2). Thus, over the range of interest, the ratio $\Gamma = \sigma(M)/\sigma_8$ increases with $\Omega_{\rm m}$, especially at

lower mass (top left panel of Fig. A.2). We can model this dependence as

$$\Gamma(M) \sim \Omega_{\rm m}^{\beta(M)}\,, \tag{A.9}$$

where $\beta(M)$ decreases with mass, and becomes negative at $M = M_8 \approx 2 \times 10^{15} M_\odot/h$ where $\sigma(M) = \sigma_8$.

Growth factor

The lower panels of Fig. A.2 show how $g(z)$ and $g(z)/g(0)$ vary with $\Omega_{\rm m}$. In a flat ΛCDM model, the growth factor $g(z)$ at a fixed redshift z is reduced for low $\Omega_{\rm m}$, as dark energy suppresses the growth of fluctuations. The relative amplitude of the effect is greater for very low values of $\Omega_{\rm m}$, or for large redshift ranges, and thus the ratio $g(z)/g(0)$ is *largest* for low $\Omega_{\rm m}$ and/or high redshift (middle panel). Mathematically, we can estimate the dependence on the density parameter by using the approximation $g(z) \propto \Omega_{\rm m}^{3/7}(z)$ (Carroll et al., 1992):

$$\frac{g(z)}{g(0)} \propto \left(\frac{\Omega(z)}{\Omega_0}\right)^{3/7} \propto \left(\frac{\bar{\rho}_m(z)\rho_c(0)}{\bar{\rho}_0\rho_c(z)}\right)^{3/7} \propto \frac{((1+z)^3)^{3/7}}{E(z)^{3/7}} \propto \Omega_{\rm m}^{-\alpha(z)}\,, \tag{A.10}$$

where

$$E(z) = \frac{H(z)}{H(0)} \sim \sqrt{(1+z)^3\Omega_{\rm m} + \Omega_\Lambda} \tag{A.11}$$

is the Hubble ratio. The index $0 < \alpha(z) < 1$ is an increasing function of z, and approaches the value 3/14 at high redshift.

Peak Height

Given these results, the overall dependence of the peak height on $\Omega_{\rm m}$ and σ_8 can be written

$$\nu(M,z) \propto \frac{1}{\sigma_8}\Omega_{\rm m}^{\alpha(z)}\Omega_{\rm m}^{-\beta(M)}\,, \tag{A.12}$$

where $\alpha > 0$ and increases with redshift to the limiting value 3/14, while $\beta(M)$ is a decreasing function of mass, and is negative at large masses. Combining the positive slope of Γ with the negative slope of $g(z)/g(0)$, the bottom right-hand panel of Fig. A.2 shows how ν generally decreases with $\Omega_{\rm m}$, but can increase for large masses and high redshifts. This explains the slight ν-dependence seen in Fig. 2.1. The summary of how the peak height varies with mass and redshift for different values of Ω_m and σ_8 is shown in Fig. A.3.

