NONLINEAR GROWTH OF STRUCTURE IN COSMOLOGICAL SIMULATIONS

BY

ZARIJA LUKIĆ

Dipl., University of Belgrade, 2002
M.S., University of Illinois at Urbana–Champaign, 2004

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Astronomy
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2008

Urbana, Illinois

Doctoral Committee:

Prof. Paul M. Ricker, Chair
Prof. Joseph Mohr
Prof. Brian Fields
Prof. Charles Gammie
Dr. Katrin Heitmann, Los Alamos Nat. Lab.

Abstract

Upcoming cosmological observations (South Pole Telescope, Atacama Cosmology Telescope, Baryon Oscillation Spectroscopic Survey, and Planck) will allow for accurately probing structures and their growth, some into highly nonlinear regimes. These observations, in combination with already very accurate measurements of the expansion rate of the universe, will not only constrain cosmological parameters to a percent level, but will also answer what is the theory of gravity on the largest scales. In order to obtain theoretical predictions for different measurables (like the distribution of masses, spatial correlations), large numerical simulations have to be carried out. In this context, their main goal is to quantify how are such measurables affected by a change of cosmological parameters. The promised high accuracy of observations make the simulation task very demanding, as the theoretical predictions have to be at least as accurate as the observations.

In this thesis, we study the formation and evolution of dark matter halos in ΛCDM models over a wide range of cosmological epochs, from redshift z=20 to the present. First, we focus on the halo mass function, likely a key probe of cosmological growth of structure. By performing a large suite (60 simulations) of nested-box N-body simulations with careful convergence and error controls, we determine the mass function and its evolution with excellent statistical and systematic errors, reaching a few percent over most of the considered redshift and mass range. Our results are consistent with a 'universal' form for the mass function, and are in a good agreement with the Warren et al. analytic fit. Next, we study the structure of halos and ramification of different halo mass definitions. This analysis is important for connecting structure formation theory with observations, and also impacts the widely used approaches of assigning visible galaxies to dark matter halos - the halo occupancy distribution models. We find that the vast majority of halos (80-85%) appear as isolated objects, allowing for an accurate mapping between the two main mass definitions (friends-of-friends and spherical overdensity). Based on results from Monte Carlo realizations of ideal Navarro-Frenk-White halos and N-body simulations we provide a mass mapping formula. Furthermore, investigation of non-isolated, bridged halos, reveals that the fraction of these halos and their satellite mass distribution is cosmology dependent, and can be expressed in a cosmology universal form. Third, we turn to the spatial distribution of halos, which serves as a 'biased' mass tracer. While this bias is scale dependent, at large distances it asymptotes

to a constant value. We show that commonly used, heuristic approach to relating the mass function to the bias (peak-background split) clearly fails at the accuracy we are interested in ($\leq$10%). Using our large set of simulations we provide universal formula for halo bias as a function of mass. This formula fit well not only our data, but the current state of the art simulation data (Millenium simulation).

Finally, we present the results of a comparison between 10 different cosmology codes. These include virtually all major codes used today, and more importantly, they completely cover the range of numerical algorithms used in cosmological N-body simulations. For the mass function, the matter power spectrum, and halo profiles – the most important statistics for this thesis – codes agree at less then 10% over wide dynamic ranges. This robustness gives us additional confidence in our numerical results.

To my family.

Acknowledgments

This Ph. D. work would not be successful without support of many people. I would like to express my deep gratitude to my advisor Prof. Paul Ricker. He showed great patience, particularly in the first years of my graduate school; at all times I had his unconditional support. I have learned tremendously from him. I am especially thankful to my mentors in Los Alamos National Laboratory, Dr. Salman Habib and Dr. Katrin Heitmann for believing in me, for challenging me with new ideas, and for encouraging me. Their suggestions and motivation were the stepping stones for the successful completion of this thesis. Also, I would like to thank my professors and committee members, Prof. Brian Fields, Prof. Charles Gammie, and Prof. Joseph Mohr for everything I have learned from them. The excellent courses they taught, as well as the many informal discussions I had with them enabled me to shape my interests in science. They provided me with invaluable knowledge.

There is no way to express properly the gratitude to my parents and my sister – they just made me what I am. Hopefully, this thesis shows that at least some things turned out fine. The same is true for my late grandparents; the memory of them provides a persistent inspiration for my journeys. If only they could have been here these moments. During several years of student life in the US, Ayumi was with me, sharing all ups and downs; on more than one occasion she was the only support I had.

Many dear people made my grad school a smooth and enjoyable experience: Tijana, Vlada, Maki, Nemanja, Sale, Djidji, Marko, Paja, Miloš, Ivan, Amit, X, Shweta, Scott, Karen, Matt... and the list goes on. I prefer not to even imagine doing a Ph. D. in a foreign country without their camaraderie. A large part of the research was done in New Mexico, and there I had great help from Miki, Lilla and Mister K, who turned my stays there into memorable experiences. Last, but not least, I acknowledge the help from the Peruvian embassy in Albuquerque, and in particular His Excellency, Ambassador Paul Valderrama Rodriguez.

Table of Contents

List of Tables

List of Figures

1 Introduction

Cosmology studies the Universe as a system, how it formed, how it evolved and its future course. Likely one of the oldest human endeavors, modern cosmology grew from ideas before recorded history. The first reasonings of what would be denoted today as 'physical cosmology' date from the early nineteenth century. In 1823, German astronomer Heinrich Olbers argued that the universe must be finite because the night sky is dark. He reasoned that if the universe is infinite, eternal, and static with stars throughout, then for any particular direction, the line-of-sight would eventually end on the surface of a star. Of course, finitness of the universe is incompatible with Newton's theory of gravity, as its always attractive forces would have caused the entire universe to collapse onto itself.

Step by step, the reasoning lead to the model of the universe in expanding space, homogeneous and isotropic on very large scales. The universe had its time of birth called the Big Bang. At all times, cosmology modeling is guided by the principle of Occam's razor, and the requirement for robustness of the initial conditions. The Big Bang, or *standard* model is currently the simplest physical formulation of cosmology which accounts for all known observations.

Fundamentally, the confidence in the Big Bang model comes from four crucial observations: The universe is expanding according to the Hubble law, such that the further an object is from us, the larger recession velocity it has. Following the cosmological principle, any observer in the universe should make the same observation, leading to the model of expanding space. Since the universe is expanding, it had to be smaller and hotter in the past than it is today. This leads to the second fundamental observation – a thermal sea of photons with virtually perfect blackbody spectrum at $T \approx 2.7°\mathrm{K}$ filling space isotropically. Third, we observe that the universe has a certain abundance of atomic elements; ratio of light elements H, D, ^{3}He, ^{4}He, ^{7}Li to hydrogen cannot be explained by stellar fusion, but is easily understood from primordial nucleosynthesis.

Finally, and this is the focus of this thesis, we can observe that structure forms in a certain way, such that smaller mass objects existed in the past, while the most massive structures are still forming. Also, we can observe that objects are not uniformly distributed throughout the space, but that they form certain patterns. By analysing statistical properties of the large-scale structure one can thus infer cosmological parameters.

The thesis can be divided into four parts:

Part I: Theoretical background and current critical observational campaings

are described in Chapter 1. Chapter 2 provides a review of important aspects of linear theory describing the first phase of structure formation. Also, it provides some analytical extensions into non-linearity.

Part II: Nonlinear evolution is discussed in Sections 3–5. Each section focuses on one statistical description of dark matter objects (halos) in N-body simulations: Chap. 3 contains results for the distribution of masses; in Chap. 4 the structure of halos is analyzed, and cosmological implications are discussed. Finally, Chap. 5 investigates the spatial distributions of halos, described via correlation functions.

Part III: As virtually all results in this thesis are obtained via computer simulations, Chapter 6 analyzes the accuracy and systematical errors in different numerical algorithms. While some important convergence tests are already covered in Chap. 3, here results from an extensive code comparison are presented, where 10 different codes, implementing different algorithms, are compared against each other.

Part IV: Finally, the thesis closes with an outlook on future developments in Chapter 7.

1.1 Canvas

In order to describe the dynamics of the universe, one first has to define a space–time manifold in which all the events will take place, similarly as a painter would first fix his canvas and only then start drifting colors. Observations show that the universe expands, so the space–time description has to take that into account; it is also assumed that on very large scales the universe appears homogeneous and isotropic, meaning there is no special point nor a preferred direction. In such a universe, the distance between two space–time points (ds) is given by the Friedmann–Lemaître–Robertson–Walker (FLRW) metric, which in spherical coordinates (r, θ, ϕ) can be written in the following form:

$$ds^2 = dt^2 - a(t)^2 \left(\frac{dr^2}{1 - \kappa r^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right) . \tag{1.1}$$

Here, κ describes the curvature, and is set by initial conditions, while $a(t)$ is the time–dependent scale factor. The metric leaves the choice of normalization free, and is usually normalized by setting the scale factor today to unity ($a(\text{now}) \equiv 1$).

Instead of time, cosmologists prefer a directly observable quantity – redshift (z), defined as a relative difference between the observed and emitted wavelength coming from some object:

$$z = \frac{\lambda_{\text{observed}} - \lambda_{\text{emitted}}}{\lambda_{\text{emitted}}} . \tag{1.2}$$

Redshift (or blueshift) may also arise from a relative motion of emitter and observer, as a Doppler shift. However, this kind of shift quickly becomes negligible as the distance between emitter and observer increases, and is dominated by the 'cosmological redshift' – the effect of wavelengths being stretched in the empty space as a consequence of space itself being expanded. In fact, in the Sloan Digital Sky Survey Release III, out of more than half million galaxies whose redshifts have been determined, fewer than a hundred exhibit blueshift, the rest being redshifted. Such a cosmological redshift (hereon simply redshift) relates to the scale factor as:

$$1 + z = \frac{a_{\text{now}}}{a_{\text{when emitted}}} \,. \tag{1.3}$$

With the normalization convention adopted here, *now* can therefore be annotated as $a = 1$, or $z = 0$.

1.1.1 Comoving Coordinates

According to general relativity one may use any desired coordinate system to formulate the laws of physics. Nevertheless, some choices are simpler than others, and the FLRW metric (eq. 1.1) has coordinates carefully chosen to make the symmetries of the space–time obvious. It is furthermore convenient to introduce *comoving* coordinates where space itself is static. The relation between comoving and physical coordinates is self-evident from equation (1.1):

$$\mathbf{r} \equiv \frac{\tilde{\mathbf{r}}}{a(t)} \,, \tag{1.4}$$

with tilde denoting the physical distance. Comoving frames are thus moving along with the Hubble flow as the universe expands. The velocity an object has relative to its comoving frame is called *proper* velocity, and is usually small for most cosmological considerations:

$$\dot{\mathbf{r}} = \frac{\dot{\tilde{\mathbf{r}}}}{a} - \mathbf{r}\frac{\dot{a}}{a} = \frac{\mathbf{u}}{a} - \mathbf{r}\frac{\dot{a}}{a} \,. \tag{1.5}$$

On the right-hand side, the first term ($\dot{\tilde{\mathbf{r}}} \equiv \mathbf{u}$) is the proper velocity, while the second one is the Hubble flow. The motion relative to the background universe is $\mathbf{v} = a\dot{\mathbf{r}}$.

Equations of motion in comoving coordinates read:

$$\frac{d\mathbf{r}}{dt} = \frac{\mathbf{v}}{a} \,,$$

$$\frac{d\mathbf{v}}{dt} = -\frac{1}{a}\nabla\Phi - \frac{\dot{a}}{a}\mathbf{v} \,, \tag{1.6}$$

where the source for the gravitational potential Φ is the fluctuating part of the mass density:

$$\nabla^2\Phi = 4\pi G a^2 \left[\rho(\mathbf{r}, t) - \rho_b(t)\right] \,. \tag{1.7}$$

1.1.2 Friedmann Equations

The relativistic analogs of Newton's law of gravity – Einstein's field equations, connect the mass–energy content in the universe to the geometry of the space–time; the 10 coupled, nonlinear, hyperbolic–elliptic partial differential equations in tensor form look deceptively simple:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \ . \tag{1.8}$$

The Einstein tensor ($G_{\mu\nu}$) describing geometry can be computed for the case of the FLRW metric, and it has two nontrivial components. The stress–energy tensor ($T_{\mu\nu}$) for an ideal fluid of density ρ and pressure p is

$$T_{\mu\nu} = \mathrm{dia}\left[\rho(t), -p(t)/c^2, -p(t)/c^2, -p(t)/c^2\right] \ . \tag{1.9}$$

Using this, the ten field equations can be reduced to the two Friedmann equations [1] [9]:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) \ , \tag{1.10}$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{\kappa c^2}{a^2} \ . \tag{1.11}$$

Combining these equations, one obtains the conservation of energy:

$$\dot{\rho}c^2 + 3\frac{\dot{a}}{a}(\rho c^2 + p) = 0 \ . \tag{1.12}$$

The Friedmann equations, together with the equation of state $p = p(\rho)$ can be used to determine the three functions of interest: $a(t)$, $\rho(t)$, and $p(t)$. Differentiating equation (1.11) and substituting into equation (1.10) we get

$$\frac{d}{dt}(\rho a^3) = -3a^2\dot{a}p \ . \tag{1.13}$$

This is the first law of thermodynamics for an expanding space $d(\rho a^3) = -pda^3$, and can be conveniently written as:

$$\frac{d}{da}(\rho a^3) = -3a^2 p \ . \tag{1.14}$$

For an ideal fluid, the equation of state takes the form $p = w\rho$, where p is a function of density and other state variables (like temperature). Equation (1.14) has to be satisfied separately for each component i that contributes to the total energy density of the universe (including a cosmological constant, or a homogeneous scalar field). In the case that the equation of state parameter does not evolve in time, $w_i = p_i/\rho_i = \text{const.}$, the density of each species will

[1] Some authors call both equations written here as 'Friedmann', while some refer by that name to only eq. (1.10).

evolve as:

$$\rho_i \propto a^{-3(1+w_i)} \ . \tag{1.15}$$

Knowing the values for w's is the heart of cosmology; no conclusions are possible, nor results derivable before assuming a certain equation of state. For ordinary matter we have laboratory experiments which, in the tested ranges of energy, confirm that matter we know indeed behaves as an ideal fluid. If non-relativistic, then $w = 0$, while for relativistic matter and radiation, $w = 1/3$. Some undetected forms of matter may also exist, for example, large scale scalar field(s), which also allow for $p = w\rho$ equation of state, and $w = 1$ (when kinetic energy dominated), or $w = -1$ (potential energy dominated). In general, this points to an important question: if observations do not match the model, is it because the model is flawed, or because of the existence of some yet undetected form of matter-energy? The answer is not always clear.

A very important quantity which appears in the Friedmann equation is

$$H(a) \equiv \frac{\dot{a}}{a} \ , \tag{1.16}$$

the time-dependant Hubble parameter. Its value today (H_0) is the Hubble constant, one of the most important numbers in cosmology. It indicates the rate at which the universe is expanding, and is often expressed through the dimensionless parameter h, such that $H_0 = 100h$ kms^{-1}Mpc^{-1}. In addition, cosmologists often use the deceleration parameter:

$$q(a) \equiv -\frac{\ddot{a}}{aH^2}, \tag{1.17}$$

and there were even some suggestions to reduce cosmology to "a search for two numbers" - H_0 and q_0 [10].

1.2 Content

The overall geometry of the universe depends on the total matter–energy content, as we can see from equation (1.8). While there is a continuum of mean densities which will lead to negative ('open' universe) or positive ('closed' universe) curvature, there is one special value which results in an exactly flat universe:

$$\rho_c = \frac{3H_0^2}{8\pi G} \approx 1.88 \times 10^{-29} h^2 \quad \mathrm{g\,cm^{-3}} \ . \tag{1.18}$$

This is called the critical density, and commonly different contributions to the energy density are represented in terms of it:

$$\Omega_m \equiv \frac{\rho_m}{\rho_c}, \quad \Omega_r \equiv \frac{\rho_r}{\rho_c}, \quad \Omega_\Lambda \equiv \frac{\Lambda}{3H^2}, \quad \Omega_\kappa \equiv \frac{-\kappa c^2}{a^2 H^2} \ , \tag{1.19}$$

$$\Omega_m + \Omega_r + \Omega_\Lambda + \Omega_\kappa = 1 \ .$$

The above corresponds to relative amounts of matter, radiation, cosmological constant, and curvature. Hereafter, when the above cosmological parameters are not explicitly written as a function of time, we refer to their values at the present time, e.g. $\Omega_m \equiv \Omega_m(z=0)$.

Each component listed in equation (1.19) evolves differently as the universe increases in size (following eq. (1.15)), and the Friedmann equation (eq. (1.11)) can be rewritten as

$$\frac{\dot{a}}{a} \equiv H^2(z) = H_0^2 \left[\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_\kappa(1+z)^2 + \Omega_\Lambda\right] , \qquad (1.20)$$

to directly show the time scaling. We can see that in the general case, the universe will evolve through phases when different components dominate, as they decay differently in time. We can also infer a few possible scenarios: if $\Omega_\Lambda < 0$ the universe will always recolapse. It will happen either because the cosmological constant will reverse the expansion once it becomes the dominant component, or because the amounts of matter or radiation were sufficiently high to stop the expansion themselves. If, on the other hand, the cosmological constant is positive, the universe will always expand unless there is enough matter or radiation to reverse the expansion before the cosmological constant prevails.

1.2.1 Current Data

Understanding the composition of the energy-density in the universe, as well as the relative amounts of different components, is crucial for understanding the history of the universe. It provides us with knowledge about the expansion rate (eq. (1.20)), which in turn determines the age of the universe, distance–redshift relation, statistical properties of the large scale structure, to name a few. In principle, any of the many cosmological probes (§1.3) can be used to constrain to some extent all cosmological parameters. Nevertheless, some observations naturally measure certain parameters with excellent accuracy, and those can be used as priors in the analysis of other observations.

Table 1.1 shows the parameters of the current 'standard' model. While the exact numbers differ from study to study, and are also sensitive to what parameters were fixed as priors, and which kept open, all the studies agree on the essence of the current model, often called the 'concordance' model to emphasize its robustness against different observations. In short, observations strongly favor a flat universe $\Omega = 1$, but only a few percent of the total content is in forms of matter and energy previously known. The rest is in forms neither experimentally detected nor theoretically predicted – cosmology provided the first indications for the existence of dark matter and dark energy.

There are excellent reviews on possible dark matter candidates [11, 12]; here we will just mention one important dynamical categorization scheme – 'hot' vs 'cold' dark matter. A dark matter candidate is called 'hot' if it is moving at

Table 1.1: Cosmological parameter values from Tegmark et al. (2006) [1].

Parameter	ΛCDM model	Flat model
Ω	1.003 ± 0.010	1 (fixed)
Ω_Λ	0.757 ± 0.021	0.757 ± 0.020
Ω_m	0.246 ± 0.028	0.243 ± 0.020
Ω_b	0.042 ± 0.002	0.042 ± 0.002
σ_8	0.747 ± 0.046	0.733 ± 0.048
n_S	0.952 ± 0.017	0.950 ± 0.016
h_0	0.72 ± 0.05	0.72 ± 0.03
q_0	-0.64 ± 0.03	-0.57 ± 0.1
w	-1 (fixed)	-0.94 ± 0.1
T_0 (K)	2.725 ± 0.001	2.725 ± 0.001
t_0 (Gyr)	13.9 ± 0.6	13.8 ± 0.2

relativistic speeds, that is, if most of its energy is kinetic energy. Similarly, it is 'cold' if it is moving non-relativistically, with its rest mass $m_{\rm dm}c^2$ dominating its energy budget. This categorization has important ramifications for structure formation, as inhomogeneities in the universe will evolve completely differently in the two scenarios. If dark matter is hot, it will erase perturbations on small scales via free-streaming processes, and structures will form through a top-down process, where most massive structures form first, and then fragment into smaller objects. On the other hand, in the cold dark matter (CDM) case, the smallest structures are the first to gravitationally collapse, and progressively larger structures form through mergers of smaller objects. While observations of large-scale structure exclude cosmologies where most of the dark matter is relativistic, it is still viable that a small fraction is 'warm'. The work here will consider only the CDM scenario, and it should be emphasised that the mass scales considered here are not affected by the possible presence of warm dark matter.

1.2.2 Dark Energy

In 1998, cosmic acceleration was discovered [13, 14], and since then the question what drives it remains one of the most striking problems in physics. Since all known forms of matter and radiation (as well as those still unknown but anticipated – dark matter) can only decelerate the expansion of the universe, a new and radical solution is required. All possible solutions are currently dubbed *dark energy*, even though the origin of acceleration might not necessarily be a contributor to the total energy density of the universe. The simplest model, and the one to which the current observations converge, is a cosmological constant, first introduced almost a century ago by Albert Einstein [15]. Originally, its sole purpose was to make the universe static, which was aesthetically pleasing in the years before the discovery of Hubble's law in 1929 (and even before the

discovery of galaxies which happened in 1924).

Assuming the universe consists only of pressureless matter (dark matter) and dark energy in the form of a cosmological constant, we can rewrite the Friedmann equation (1.10) as:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\rho_m + \frac{\Lambda}{3} \,, \tag{1.21}$$

which gives us the effective force law

$$f = -\frac{GM}{a^2} + \frac{\Lambda}{3}a \,. \tag{1.22}$$

We see that the cosmological constant enters as a 'repulsive gravity', and its intensity increases with distance. Clearly, if dominant, the Λ term leads to an accelerated expansion of the universe.

The above considerations can be generalized to dark energy, for the moment defined via $w \neq -1$. Analyzing the Friedmann equation further, we can define the condition for acceleration (again, assuming pressureless matter and $\Omega_m + \Omega_\Lambda = 1$) as:

$$\sum_i (\rho_i + 3p_i) = \sum_i \rho_i(1 + 3w_i) < 0 \,, \tag{1.23}$$

which leads to

$$w < \frac{1}{3(1 - \Omega_m)} \,. \tag{1.24}$$

For the approximate value $\Omega_m = 1/3$ (table 1.1) cosmic acceleration requires $w \leq -1/2$, violating the strong energy condition ($w \geq -1/3$). Of course, it is also possible for any value of Ω_m, to find Ω_Λ which will exactly result in $\dot{a} = 0$ and $\ddot{a} = 0$ – a static solution – which is why Einstein introduced Λ in the first place.