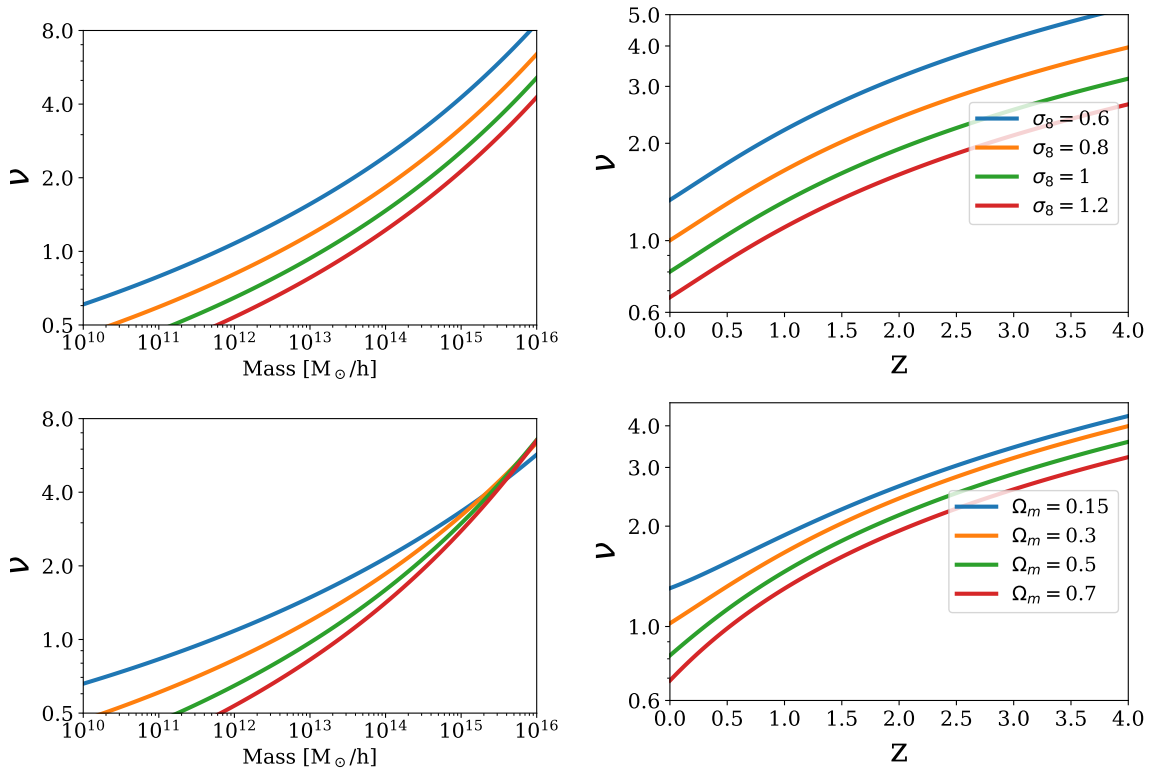

Figure A.3: Variation of the peak height ν with mass (left panels) and redshift (right panels), for different values of the cosmological parameters. The two top panels show the dependence on σ_8 at fixed $\Omega_{\rm m} = 0.3$, while the bottom panels show the dependence on $\Omega_{\rm m}$ at fixed $\sigma_8 = 0.8$.

Collapsed Fraction

Combining the peak height dependence described above with the functional form of the collapsed fraction as a function of peak height (Eqs. 2.3 or 2.4) we obtain the collapsed fraction as a function of mass for different values of σ_8 and Ω_m, shown in the left and right panels of Fig. A.4 respectively. It shows two regimes, a power-law increase at low masses, followed by an exponential decrease for cluster mass haloes. The figure also shows why the collapsed fraction sensitivity to σ_8 and Ω_m varies with mass, and the transition between the two regimes occurs at different masses for different values of σ_8.