One important issue, making theorists feel uncomfortable about dark energy is the fine tuning problem. If we ask when did the transition between deceleration and acceleration occur ($\ddot{a} = 0$), we find the condition:

$$(1 + z_{\text{eq}})^{-3w} = -(1 + 3w)\frac{\Omega_\Lambda}{\Omega_m} \,. \tag{1.25}$$

For the measured values of the cosmological parameters (see table 1.1) this leads to $z_{\text{eq}} \simeq 0.7$ – fairly close to today. Having in mind that both matter and radiation density are rapidly decreasing in time (see eq. (1.15)), the initial condition for ρ_Λ/ρ_m at the epoch when inflation ends has to be tuned to parts in $\sim 10^{-81}$ in order to become close to unity today. Moreover, the fact that acceleration is a relatively recent phenomenon is confirmed by supernova observations which favor $z_{\text{eq}} \simeq 0.5$ [16]. Note also that the coincidence problem cannot be avoided with $w \neq -1$, but only if the equation of state parameter is time-dependant $w = w(z)$.

Finally, the most bizarre possibility should be mentioned – dark energy vi-

olating even the weak energy condition ($w \geq -1$). Due to Caldwell [17] this option is named *phantom energy*. Equation (1.15) shows that the dark energy density will grow to infinity for the case of phantom energy. Also, if we write the evolution of the Hubble parameter for the case of interest (flat universe, negligible radiation):

$$H^2(z) = H_0^2 \left[\Omega_m(1+z)^3 + \Omega_\Lambda(1+z)^{3(1+w)}\right] , \tag{1.26}$$

we immediately see that the case $w < -1$ leads to an infinite expansion rate, achieved in a finite time. At that point, our notion of space–time fails.

In summary, there is not yet a clear idea what dark energy might be, and it's not even clear if it is really a form of energy or something else. Currently 'dark energy' is just another way to state that the expansion of the universe accelerates. An excellent review on dark energy candidates is written by Weinberg [18].

1.3 Measurables

It is of course, possible to develop numerous thoeries about the content and structure of the universe. In the end, we have to use observations to constrain – and rule out – competing theoretical models. All cosmological tests can roughly be reduced to effective measurements of two quantities: the expansion history – $H(z)$ (eq. 1.16), and the linear growth factor – $D_+(z)$ (which will be defined later in §2). Therefore, the methods in the first category probe the homogeneous universe, while the second group characterizes how inhomogeneities evolve.

1.3.1 Expansion History Probes

The most prominent and accurate $H(z)$ probe nowadays is obtained from observations of type Ia supernovae [19, 20, 13, 14, 21]. By measuring the incoming flux $\mathcal{F}$ from a source of known intrinsic luminosity L, one can infer the luminosity distance d_L to the object. This distance bears a cosmology dependence:

$$d_L(z) \equiv \sqrt{\frac{L}{4\pi\mathcal{F}}} = (1+z)r(z) , \tag{1.27}$$

where $r(z)$ is the comoving distance to an object at redshift z,

$$\begin{aligned} r(z) &= \int_0^z \frac{dz'}{H(z')} = \int_{1/(1+z)}^1 \frac{da}{a^2 H(a)} \qquad (\kappa = 0) , \\ r(z) &= |\kappa|^{-1/2} \chi\left[|\kappa|^{1/2} \int_0^z \frac{dz'}{H(z')}\right] \qquad (\kappa \neq 0) , \end{aligned} \tag{1.28}$$

and where $\chi(x) = \sin(x)$ for $\kappa > 0$ and $\sinh(x)$ for $\kappa < 0$. In particular, for a flat universe and constant w (see eq. 1.20):

$$r(z) = \frac{1}{H_0} \int_0^z \frac{dz'}{\sqrt{\Omega_m(1+z')^3 + \Omega_\Lambda(1+z')^{3(1+w)} + \Omega_r(1+z')^4}} \,. \tag{1.29}$$

As a result, measurements of supernova magnitudes can be efficiently used to constrain cosmological parameters.

Similarly to the luminosity distance, if we know the transverse physical size D of an object, or physical phenomena, the angle subtending D at some redshift z is:

$$\theta = \frac{D}{d_A(z)} \,. \tag{1.30}$$

Here, $d_A(z)$ is the angular diameter distance

$$d_A(z) = \frac{d_L(z)}{(1+z)^2} \propto \int_0^z \frac{dz'}{H(z')} \tag{1.31}$$

(following equations (1.27) and (1.28)). Measuring angular diameter distance to a 'standard ruler', thus provides the same leverage over cosmology as standard candles.

Perturbations that exist in the early universe excite sound waves in the photon-baryon fluid [22, 23], called 'baryon acoustic oscillations' (BAO). The most prominent first peak defines a standard ruler whose length is equal to the distance sound can travel before baryons decouple from photons. This sound horizon scale is

$$s = \int_{z_d}^{\infty} \frac{c_s}{H(z)} dz \,. \tag{1.32}$$

As the cosmic microwave background (CMB) as well as the Big Bang nucleosynthesis, constrains the baryon to photon ratio to great precision, the sound speed is also known and given by [24, 25]:

$$c_s = \frac{1}{\sqrt{3}} \left(1 + \frac{3}{4}\frac{\rho_b}{\rho_\gamma}\right)^{-1/2} \,. \tag{1.33}$$

Moreover, measurements of the peaks in the CMB angular power spectrum constrain the physical scale of the sound horizon to a high precision. For example, WMAP 5 year data suggest $s = 146.8 \pm 1.8$Mpc [26]. Looking at the distribution of luminous red galaxies in the Sloan Digital Sky Survey, Eisenstein et al. [27] were able to detect the BAO peak, whose height is consistent with the universe containing 30% of matter and 70% dark energy.

1.3.2 Growth of Structure

Cosmological probes described above (SN Ia, BAO) measure expansion of the universe, and they find that the expansion is accelerating. The reason for the

acceleration might be due to the existence of dark energy, or might indicate that our understanding of gravity on very large scales is incomplete [2]. Measurements of the expansion rate alone cannot provide the answer to this, but in combination with measurements of the growth of structure, the degeneracy can be broken. Weak lensing [28, 29, 30, 31], for example, is sensitive to both the geometry of the universe and the growth of structure.

For any theory of gravity, one can determine the relation between the growth rate and the expansion rate (§2.2.1). Independent measurements of both can then confirm or reject the theory. The most promising path to measuring the growth factor is via the most massive gravitationally collapsed structures in the universe – clusters of galaxies. As they require billions of years to assemble, they are strongly influenced by the cosmic expansion history, but are also very sensitive to the amount of matter Ω_m, and the amplitude of density fluctuations σ_8. Thus, their statistical measures – their mass distribution and spatial correlations – can be powerfull probes of cosmology. Moreover, as clusters form in the epoch when dark energy is dominant, they are exponentially sensitive to the effects of the recently observed acceleration of the Universe.

This thesis presents theoretical and numerical results on statistics of very rare overdense structures which correspond to clusters at the current epoch. It is said 'correspond' rather than 'they are', as here are analyzed dark matter only structures. While the dark matter is indeed the dominant component (see table 1.1) that gravitationally clusters, its evolution is different from that of baryons. It does not radiate energy, and thus once virialized, dark matter halos cannot contract further. Baryons do cool on the other hand, and therefore at smaller and smaller scales the results from dark matter only simulations will be more and more inaccurate. For that reason, we focus in this thesis on the very large structures. Still, the clear downside is that connecting this theoretical framework to the observations is not always straightforward, as in general, dark matter structures are not directly observable. The most important methods for measuring growth of structure via galaxy clusters are the mass function, as well as the amplitude and the shape of the power spectrum.

The Mass Function

The distribution of masses in the universe is termed the mass function and constitutes one of the most important probes of cosmology. At low redshifts, $z \leq 2$, the mass function at cluster scales (high-mass end) is exponentially sensitive to variations in cosmological parameters, such as the matter content of the Universe Ω_m, the dark energy content along with its equation-of-state parameter, w [32], and the normalization of the primordial fluctuation power spectrum,

[2] Historically, it would not be the fist time that measurements lead to an alteration of the current theory of gravity. After successfuly predicting that Uranus's orbit is altered by an invisible planet (Neptune), Le Verrier interpreted variations in Mercury's orbit as caused by a dark planet Vulcan. The planet was never found, but in 1915 Einstein extended Neewton's gravity, which accurately accounted for Mercury's perihelion shift.

σ_8. This is especially interesting today, as several observational campaigns are starting to map clusters (Planck [33], Atacama Cosmology Telescope [34], South Pole Telescope [35]).

In this thesis we investigate the halo mass function – the number density of dark matter halos. We provide accurate predictions for the mass function from numerical simulations, covering 7 orders of magnitude in mass, and a wide range of redshifts ($z \in [20 - 0]$). While the mass function at high redshifts naturally does not include clusters, it does include very rare density peaks, which at a given epoch are dynamically analogous to clusters of galaxies today. Since it is difficult to imagine baryonic process which would disrupt an object with mass $M \geq 10^{14} M_{\odot}$, the number density of halos is an excellent approximation to the number density of clusters. The major problem which arises on the intersection between theory and observations is defining and measuring masses. Chapter 4 addresses some of the issues, and determines relations between some of the commonly used mass definitions.

The Power Spectrum

The power spectrum $P(\mathbf{k})$, or its Fourier space analog, the two-point correlation function $\xi(\mathbf{r})$, is one of the main measures of large-scale structure. The amplitude of the power spectrum cannot be predicted by theory, and is empirically determined. Commonly, the normalization is set by σ_8 – the root mean square of density fluctuations in $8\,\mathrm{h}^{-1}$Mpc spheres [3]. The mass corresponding to that scale is $\sim 5 \times 10^{14} M_{\odot}$, which is the mass of a rich galaxy cluster. Roughly, $\sigma_8 \sim 1$, and thus stands (again very roughly) in between nonlinear, smaller scales, and linear scales, larger than $8\,h^{-1}$Mpc. Here, esspecially in Chapter 4 we will analyze how statistical properties of dark matter halos change with σ_8, and we will present a new way to measure it, using statistics of merging objects.

The second interesting property is the shape of the power spectrum. As it is possible to make theoretical predictions on what it should be, its measurements can provide insight in both the early universe, and the subsequent evolution of perturbations. Inflation, the most promising theory for the creation of density fluctuations (see §2.1) fixes the shape of the 'primordial' power spectrum. Its evolution can be predicted for any cosmology of interest, and is usually given through the transfer function $T(\mathbf{k}, t)$, such that

$$P(\mathbf{k}, t) = T(\mathbf{k}, t)^2 P_{\mathrm{in}}(\mathbf{k}) \ . \tag{1.34}$$

The effects of evolution are thus scale-dependent, and two important scales will be imprinted on the final power spectrum: the sound horizon size at decoupling which provides a standard ruler (§1.3.1), and the horizon size at matter radiation equality. The density fluctuations on scales smaller than the horizon are suppressed when the radiation is dominant, while those larger than the horizon

[3] For a more precise definition of σ, see §2.3.1.

grow linearly (§2.1). As a result the power spectrum will have a global maximum at $k = 2\pi/R_{\rm eq}$, and this value will strongly depend on the matter content in the universe, as higher values for Ω_m result in an earlier epoch of equality.

A practical problem is that the matter power spectrum is not directly observable. As one can only observe spatial correlation for a certain class of objects (clusters, galaxies, hydrogen clouds...) knowing how such objects are 'biased' with respect to the overall matter is crucial. It turns out that more massive objects are more biased, meaning that their power spectrum is more amplified [36, 37]. In Chapter 5 the mass dependence of the bias is analyzed, and an accurate analytical fit is given. As before, the results are obtained from dark matter simulations, thus they cannot be directly used for statistics of galactic objects, as (not yet understood) galaxy formation highly depends on baryonic processes. However, the most massive dark matter structures must resemble clusters very closely, and this is particularly important in the light of upcoming surveys such as NORAS2 and REFLEX2 which together will map ~ 1800 clusters in the local universe.

2 Formation of Structure

2.1 Evolution of Structure

The most promising theory of the early universe is inflation [38]. Apart from solving the flatness, horizon and relics problems, it provides a mechanism for producing initial density fluctuations. These fluctuations will serve as seeds for future structures, and effectively they provide initial conditions for any structure formation research.

During inflation, the quantum fluctuations in the inflaton field ϕ are exponentially stretched due to the rapid expansion. The proper wavelength of the fluctuations are stretched out of the Hubble-horizon scale at the time, H^{-1}. Once outside the horizon, the characteristic r.m.s. amplitude of these fluctuations is $\sigma_\phi \ll H/(2\pi)$. As causal physics cannot affect perturbations outside the horizon, they become 'frozen', and simply linearly grow in time. Different inflation models result in different perturbation spectra [39], and in general have to be calculated numerically [40]. Still, most inflation models result in the following perturbation spectrum:

$$|\delta(k)| \propto k^{1-\epsilon} , \tag{2.1}$$

where ϵ is the tilt. The spectrum is therefore expected to be very close to the scale-invariant (or Harrison-Zel'dovich) power spectrum – $P(k) \propto k^1$. The amplitude of the perturbations depends on the inflaton potential at the time when the scale crossed the Hubble radius during inflation. Due to the unknown nature of the potential, inflation does not make accurate predictions of the normalization of the above amplitudes.

As a result, inflation defines initial conditions for the standard big-bang model. These initial conditions result in a universe which is flat, homogeneous, and isotropic on large scales, and also imprints the perturbations on the metric which produce all the observable structure. After inflation (followed by reheating), the standard scenario of gravitationally driven growth of structures starts. As the universe decelerates, at some point the fluctuations re-enter the Hubble horizon, and seed gravitational instabilities in the universe. As perturbations of virtually any scale will be larger than the horizon at early enough times, its evolution has to be, at least initially, described by general relativity.

In the linearised approximation, perturbations of the metric are assumed to

be small $h_{ij}h^{ij} \ll 1$, and Einstein's equations can be linearized to

$$\mathcal{L}(g_{\alpha,\beta})\delta g_{\alpha,\beta} = \delta T_{\alpha,\beta} \ . \tag{2.2}$$

The main difficulty in this approach is that, depending on the choice of coordinates (gauge) $\delta T_{\alpha,\beta}$ can be arbitrarily large, and even generate additional quantities, which would be absent for different gauge choice. In general, if all the quantities (and the coordinate system) are well motivated, and physically interpretable, the gauge artifacts will be present; if on the other hand, a convenient transformation of physical variables is used, Einstein's equations can be written in gauge invariant form [41]. The problem there is that those transformed quantities usually do not have a good physical interpretation.

The linearized treatment is developed in several works [42, 43, 44]; here, only some important results will be noted. The amplitude of perturbations larger than the horizon will always grow, and the growth will be more rapid in the epoch when radiation dominates (before matter-radiation equality, $t_{\rm eq}$ time):

$$\left(\frac{\delta\rho}{\rho}\right) \propto \begin{cases} a^2; & t < t_{\rm eq} \\ a; & t > t_{\rm eq} \end{cases} \ . \tag{2.3}$$

Modes which are inside the horizon can grow, but but can also be supressed either due to pressure forces, or because of the expansion of the universe. In the radiation dominated phase, the expansion timescale is shorter than the gravitational collapse timescale; thus no perturbations on scales smaller than the horizon can grow. After the matter-radiation equality epoch, when matter becomes the dominant component, fluctuations in dark matter resume growth. Thus, dark matter perturbations inside the horizon evolve as:

$$\left(\frac{\delta\rho}{\rho}\right)_{\rm DM} \propto \begin{cases} const.; & t < t_{\rm eq} \\ a; & t > t_{\rm eq} \end{cases} \ . \tag{2.4}$$

For baryons the evolution is somewhat more complicated as they interact by pressure forces as well. The epoch of baryon-radiation equality is $a \approx 1/(4 \times 10^4 \Omega_b h^2)$, which is for $\Omega_b h^2 > 0.026$ earlier than the decoupling epoch. Thus, for the case of interest, even after $t_{\rm eq}$ for baryons, they will still interact with photons via Thomson scattering. Due to the large sound speed in the photon-baryon fluid – $\sim 1/3$ speed of light – the pressure will provide support against the collapse. Thus the evolution will look similar as in the dark matter sector, but with a time offset:

$$\left(\frac{\delta\rho}{\rho}\right)_b \propto \begin{cases} const.; & t < t_{\rm decoupling} \\ a; & t > t_{\rm decoupling} \end{cases} \ . \tag{2.5}$$

As the perturbations in dark matter resume growth before the baryons, they will form potential wells. Once baryons decouple, and lose pressure support,

they will quickly be driven inside those potential wells. As a result, right after decoupling baryon inhomogeneities will rapidly grow until their perturbations become the same as for the dark matter. From there on, their growth will be indistinguishable from the dark matter growth, as long as we consider scales much larger than Jeans scale for baryons. This will be discussed later in more detail. When the length scale of the perturbations is much smaller than the Horizon scale $\lambda_H \simeq 2ct$, then a Newtonian treatment of the subject is valid, as relativistic effects due to the curvature of space-time are negligible.

2.2 Linear Evolution

When treating dynamics in the Newtonian limit, matter can be approximated as a collection of identical particles of mass m, moving in the gravitational potential Φ. In classical mechanics, a complete description of such a system is given through the number of particles residing in a small volume d^3r centered on $\mathbf{r}$, whose velocities are confined to d^3u, around $\mathbf{u}$:

$$dN = f(\mathbf{r}, \mathbf{u}, t) d^3r d^3u \ , \tag{2.6}$$

where $f(\mathbf{r}, \mathbf{u}, t)$ is the phase-space density. Evolution of the system is described by the Boltzmann equation:

$$\frac{\partial f}{\partial t} + \dot{\mathbf{r}} \nabla f - \nabla \Phi \frac{\partial f}{\partial \dot{\mathbf{r}}} = \left. \frac{\partial f}{\partial t} \right|_{\text{coll}} \ . \tag{2.7}$$

In the limit where the gravitational potential can be considered smooth, that is, when the mean free path of a particle is very long (compared to the characteristic scale of the problem), the collisional term on the right-hand side of equation (2.7) can be neglected. We are interested in solving the Boltzmann equation in comoving coordinates where:

$$\frac{\partial f}{\partial t} + \frac{\mathbf{v}}{a} \nabla f - \frac{1}{a} \nabla \Phi \frac{\partial f}{\partial \mathbf{v}} = 0 \ . \tag{2.8}$$

This collisionless Boltzmann, or Vlasov, equation has to be solved self-consistently together with the Poisson equation:

$$\nabla^2 \Phi(\mathbf{r}, t) = 4\pi G a^2 \rho(\mathbf{r}, t), \tag{2.9}$$

where

$$\rho(\mathbf{r}, t) \equiv \int f(\mathbf{r}, \mathbf{v}, t) d^3v \ . \tag{2.10}$$

The analysis of the Vlasov equation can be performed by taking its moments[1] [45, 46]. The N-th moment is obtained by multiplying the Vlasov equation by $\mathbf{v}^{N-1}$, and integration over velocity space. The first two moments of the Vlasov

[1] f itself cannot be obtained through observations, but it is possible to measure its moments.

equation are analogous to the continuity and Euler equations in ordinary fluids, but where the velocity dispersion plays the role of effective pressure. However, taking moments alone cannot close the system, as every N-th moment will introduce a new variable. In fluid dynamics, the equation of state is used for closure, describing material properties through the relation between ρ and p (and possibly other thermodynamical quantities). With collisionless fluids, one is forced to either truncate in some way this regression, or to use some knowlegde of the σ_{ij} for the particular case under consideration.

In the case when the mean free path of a particle is short, matter can be described as an ideal fluid. The description of the ideal fluid consists of mass and momentum conservation equations [47]. The continuity equation in comoving coordinates is [45]:

$$\frac{\partial \rho}{\partial t} + 3H\rho + \frac{1}{a}\nabla(\rho \mathbf{v}) = 0 \ , \tag{2.11}$$

while the Euler equation is:

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{a}(\mathbf{v}\cdot\nabla)\,\mathbf{v} + H\mathbf{v} = -\frac{1}{a\rho}\nabla p - \frac{1}{a}\nabla\Phi \ . \tag{2.12}$$

The density contrast at a given point can be defined as:

$$\delta(\mathbf{r},t) = \frac{\rho(\mathbf{r},t) - \rho_b(t)}{\rho_b(t)} \ , \tag{2.13}$$

where ρ_b includes matter only, as it is the only component of matter-energy which clusters, and therefore $\rho_b \propto a^{-3}$. Furthermore, we can write the pressure as $p = p_0 + \delta p$, and if the pressure is a function of density alone, then:

$$\nabla p = p(\rho_b) + c_s^2 \rho_b \delta \ , \tag{2.14}$$

where c_s is the speed of sound: $c_s^2 = dp/d\rho$. Moreover, $\mathbf{v}$ can also be considered a perturbation, as $\mathbf{v} = \mathbf{v}_0 + \delta\mathbf{v}$, but we can always choose an inertial reference frame where $\mathbf{v}_0 = 0$. Similarly, Φ is already a perturbed potential (see equation (1.7)). Substituting these into equations (2.11) and (2.12), and considering small perturbations where $\delta \cdot \mathbf{v}$ and v^2 terms can be dropped, we find the following system of equations:

$$\begin{aligned} &\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla\mathbf{v} = 0 \\ &\frac{\partial \mathbf{v}}{\partial t} + H\mathbf{v} + \frac{1}{a\rho_b}\nabla p + \frac{1}{a}\nabla\Phi = 0 \ , \end{aligned} \tag{2.15}$$

as $\dot{\rho}_b = -3H\rho_b$. We can take the time derivative of the first equation, and multiply it by $-a$; also take the divergence of the second equation. After adding the two we obtain one equation for the density perturbations:

$$\frac{\partial^2 \delta}{\partial t^2} + 2H\frac{\partial \delta}{\partial t} = \frac{1}{a^2\rho_b}\nabla^2 p + \frac{1}{a^2}\nabla^2\Phi \ . \tag{2.16}$$

Velocity was completely eliminated via $\nabla \mathbf{v} = -a\dot{\delta}$. Writing the Poisson equation as

$$\nabla^2 \Phi = 4\pi G \rho_b a^2 \delta \ , \tag{2.17}$$

we end with

$$\frac{\partial^2 \delta}{\partial t^2} + 2H\frac{\partial \delta}{\partial t} = \frac{1}{a^2 \rho_b}\nabla^2 p + 4\pi G \rho_b \delta \ . \tag{2.18}$$