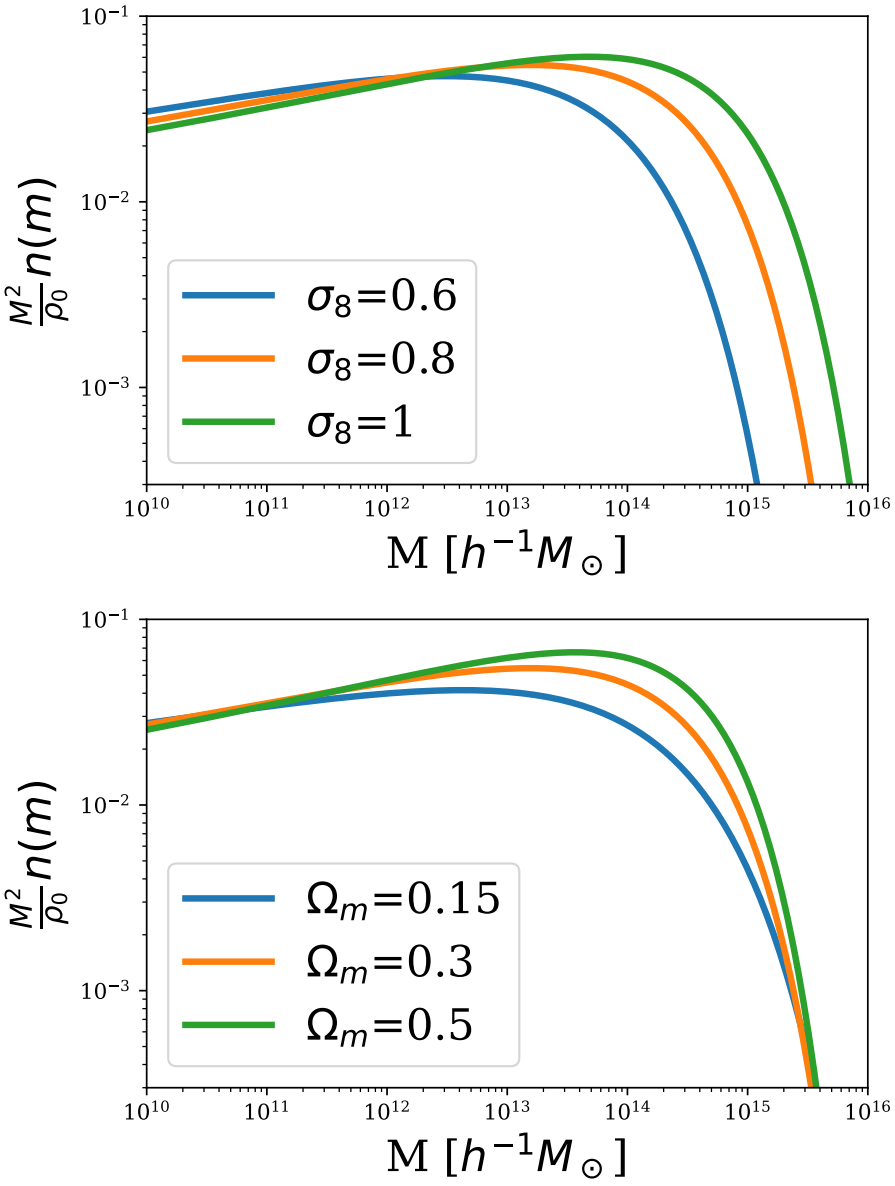

Figure A.4: Variation of the collapsed fraction with σ_8 (up) and Ω_m (bottom). The collapsed fraction increases slowly for low values of ν, but then drops exponentially at high values; σ_8 changes the mass at which the transition to exponential behaviour occurs, while Ω_m changes the sharpness of the transition.

Appendix B

Parametric Dependence of the Formation Time

To understand the dependence of formation time on the cosmological parameters $\Omega_{\rm m}$ and σ_8, we consider the simplest, SC or Press-Schechter (PS) expression for z_{50}:

$$P_{PS}(z_{50} > z|M_0, z_0) \equiv \int_{M_0/2}^{M_0} \frac{M_0}{M} f_{PS}(M, z|M_0, z_0) dM \,, \tag{B.1}$$

where

$$\begin{aligned} f_{PS}(M_1, z_1|M_0, z_0) dM_1 = & \frac{1}{\sqrt{2\pi}} \frac{\delta_c(z_1) - \delta_c(z_0)}{(S(M_1) - S(M_0))^{3/2}} \\ & \times \exp\left(-\frac{(\delta_c(z_1) - \delta_c(z_0))^2}{2(S(M_1) - S(M_0))}\right) dS_1 \,. \end{aligned} \tag{B.2}$$

We can express f_{PS} as a function of a single variable

$$D\nu \equiv \frac{\delta_c(z_1) - \delta_c(z_0)}{\sqrt{S(M_1) - S(M_0)}} \,,$$

as follows:

$$f_{PS}(M_1, z_1|M_0, z_0) dM_1 = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{(D\nu)^2}{2}\right) |d(D\nu)| \,. \tag{B.3}$$

Thus, we see that the conditional probability has the same form as the unconditional one, but with the argument $D\nu$ rather than ν.