First, we try to find a stable solution where the pressure gradient can stop the runaway gravitational collapse. The pressure gradient is

$$\nabla p = \frac{dp}{d\rho}\nabla \rho = c_s^2 \rho_b \nabla \delta \ , \tag{2.19}$$

resulting in

$$\frac{\partial^2 \delta}{\partial t^2} + 2H\frac{\partial \delta}{\partial t} = \frac{c_s^2}{a^2}\nabla^2 \delta + 4\pi G \rho_b \delta \ . \tag{2.20}$$

As none of the coefficients with δ or its derivatives depends on $\mathbf{r}$, we can try to find the solution in the form of plane waves:

$$\delta = A(t)\exp(i\mathbf{k}\cdot\mathbf{r}) \ . \tag{2.21}$$

Substituted into eq. (2.20), it shows us how the amplitudes of different k modes evolve:

$$\ddot{A}_k + 2H\dot{A}_k = \left[4\pi G\rho_b - \frac{c_s^2 k^2}{a^2}\right] A_k \tag{2.22}$$

This is the standard equation for the fluid instability, except for the second term on the left-hand side, called 'Hubble drag', which acts as a friction term. The gravitational source term vanishes for the wavenumber

$$k_J = \left(\frac{4\pi G\rho_b}{c_s^2}\right)^{1/2} \ , \tag{2.23}$$

which defines the Jeans wavelength $\lambda_J = 2\pi a/k_J$. Wavelengths shorter than this will oscillate, as pressure can provide support against gravity. Neglecting the Hubble drag for a moment, the solution for the amplitudes become $A \propto \exp(-i\omega t)$, resulting in a dispersion relation:

$$\omega^2 = -\omega_J^2 + k^2 c_s^2 \ . \tag{2.24}$$

Modes of wavelengths larger than the Jeans wavelength, will be unstable: in standard fluid dynamics they exponentially grow, but in the expanding universe Hubble drag slows them down into power-law behaviour as will be demonstrated in the following section.

2.2.1 Linear Growth Factor

In the limit of very long wavelengths, $\lambda_J \ll \lambda$, equation (2.22) becomes the same as for a pressureless fluid. The linear growth factor for that case can be calculated from

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\rho_b\delta \ . \tag{2.25}$$

By using the chain rule, it is straightforward to switch from time as independent variable, to the scale factor:

$$\dot{\delta} \equiv \frac{\partial\delta}{\partial t} = \dot{a}\frac{\partial\delta}{\partial a} \ , \tag{2.26}$$

and similarly

$$\ddot{\delta} = \frac{\partial}{\partial t}\left(\dot{a}\frac{\partial\delta}{\partial a}\right) = \dot{a}^2\frac{\partial^2\delta}{\partial a^2} + \left(H\dot{a} + a\dot{a}\frac{\partial H}{\partial a}\right)\frac{\partial\delta}{\partial a} \ . \tag{2.27}$$

The term in brackets in the above equation comes from

$$\ddot{a} \equiv \frac{\partial\dot{a}}{\partial t} = \frac{\partial}{\partial t}(Ha) = H\dot{a} + a\dot{a}\frac{\partial H}{\partial a} \ . \tag{2.28}$$

Equation (2.25) thus reads:

$$\dot{a}^2\frac{\partial^2\delta}{\partial a^2} + \left(3H\dot{a} + a\dot{a}\frac{\partial H}{\partial a}\right)\frac{\partial\delta}{\partial a} = 4\pi G\rho_b\delta \ . \tag{2.29}$$

It can further be transformed into

$$\frac{\partial^2\delta}{\partial a^2} + \left(\frac{3}{a} + \frac{\partial \ln H}{\partial a}\right)\frac{\partial\delta}{\partial a} = \frac{3\Omega_m H_0^2}{2a^5H^2}\delta \ , \tag{2.30}$$

as smoothly distributed components do not cluster and ρ_b is simply:

$$\rho_b(a) = \Omega_m(a)\frac{3H_0^2}{8\pi G} \ , \tag{2.31}$$

and $\Omega_m(a) = \Omega_m a^{-3}$. Currently, the observations point to the cosmology where $\Omega \approx \Omega_m + \Omega_\Lambda$, where Ω_Λ might be the cosmological constant, but also an arbitrary form of dark energy. In general, equation (2.30) cannot be solved analytically, as dark energy has an undetermined equation of state parameter w, which might also be time-dependent.

One interesting case is $\Omega_m = \Omega = 1$, as matter was dominating the density content throughout a significant fraction of the cosmic history. There, the Hubble parameter is simply $H = H_0\Omega_m^{1/2}a^{-3/2}$, and eq. (2.30) becomes:

$$\frac{\partial^2\delta}{\partial a^2} + \frac{3}{2a}\frac{\partial\delta}{\partial a} - \frac{3}{2a^2}\delta = 0 \ . \tag{2.32}$$

The density contrast δ is the sum of two solutions to this equation, one that

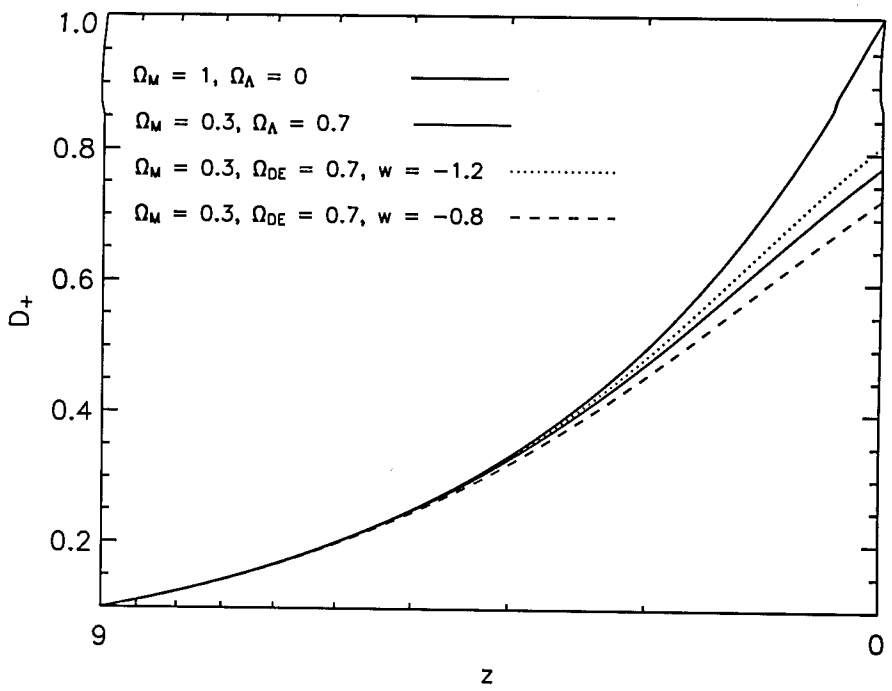

Figure 2.1: Linear growth factors for four different cosmologies – CDM, ΛCDM, and two CDM + dark enery with constant w, as a function of redshift ($a = (1+z)^{-1}$).

grows in time, and one that decays:

$$\delta = A(\mathbf{r})D_+ + B(\mathbf{r})D_-$$

$$D_+ \propto a \ , \quad D_- \propto H \propto a^{-3/2} \ . \tag{2.33}$$

The more interesting solution is naturally the growing one, D_+, as it drives gravitational instabillity over time.

Another interesting solution is when dark energy is the cosmological constant, that is, $w = -1$, and $H = H_0 \left(\Omega_m a^{-3} + \Omega_\Lambda\right)^{1/2}$. The decaying mode is again $D_- \propto H$, but the growing mode differs. The easiest way to find a growing mode is to look for a solution of the form $D_+ = uH$, and equation (2.30) as a function of u is [48]:

$$\frac{\partial^2 u}{\partial a^2} + 3\left(\frac{1}{a} + \frac{\partial \ln H}{\partial a}\right)\frac{\partial u}{\partial a} = 0 \ . \tag{2.34}$$

This equation can be reduced to first order in u', and solved by simple integration as:

$$\frac{\partial u}{\partial a} \propto (aH)^{-3} \ , \tag{2.35}$$

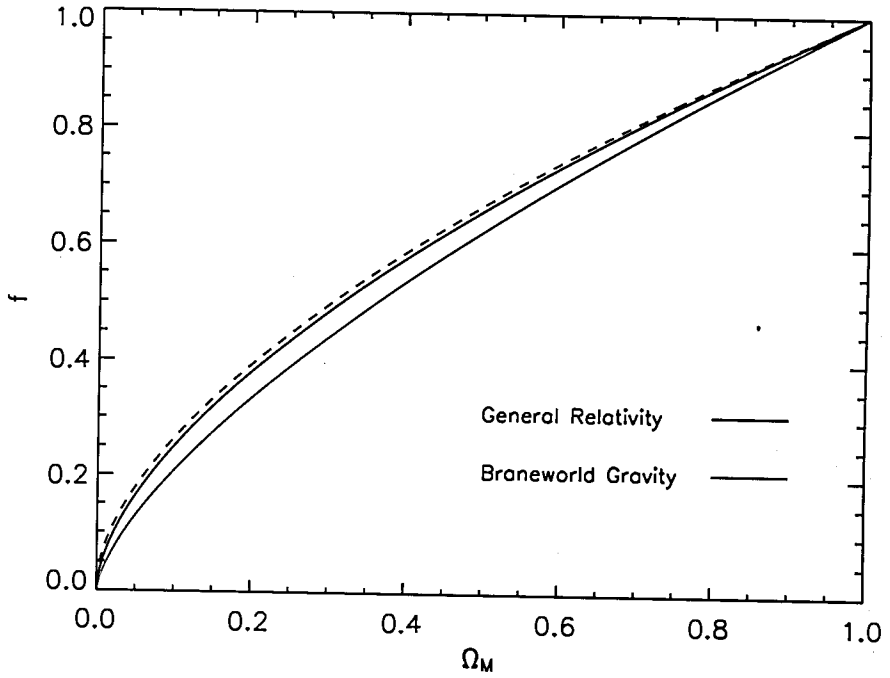

Figure 2.2: Linear growth rate in general relativity (black, solid for $\Omega = \Omega_m$, dashed for $\Omega = \Omega_m + \Omega_\Lambda$), and in braneworld gravity (red), where the late time acceleration emerges from the gravity sector rather than from the matter-energy content.

giving a growth factor

$$D_+(a) = \frac{5\Omega_m}{2}\frac{H}{H_0}\int_0^a \frac{da'}{[a'H(a')/H_0]^3}\,, \tag{2.36}$$

where the proportionality constant in front of the integral is obtained by requiring that for $a \ll 1$ the growth factor should be the same as for a matter dominated universe.

When $w \neq -1$, one has to numericaly solve equations (2.30) and (1.26) for desired values of Ω_m and w[2]. Examples of linear growth factors for four different flat cosmologies are presented in the figure 2.1: matter dominated, matter + cosmological constant, matter + dark energy with two different w's. As expected, dark energy, which effectively enters as repulsive gravity, slows down the formation of structures.

For many studies (peculiar velocity field, redshift distortions...) it is convenient to introduce the linear growth rate

$$f \equiv \frac{d\ln D}{d\ln a}\,. \tag{2.37}$$

If gravity is described by general relativity, f is, in the matter dominated

[2]It is convenient to first reduce the second order equation (2.30) into two first order differential equations.

universe, well fitted by [45] $f \approx \Omega_m^{0.60}$. A cosmological constant does not change this relation significantly [49]:

$$f \approx \Omega_m^{0.60} + \frac{1}{70}\Omega_\Lambda \left(1 + \frac{1}{2}\Omega_m\right) , \tag{2.38}$$

and for all practical purposes depends on the matter content only. Even more convenient is that in alternative theories of gravity, the growth rate still preserves its functional dependence, but with a different exponent. In braneworld gravity [50] for example, the linear growth rate is given by $f \approx \Omega_m^{0.68}$ [51, 52, 53]. Figure 2.2 shows the difference between these two theories. Also, we see that the difference between CDM and ΛCDM is very small in general relativity.

2.3 Spectrum of Perturbations

The disadvantage of treating perturbations in real space is that if we divide it into small volumes, their evolution will not be independent as gravity will move matter from one volume element into another. It is therefore more convenient to analyze the evolution of structure in Fourier space. Considering a volume V, with length L on a side – which is much greater than the 'homogeneity' scale – we can expand the perturbation field as:

$$\delta(\mathbf{r}) = \sum_{\mathbf{k}} \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{\mathbf{k}} \delta^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}} . \tag{2.39}$$

The components of the wavevector $\mathbf{k}$ are $k_x = 2\pi n_x/L$, $k_y = 2\pi n_y/L$, $k_z = 2\pi n_z/L$, where n_x, n_y, n_z are integer numbers. The fourier components $\delta(\mathbf{k})$ are

$$\delta(\mathbf{k}) = \frac{1}{V}\int \delta(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d^3r . \tag{2.40}$$

As $\delta(\mathbf{r})$ has to be real, $\delta^*(\mathbf{k}) = \delta(-\mathbf{k})$. If the density field on different locations is uncorrelated (or weakly correlated), then the integral in equation (2.40) becomes a sum over a large number of random variables. According to the central limit theorem, the probability distribution of $\delta(\mathbf{k})$ is then Gaussian:

$$\delta(\mathbf{k}) \propto \exp\left(-\frac{\delta^2}{\sigma^2}\right) . \tag{2.41}$$

As the mean value of perturbations $\langle\delta\rangle = 0$ by definition, the variance σ^2 is:

$$\sigma^2 = \langle\delta^2\rangle = \frac{1}{V}\sum_{\mathbf{k}} \delta^2(k) , \tag{2.42}$$

and it has to be a function of $k \equiv |\mathbf{k}|$ only because of the isotropy of the universe.

In the continuum limit ($V \to \infty$) it becomes

$$\sigma^2 = \frac{1}{(2\pi)^3} \int P(k) d^3k = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) dk \ . \tag{2.43}$$

In the case of a Gaussian random field, the power spectrum

$$P(k) = \langle \delta(\mathbf{k}) \delta^*(\mathbf{k}) \rangle = \delta^2(k) \tag{2.44}$$

contains the complete statistical characterisation of the perturbation field. As the Gaussian random field is ergodic, the ensemble average in k space is equal to the spatial average in real space, thus

$$\sigma^2 = \left\langle \left(\frac{\delta\rho}{\rho} \right)^2 \right\rangle \ . \tag{2.45}$$

This definition of variance accounts for perturbations on all scales, and can diverge for certain forms of $P(k)$.

2.3.1 Mass Variance

Variance can also be considered on a given scale R, which on average contains a mass:

$$\langle M \rangle \propto \rho_b R^3 \ . \tag{2.46}$$

As δ is a Gaussian variable, then all its linear combinations (like M) will also have a Gaussian distribution. The mass variance is thus:

$$\sigma^2(R) = \frac{\langle (M - \langle M \rangle)^2 \rangle}{\langle M \rangle^2} \ . \tag{2.47}$$

Expanding the density in Fourier series as in equation (2.40) the variance becomes [24, 41]:

$$\begin{aligned}
\sigma^2(R) &= \frac{1}{V^2} \left\langle \int \int \sum_{\mathbf{k}} \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{k}'} \delta(\mathbf{k}') e^{i\mathbf{k}'\cdot\mathbf{r}'} d\mathbf{r} d\mathbf{r}' \right\rangle = \\
&= \frac{1}{V^2} \left\langle \sum_{\mathbf{k},\mathbf{k}'} \delta(\mathbf{k}) \delta^*(\mathbf{k}') \int e^{i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} \int e^{-i\mathbf{k}'\cdot\mathbf{r}'} d\mathbf{r}' \right\rangle = \\
&= \frac{1}{V^2} \left\langle \sum_{\mathbf{k},\mathbf{k}'} \delta(\mathbf{k}) \delta^*(\mathbf{k}') e^{[i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}_c]} I(\mathbf{r}-\mathbf{r}_c) I'(\mathbf{r}'-\mathbf{r}_c) \right\rangle \ ,
\end{aligned} \tag{2.48}$$

where $\mathbf{r}_c$ is the center of the volume V, and $I = \int \exp[i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_c)] d(\mathbf{r}-\mathbf{r}_c)$. As the first exponent is a delta function, the whole sum is:

$$\sigma^2(R) = \sum_{\mathbf{k}} \langle |\delta(\mathbf{k})|^2 \rangle \left[\frac{1}{V} \int e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_c)} d(\mathbf{r}-\mathbf{r}_c) \right]^2 = \sum_{\mathbf{k}} \langle |\delta(\mathbf{k})|^2 \rangle I^2 =$$

$$= \frac{1}{V}\sum_{\mathbf{k}} \delta^2(\mathbf{k}) W^2(k;R) \ . \tag{2.49}$$

W is a Fourier transform of a filter ('window') function used for smoothing the density field in real space. Clearly, the particular choice of filter will determine the proportionality constant in eq. (2.46). The continuum case for $\sigma(R)$ is straightforward:

$$\sigma^2(R) = \frac{1}{2\pi^2}\int_0^\infty k^2 P(k) W^2(k;R) dk \ . \tag{2.50}$$

Comparing it to equation (2.43), we see that $\sigma^2(R) < \sigma^2$. While for the variance all modes are summed, in the mass variance case modes with wavelenghts much smaller than R will not be important as the window function will average them out. If the power spectrum falls off with decreasing k, the modes on scales much larger than R will also contribute very little.

2.4 Extensions of Linear Theory

In the linear regime each Fourier mode $\delta(\mathbf{k}, t)$ evolves independently preserving statistical properties of the modes, like the power spectrum. As inhomogeneities grow further, non-linear features – most importantly mode to mode coupling – develop [54], and there is no full analytical theory describing this regime.

In the following the most often used non-linear approximations will be discussed, which involve extrapolations of the linear properties of the density field (well) into the non-linear regime. While these models do not attempt to correctly account for the dynamical evolution of clustering, but rather are based on phenomenological approaches, often a simple intuition, they still provide interesting insights into the properties of the non-linear mass distribution. A detailed review of non-linear approximation methods for the gravitational collapse is given by Sahni and Coles [55].

2.4.1 Zel'dovich Approximation

The Zel'dovich approximation [56] considers perturbations to fluid element's (particle) trajectories, rather than perturbations at fixed point in space. Therefore, this represents going from the Eulerian treatment, discussed in §2.1 and §2.2, to a Lagrangian formalism. There, one considers the change in the final (Eulerian) coordinate $\mathbf{r}$ from its initial (Lagrangian) position $\mathbf{q}$:

$$\mathbf{r} = \mathbf{q} + \mathbf{\Psi}(\mathbf{q}, t) \ . \tag{2.51}$$

The displacement term can be perturbatively expanded as

$$\mathbf{\Psi} = \mathbf{\Psi}^{(0)} + \mathbf{\Psi}^{(1)} + ... \tag{2.52}$$

The Zel'dovich approximation consists of truncating the above expansion at linear order

$$\mathbf{\Psi} \approx \mathbf{\Psi}^{(0)} = D_+(t)\mathbf{s}(\mathbf{q}) = D_+(t)\nabla\Phi(\mathbf{q}) \ . \tag{2.53}$$

Thus, the motion of each fluid element is determined by the linear growth rate, and the spatial perturbations $\mathbf{s}(\mathbf{q})$, which can be expressed as the gradient of the gravitational potential. A simple substitution: $\tau = D_+(t)$, $\mathbf{v} = -\mathbf{s}(\mathbf{q})$, leads to $\mathbf{r}(\mathbf{q},\tau) = \mathbf{q} + \tau\mathbf{v}$ – a transformation from Lagrangian to Eulerian coordinates for the case when no forces are present. We see that Zel'dovich approximation does not account for the dynamics of gravitational collapse, but rather treats motion in a ballistic way. Also, it assumes a velocity field of potential type $\mathbf{v} = -\nabla\Phi(\mathbf{q})$, which is not very restrictive as (possible) initial rotations in the fluid would be damped by the expansion of the universe.

As long as the trajectories of different fluid elements do not cross, the relation

$$\mathbf{r} = \mathbf{q} + D_+(t)\nabla\Phi(\mathbf{q}) \tag{2.54}$$

provides a unique mapping between Eulerian and Lagrangian space. Applying mass conservation, the relation among densities in two systems has to be:

$$\rho(\mathbf{q},t) = \frac{\rho_0}{|\mathcal{D}_{ij}|} \ , \tag{2.55}$$

where $\mathcal{D}_{ij}$ is the tensor of deformation:

$$\begin{aligned} \mathcal{D}_{ij} &= \frac{\partial r_i}{\partial q_j} = \delta_{ij} + D_+ \frac{\partial v_i}{\partial q_j} = \\ &= \left\| \begin{array}{ccc} 1 - D_+\lambda_1(\mathbf{q}) & 0 & 0 \\ 0 & 1 - D_+\lambda_2(\mathbf{q}) & 0 \\ 0 & 0 & 1 - D_+\lambda_3(\mathbf{q}) \end{array} \right\| \ , \end{aligned} \tag{2.56}$$

whose eigenvalues are λ_i, and in the general case, each can be positive or negative. The probability of a coincidence $\lambda_1 = \lambda_2 = \lambda_3$, or even only two eigenvalues being exactly the same is effectively zero.

For a Gaussian random field in a $\Omega = \Omega_m = 1$ universe, the set of eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$ will have the probability distribution ([57], english translation in [58]):

$$P(\lambda_1,\lambda_2,\lambda_3) = \frac{5^3 \cdot 27}{8\pi\sigma_{in}^6\sqrt{5}} \exp\left(\frac{-3I_1^2 + 15I_2/2}{\sigma_{in}^2}\right)(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3) \ , \tag{2.57}$$

where $I_1 = \lambda_1 + \lambda_2 + \lambda_3$ and $I_2 = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3$ are the invariants[3] of the deformation tensor $\mathcal{D}_{ij}$, and σ_{in} is the variance of the initial density field. The probability of only one eigenvalue being positive is 42%, and this leads to a one dimensional collapse, into two dimensional *pancakes*. The same is the probability of having only one negative eigenvalue, resulting in a collapse along two axis into one dimensional *filaments*, typically forming at intersections of pancakes. In 8% of cases all λ_i will be positive, and collapse will occure in all three directions resulting in clumps (strictly speaking a point) at filament crossings. Note that in all these cases, the initial collapse is one dimensional forming disk structures [59]; at later stages however, all three eigenvalues are important (Eq. (2.55)). Finally, the remaining 8% of cases, when all eigenvalues are negative, corresponds to formation of expanding voids.