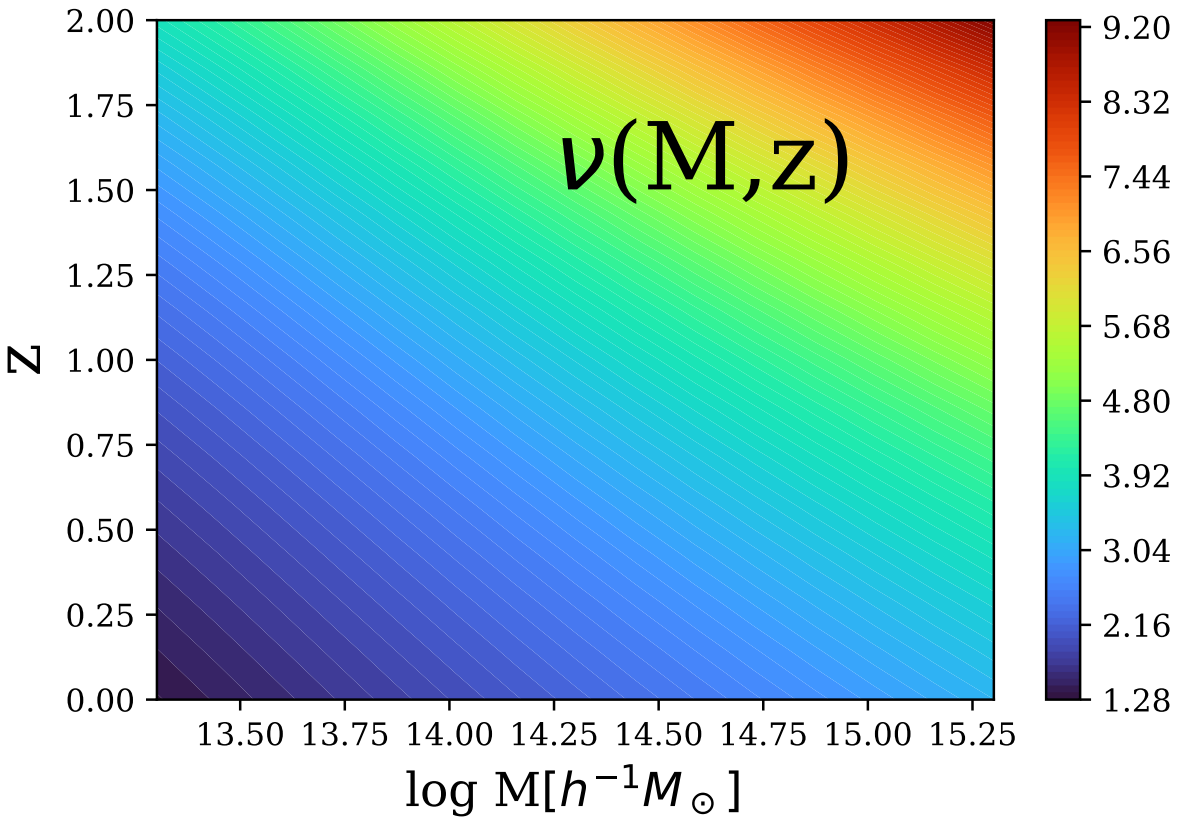

Figure B.1: Peak height ν as a function of mass and redshift, in our fiducial ($\Omega_{\rm m} = 0.3, \Omega_\Lambda = 0.7$) cosmology.

The formation redshift distribution is proportional to the average value of the PMF between $M_0/2$ and M_0, and since the factor M_0/M does not vary much over this range, it is also approximately equal to the conditional probability evaluated around the middle of the range. This probability f_{PS} in turn goes as $\exp[-(D\nu)^2/2]$, so we expect the parametric dependence of $\langle z_{50} \rangle$ to resemble an inverted, logarithmic version of the dependence for $D\nu$.

Fig. B.1 shows ν as a function of mass and redshift, while the top 4 panels of Fig. B.2 show the value of $D\nu$ as a function of z_1 and mass fraction M_1/M_0, for various values of M_0 and z_0. We see that the shape of the $D\nu$ contours is generally similar to those for ν, except when z_1 is close to z_0 (bottom of the plot), or when the mass fraction is close to 1 (right hand side of the plot).

The second set of 4 panels in Fig. B.2 shows the value of the conditional probability. As expected, the conditional probability is similar to an inverse, logarithmic mapping of $D\nu$.