From equations (2.55) and (2.57) one can obtain the probability distribution function (PDF) for the density contrast δ (see [60]):

$$P(\delta, z) = \frac{9 \cdot 5^{3/2}}{4\pi(1+\delta)^3 N_s \sigma^4} \times \tag{2.58}$$
$$\times \int_{3(1+\delta)^{-1/3}}^{\infty} e^{-\frac{(s-3)^2}{2\sigma^2}} (1 + e^{-\frac{6s}{\sigma^2}})(e^{-\frac{\beta_1^2}{2\sigma^2}} + e^{-\frac{\beta_2^2}{2\sigma^2}} - e^{-\frac{\beta_3^2}{2\sigma^2}}) ds \ ,$$

where N_s is the mean number of streams in the flow ($N_s = 1$ means there are no shell crossings), and σ is the variance of the density field at redshift z according to linear theory, $\sigma(z) = D_+(z)\sigma_{in}$, and

$$\beta_n(s) = s\sqrt{5}\left(\frac{1}{2} + \cos\left[\frac{2}{3}(n-1)\pi + \frac{1}{3}\arccos\left(\frac{54}{(1+\delta)s^3} - 1\right)\right]\right) \ . \tag{2.59}$$

At very early times, when the variance is small, $\sigma \ll 1$, the density distribution remains Gaussian as expected from linear theory:

$$P(\rho) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\rho - \bar{\rho})^2}{2\sigma^2}\right] \ . \tag{2.60}$$

As evolution proceeds, the density contrast in some regions grows to very high values, while on the other side, has to be bounded by $\rho \geq 0 \ \Rightarrow \ \delta \geq -1$ requirement. Thus, the density PDF will at later times become non-Gaussian (for a detailed discussion of density PDFs, see [61]). Figure 2.3 shows $P(\delta)$ at the current epoch, $z = 0$, and it is interesting to note that the Zel'dovich approximation successfully describes the statistics of the density field even much beyond the validity of linear theory, $\delta \ll 1$.

It has to be emphasised that the picture of structure evolution leading to pancakes, filaments, clumps, and voids, is in a remarkable agreement with simulations (and after all, with the real universe). When perturbations are small, both Zel'dovich approximation and standard Eulerian linear theory (§2.2) will agree

[3]Third invariant being $I_3 = \lambda_1\lambda_2\lambda_3$.

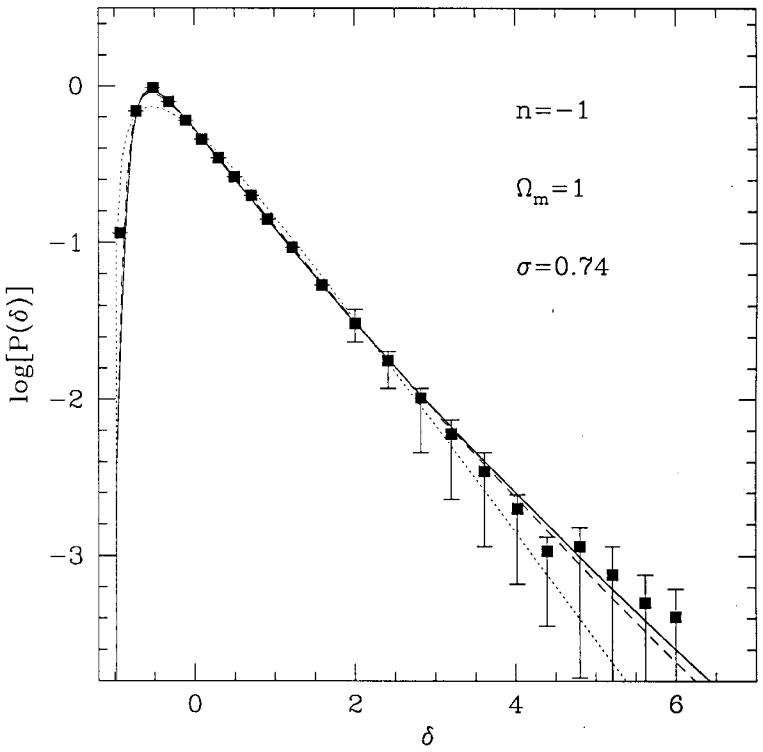

Figure 2.3: Probability distribution function of density contrasts in an $\Omega = \Omega_m = 1$ universe, measured in spheres of R=10 Mpc/h (corresponding to density RMS σ=0.74). The red dotted line shows the Zel'dovich approximation (analogous to starting an N-body simulation at $z_i = 0$). Blue and black lines are for $z_i = 5$, and $z_i \to \infty$. The data points are from an N-body simulation. [Figure from Valageas (2002) [5].]

and provide virtually exact results. However, extrapolation of the linearized Euler equations into the non-linear regime will lead to meaningless results, for example regions with negative densities, while the Zel'dovich approximation still maintains finite errors even when extrapolated to $\rho \to \infty$.

These regions of infinite densities (caustics) form when trajectories of (collisionless, zero temperature) particless cross; in Zel'dovich approximation, it will happen instantaneously, followed by subsequent decrease of density as particle ballistically drift away from multi-strem regions. Fourier space analysis [62] shows that caustics form not because of extreme growth of wave amplitudes, but because of coordination of phases. As in reality, gravity will keep particles close to a pancake region, it is only the formation of caustics when the Zel'dovich approximation becomes meaningless. Many extensions of Zel'dovich theory have been made, in an attempt to extend the analytical approach further (e.g. truncated Zel'dovich approximation, adhesion model), but they will not be reviewed here (see [55]).

Finally, due to its lagrangian nature, as well as for its straightforward implementation, the Zel'dovich approximation became a common approach for

setting up initial conditions in cosmological N-body simulations.

2.4.2 The Spherical Collapse Model

We have seen that perturbations in Lagrangian formalism can be extrapolated much further than $\delta\rho/\rho_b \ll 1$ with results being at least qualitatively correct. However, the Zel'dovich approximation (§2.4.1) cannot produce virialized strucures – which are most interesting, as they are observationally more accessible than pancakes, filaments or voids. Another simple and very popular model of gravitational instability in the nonlinear regime is the spherical collapse model. There, one considers an isolated and homogeneous spherical density perturbation embeded in a homogeneous expanding background. The effects of neighbouring density perturbations and tidal forces on the evolution of the considered perturbation are ignored.

These simplifying assumptions make analytical considerations possible, and despite its simplicity, the model has been widely used not only to explain properties of a single virialized object, but also for statistical properties of an ensemble of gravitationally collapsed structures. As it was pointed out that the probability of uniform collapse in all three directions is effectively zero (§2.4.1), the success of spherical collapse might seem strange. Arguably, the explanation is that caustics (on a certain mass scale) form before the first objects (of similar masses); thus the Zel'dovich approximation is an excellent ansatz for initial collapse of matter up to the first crossing of particle trajectories, at which point other contraction eigenvalues become as important. An alternative way to think about this problem is to note that within a filament, many virialized objects will form.

According to Birkhoff's theorem (the relativistic analog to Gauss's law), an isolated spherical perturbation evolves as a homogeneous FLRW universe of its own [63, 64], whose line element is given by eq. (1.1), and whose evolution is determined by Friedmann equations (1.10) & (1.11). If the total energy density of the fluctuation (Ω') is greater than that of the background universe (Ω), it is an overdensity, which in this idealized scenario is the progenitor of a galaxy, group, or cluster. On the other hand, if $\Omega' < \Omega$, it is an underdensity which corresponds to a (idealized, spherical) void. Clearly, the evolution of the spherical density perturbation will crucially depend upon the value of Ω'.

Let us consider a spherical overdensity of radius R with an initial overdensity δ_i and mass

$$M = \frac{4\pi}{3} R^3 \rho_b (1 + \delta_i) \ , \tag{2.61}$$

Note that, since other perturbations do not interfere with this one, the above mass is conserved, but the overdensity will of course change. Furthermore, the total energy will be conserved as well:

$$E = \frac{1}{2}\dot{R}^2 - \frac{GM}{R} = \text{ const.} \tag{2.62}$$

Assuming that initially the difference between the spherical perturbation and the background universe is infinitely small, the expansion rates of the perturbation and the universe will be the same: $H_{\rm pert} = H_{\rm univ} = H_i$. Furthermore, we can consider perturbation with no peculiar velocity; the radius of any shell R_i then evolves as $\dot{R}_i = H_i R_i$.

The kinetic and potential energies per unit mass are therefore

$$K_i = \frac{1}{2} H_i^2 R_i^2$$

$$U_i = -\frac{GM}{R_i} = -K_i \Omega_i (1 + \delta_i) \ , \tag{2.63}$$

where the density parameter Ω_i is the ratio of the mean mass density to the density in an $\Omega_m = \Omega = 1$ universe with the same Hubble constant at t_i:

$$\Omega_i = \frac{8\pi G}{3} \rho_b H_i \ . \tag{2.64}$$

As a result we get the total energy:

$$E = K_i + U_i = K_i \Omega_i \left[\Omega_i^{-1} - (1 + \delta_i) \right] \ . \tag{2.65}$$

Assuming different shells do not cross, the requirement for the collapse is $E < 0$. In a universe consisting only of pressureless matter, the Hubble parameter is:

$$\begin{aligned} H(z) &= H_0 \left[\Omega_m (1+z)^3 + (1 - \Omega_m)(1+z)^2 \right]^{1/2} = \\ &= H_0 (1+z)(1 + \Omega_m z)^{1/2} \ , \end{aligned} \tag{2.66}$$

thus the density evolves as $\Omega_i \equiv \Omega_m(z) = \Omega_m(1+z)/(1+\Omega_m z)$. With that, the condition for collapse becomes:

$$\delta_i(z) > \frac{1 - \Omega_m}{\Omega_m (1+z)} \ . \tag{2.67}$$

In an open universe $\Omega_m < 1$, the density perturbation must exceed some value $\delta_i > \Omega_i^{-1} - 1$ in order to prevent eternal expansion; in the case of a closed or flat universe, even infinitesimal departure from the mean density will result in a collapsed structure at some redshift.

The time evolution of a spherical shell is identical to that of a spatially open or closed universe:

$$\begin{aligned} R &= A(1 - \cos\theta) \\ t &= B(\theta - \sin\theta) \end{aligned} \tag{2.68}$$

when $E < 0$, and

$$R = A(\cosh\theta - 1)$$

$$t = B(\sinh\theta - \theta) \tag{2.69}$$

for the $E > 0$ case. Combining the eqations for R and t we get $A^3 = GMB^2$.

As a result, an overdense spherical shell will expand, but with a slower rate than the background universe, increasing its density contrast. At a turn-around time t_{ta}, it will reach its biggest size R_{ta}, and starts collapsing. At the moment of turn-around, the shell is instantaneously at rest, and from conservation of energy we have:

$$\begin{aligned} E = K_i\Omega_i\left[\Omega_i^{-1} - (1+\delta_i)\right] &= -\frac{GM}{R_{ta}} = \\ &= \frac{R_i}{R_{ta}} K_i\Omega_i(1+\delta_i)\ , \end{aligned} \tag{2.70}$$

wherefrom

$$R_{ta} = R_i \frac{1+\delta_i}{\delta_i - (\Omega_i^{-1} - 1)}\ . \tag{2.71}$$

Clearly, $\dot{R} = 0$ occurs when $\theta = \pi$ (equations 2.68), and one can use that to find A and B: $A = R_{ta}/2, B = t_{ta}/\pi$. In terms of initial perturbation quantities these coefficients are

$$\begin{aligned} A &= \left(\frac{R_i}{2}\right)\frac{1+\delta_i}{\delta_i - (\Omega_i^{-1} - 1)} \\ B &= \frac{1+\delta_i}{2H\Omega_i^{1/2}[\delta_i - (\Omega_i^{-1} - 1)]^{3/2}}\ . \end{aligned} \tag{2.72}$$

For a flat universe, useful approximations are $A \approx R_i/(2\delta_i)$ and $B \approx 3t_i\delta_i^{-3/2}/4$.