Finally, we can consider the behaviour of $D\nu$ and the PMF in the $\Omega_{\rm m}$–σ_8 plane. The top 4 panels of Fig. B.3 show the value of $D\nu$ in this plane, for the values of (M_0, z_0) indicated, a mass fraction $M_1/M_0 = 0.5$, and $\Delta z = z_1 - z_0 = 0.1$. The overall pattern is very similar to that of ν (cf. Fig. 2.1). The bottom four panels show the value of the PMF, for the same choices of (z_0, z_1, M_0, M_1). Relative to the top panels, we see that the colour scale is inverted and logarithmic, as expected. The overall behaviour explains the shape of the contours in Fig. 2.5, and their relative orthogonality to abundance contours in the same plane.

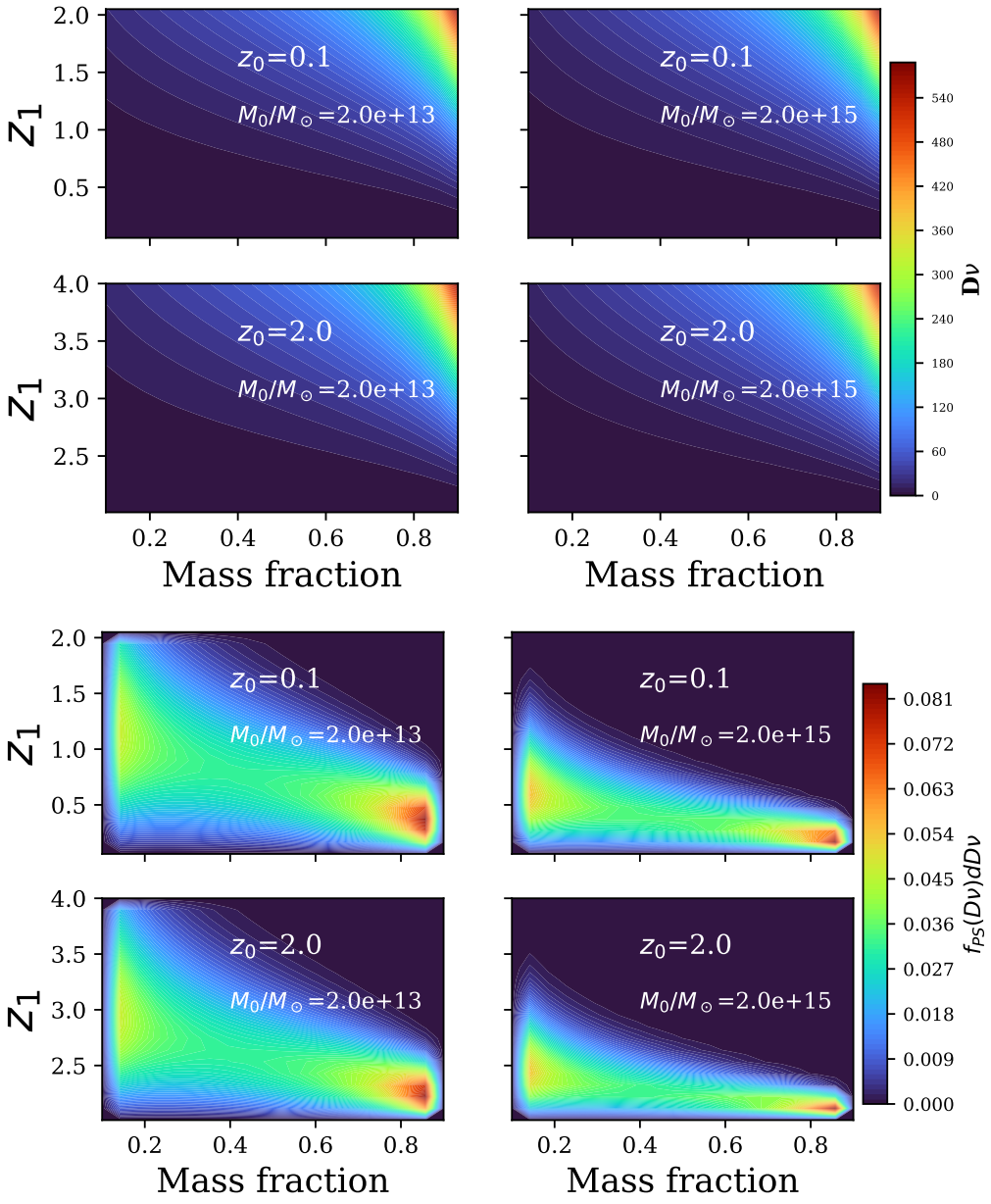

Figure B.2: Top 4 panels: The variable $D\nu$ as a function of mass fraction M_1/M_0 and z_1, for the values of (M_0, z_0) indicated, in our fiducial cosmology. Bottom 4 panels: corresponding conditional probability $f_{PS}(D\nu)d(D\nu)$.

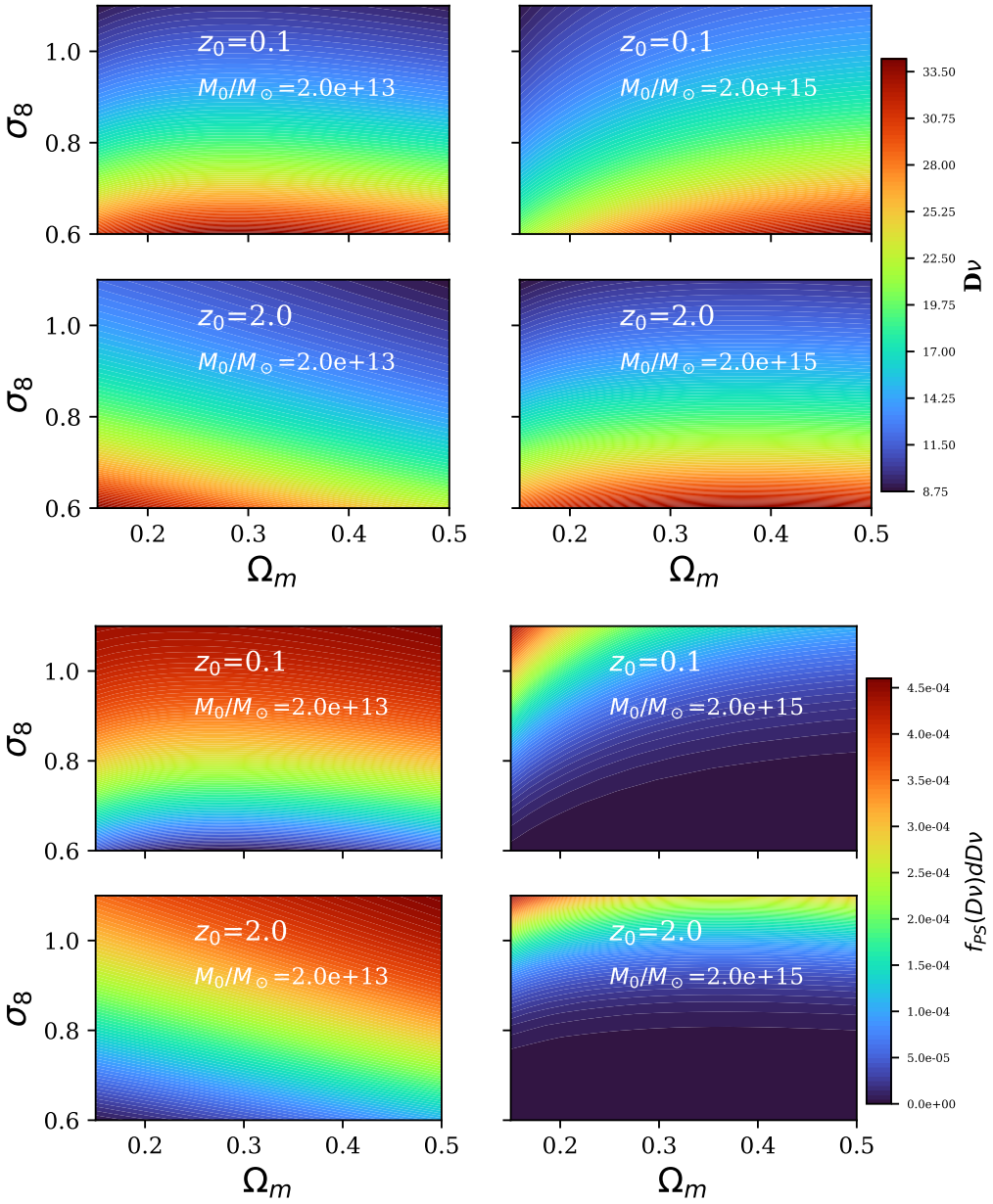

Figure B.3: Top : $D\nu$ as a function of $\Omega_{\rm m}$ and σ_8, for the values of (M_0, z_0) indicated, a mass fraction $M_1/M_0 = 0.5$, and $\Delta z = z_1 - z_0 = 0.1$. Bottom: The corresponding conditional probability $f_{PS}(D\nu)d(D\nu)$ values as a function of $\Omega_{\rm m}$ and σ_8.

Appendix C

Fits to the Average Halo Growth and Large Growth

Analytical models and simulations show that the main cosmological dependence of both the average halo growth (AHG) and the fraction of haloes with large growth (LGS fraction) is through σ_8, while they depend only weakly with Ω_m. In order to capture that dependence, we fit a power-law of the form

$$f(M|A, M_0, \alpha, p) = A(M/M_0)^{\alpha} + p \tag{C.1}$$

for both these quantities. Note that for the average halo growth at z=1, a two-power-law function seemed more appropriate

$$f(M|A, M_0, \alpha, p) = A(M/M_0)^{\alpha}(1 + M/M_0)^{\beta} + p \tag{C.2}$$

The parameters A, M_0, α and β capture the overall shape of the function and depend on Ω_m and z, but are all made constant with σ_8. In contrast, the overall normalisation p is the only parameter that varies with σ_8. This was motivated by the fact that in analytical models, the value of σ_8 does not affect the shape of either function.

We show the fits to the LGS fraction and the AHG in Figs. C.1 and C.2 respectively, for a range of masses $10^{12}M_{\odot}/h < M < 10^{14}M_{\odot}/h$.

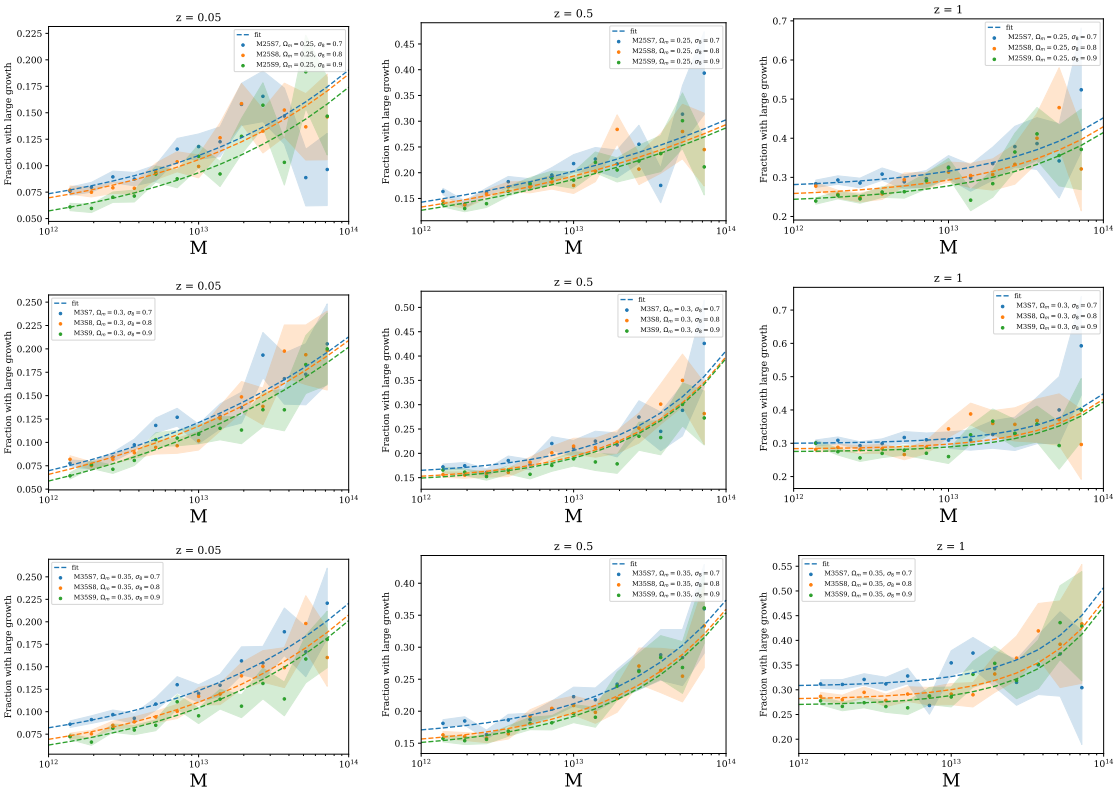

Figure C.1: The cosmological dependence of the LGS fraction. Dashed lines represent power-law fits where only the normalisation between each simulation is fitted.

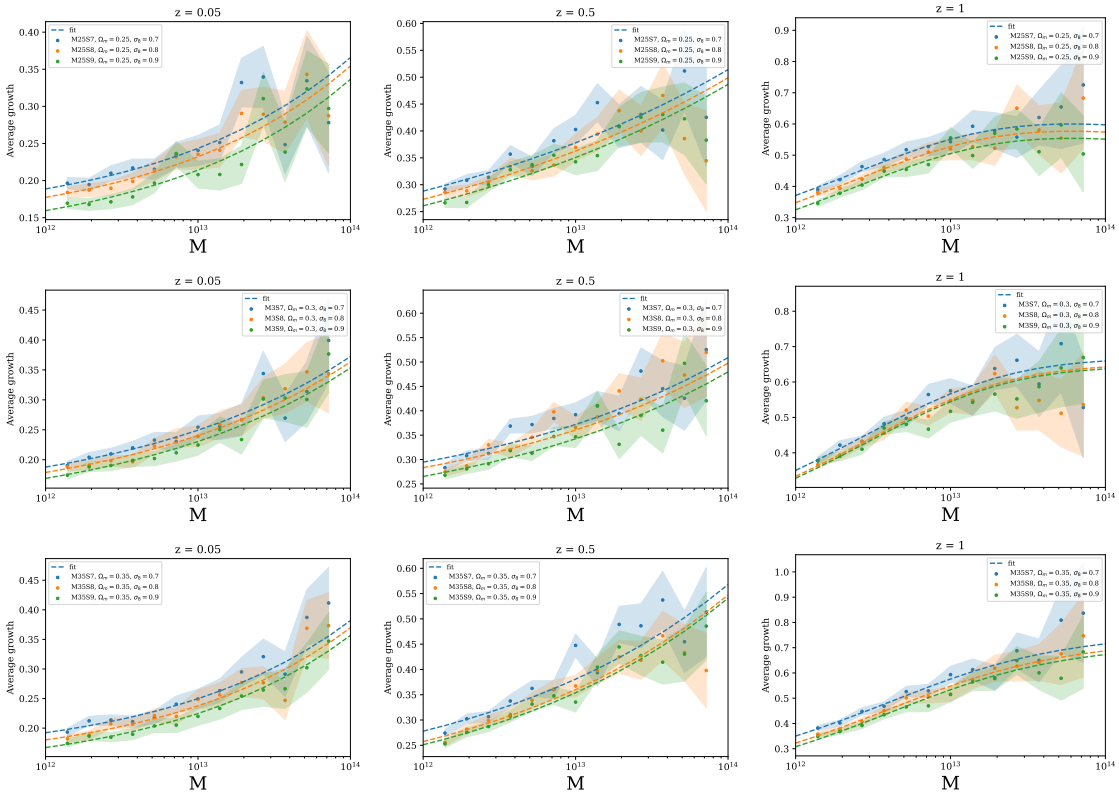

Figure C.2: The cosmological dependence of the AHG. Dashed lines represent power-law fits where only the normalisation between each simulation is fitted.