At early times, when $\theta \ll 1$, equations 2.68 can be solved by expanding in θ:

$$\frac{t}{t_{ta}} = \frac{1}{\pi}\left(\frac{\theta^3}{6} - \frac{\theta^5}{120} + \cdots\right)\ . \tag{2.73}$$

From here θ is

$$\theta^2 = \left(\frac{6\pi t}{t_{ta}}\right)^{2/3}\left[1 + \frac{1}{30}\left(\frac{6\pi t}{t_{ta}}\right)^{2/3} - \cdots\right] \tag{2.74}$$

resulting in R:

$$\begin{aligned} \frac{R}{R_{ta}} &= \frac{\theta^2}{4} - \frac{\theta^4}{48} + \cdots = \\ &= \frac{1}{4}\left(\frac{6\pi t}{t_{ta}}\right)^{2/3}\left[1 - \frac{1}{20}\left(\frac{6\pi t}{t_{ta}}\right)^{2/3} + \cdots\right]\ . \end{aligned} \tag{2.75}$$

Keeping only the dominant terms we find that the overdensity in the linear, $\theta \ll 1$, regime evolves as:

$$\delta_L = \frac{3}{20}\left(\frac{6\pi t}{t_{ta}}\right)^{2/3}\ . \tag{2.76}$$

Table 2.1: Linear theory and exact overdensities δ at various stages in the spherical collapse model.

θ	δ_L	δ
$\theta \to 0$	$\delta_L \propto \theta^2$	$\delta \propto \theta^2$
$\frac{\pi}{2}$	0.341	0.466
$\frac{2\pi}{3}$	0.568	1.01
π	1.063	4.6
2π	1.686	∞

Coming back to the full evolution: the density in a shell is

$$\rho(t) = \frac{3M}{4\pi A^3(1-\cos\theta)^3} , \tag{2.77}$$

while the background in the case of a flat, matter dominated universe evolves as:

$$\bar{\rho}(t) = \frac{1}{6\pi G t^2} = \frac{1}{6\pi G B^2(\theta - \sin\theta)^2} . \tag{2.78}$$

Combining the above equations we get the change of overdensities

$$\delta(\theta) = \frac{9}{2}\frac{(\theta-\sin\theta)^2}{(1-\cos\theta)^3} - 1 , \tag{2.79}$$

and similarly for underdense regions:

$$\delta(\theta) = \frac{9}{2}\frac{(\theta-\sinh\theta)^2}{(\cosh\theta-1)^3} - 1 . \tag{2.80}$$

Again the linear limit for small θ is:

$$\lim_{\theta\to 0}\delta(\theta) \approx \frac{3\theta^2}{20} \approx \frac{3}{20}\left(\frac{6\pi t}{t_{ta}}\right)^{2/3} . \tag{2.81}$$

With these formulae, we can find overdensities at some interesting times, and Table 2.4.2 shows a comparison of exact and extrapolated linear overdensities. At the turnaround, $\delta(\theta=\pi) \approx 4.6$, and is thus clearly in the nonlinear regime. The extrapolated linear density contrasts would be $\delta_L(\pi) \approx 1.063$. At $\theta = 2\pi$, all the mass will collapse to a point and $\delta(2\pi) \to \infty$. The collapse treshold in linear theory is $\delta_L(2\pi) \equiv \delta_c \approx 1.686$. This value actually changes only little if we include dark energy; while the calculations for that case have to be done numerically, an excellent fit is given by [65]:

$$\delta_c = \frac{3(12\pi)^{2/3}}{20}(1 + \alpha\log\Omega_{DE}) , \tag{2.82}$$

where $\alpha = 0.353w^4 + 1.044w^3 1.128w^2 + 0.555w + 0.131$.

In reality, $\delta_{coll} \to \infty$ will never be achieved since the exact spherical collapse

where particles do not have any velocity dispersion, is quite a crude approximation. As the overdensity is contracting, dynamical relaxation and shocks will force the system into virial equilibrium at a finite density. Therefore, the maximum density an object will reach can be estimated using the virial theorem and the fact that at $R = R_{ta}$ all the energy in the system is potential:

$$U(R = R_{vir}) = 2E = 2U(R = R_{ta}), \tag{2.83}$$

since $U = -GM/R$ we get $R_{vir} = R_{ta}/2$ and $\rho_{vir} = 8\rho_{ta}$ [66]. The mean density of an object at turnaround is $\rho_{ta} = \delta_{ta}+1 \approx 5.6\rho_b(t_{ta})$. Since $\rho_b(t) = (6\pi G t^2)^{-1}$ and with $t_{vir} \simeq t_{coll} \simeq 2t_{ta}$ we finally get $\rho_{vir} \simeq 8 \times 5.6 \times 4\rho_b(t_{vir})$ or since $\bar{\rho}_{vir} = (1+z)^3\rho_0$

$$\Delta_{vir} \equiv \frac{\rho_{vir}}{\rho_b} \approx 180(1+z_{vir})^3 \ , \tag{2.84}$$

where z_{vir} is the virialization redshift. Equation (2.84) permits us to relate the virialised density of a collapsed object to the epoch of its formation: $z_{vir} \simeq 0.18(\rho/\rho_0)^{1/3} - 1$.

While δ_c carries only a small dependence on cosmology, that is not the case with Δ_{vir}. For the family of flat cosmologies ($\Omega_m + \Omega_\Lambda = 1$), which is of most interest, Δ_{vir} can be approximated with 1% accuracy [67] to be:

$$\Delta_{vir} \approx \frac{18\pi^2 + 82x(z) - 39x^2(z)}{\Omega_m(z)} \ , \tag{2.85}$$

where $x(z) = \Omega_m(z) - 1$. For example, for some fiducial value of $\Omega_m = 0.3$, $\Delta_{vir} \approx 337$. Similarly, for the case of more general dark energy models, in the range of $-1 \leq w \leq -0.3$, the fit is [65]:

$$\Delta_{vir} \approx 18\pi^2 \left[1 + \Theta^b(z)\right] \ , \tag{2.86}$$

where $\Theta(z) = 1/\Omega_m(z) - 1$.

Similarity Solution

An interesting extension of the spherical collapse model is the case where the energy of a perturbation is a function of mass

$$E = E_0 \left(\frac{M}{M_0}\right)^{\frac{2}{3}-\epsilon} < 0 \ . \tag{2.87}$$

Turnaround radius and time are [68]:

$$R_{ta} = -\frac{GM}{E} = \frac{GM_0}{-E_0}\left(\frac{M}{M_0}\right)^{\frac{1}{3}+\epsilon}$$

$$t_{ta} = \frac{\pi}{2}\sqrt{\frac{R_{ta}^3}{2GM}} = \pi GM\left(\frac{-E_0}{2}\right)^{-3/2}\left(\frac{M}{M_0}\right)^{\frac{3\epsilon}{2}} \ . \tag{2.88}$$

The density profile of a virialized object is

$$\rho(r) \propto \frac{M}{r^3} \propto r^{-9\epsilon/(1+3\epsilon)} \ . \tag{2.89}$$

An interesting case would be $\epsilon = 1$, resulting in an energy of the shell equal to the one due to a point mass, $E \propto M^{-1/3}$, and the density profile $\rho \propto r^{-9/4}$. Another is $\epsilon = 2/3$, resulting in an isothermal sphere, with energy being independant of M, and density $\rho \propto R^{-2}$.

3 Mass Function

As discussed in the previous chapters, a broad suite of astrophysical and cosmological observations provides compelling evidence for the existence of dark matter. Although its ultimate nature is unknown, the large-scale dynamics of dark matter is essentially that of a self-gravitating collisionless fluid. In an expanding universe, gravitational instability leads to the formation and growth of structure in the dark matter distribution. The existence of localized, highly overdense dark matter clumps, or halos, is a key prediction of cosmological nonlinear gravitational collapse. At low redshifts, the mass function is an important probe of cosmological parameters, orthogonal to the geometrical probes (see §1.3.2). At higher redshifts, the halo mass function is important in probing quasar abundance and formation sites [69], as well as the reionization history of the Universe [70].

Many recently suggested reionization scenarios are based on the assumption that the mass function is given reliably by modified Press-Schechter type fits (Press & Schechter 1974 [71], hereafter PS; Bond et al. 1991 [72]). However, the theoretical basis of this approach is at best heuristic and careful numerical studies are required in order to obtain accurate results. Two examples serve to illustrate this statement. Reed et al. in 2003 [3] reported a discrepancy with the Sheth-Tormen fit (Sheth & Tormen 1999 [37], hereafter ST) of ~50% at a redshift of $z = 15$ (the different fitting formulae and their origin will be explained in §3.1). In addition, Heitmann et al. [73] show that the Press-Schechter form can be severely incorrect at high redshifts: at $z \geq 10$, the predicted mass function sinks below the numerical results by an order of magnitude at the upper end of the relevant mass scale. Consequently, incorrect, or at best imprecise, predictions for the reionization history can result from the failure of fitting formulae.

The halo formation is a complicated nonlinear gravitational process, and the current theoretical understanding of the mass, spatial distribution, and inner profiles of halos remains at a relatively crude level. Numerical simulations are therefore crucial as drivers of theoretical progress, having been instrumental in obtaining important results such as the Navarro-Frenk-White (NFW) profile [74] for dark matter halos and an (approximate) universal form for the mass function (Jenkins et al. 2001 [75], hereafter Jenkins). In order to better understand the evolution of the mass function at high redshifts, a number of numerical studies have been carried out. High–redshift simulations, however, suffer from their own

set of systematic issues, and simulation results can be at considerable variance with each other, differing on occasion by as much as an order of magnitude!

Motivated by all of these reasons, we carry out a numerical investigation of the evolution of the mass function with the aim of attaining good control over both statistical and, more importantly, possible systematic errors in N-body simulations. We first pay attention to simulation criteria for obtaining accurate mass functions with the aim of reducing systematic effects. Our two most significant points are that simulations must be started early enough to obtain accurate results and that the box sizes must be large enough to suppress finite-volume artifacts. As in most recent work following that of Jenkins, we define halo masses using a friends-of-friends (FOF) halo finder with linking length $b = 0.2$. This choice introduces systematic issues of its own (e.g., connection to spherical overdensity mass as a function of redshift), which we touch on as relevant below. While not quantitatively significant in the context of this chapter, the choice of halo definition will be discussed in details in Chapter 4 (see also [4]).

The more detailed analysis presented here enable studying the mass function at statistical and systematic accuracies reaching a few percent over most of our redshift range, a substantial improvement over previous works. At this level we find discrepancies with the 'universal' fit of Jenkins at low redshifts ($z < 5$), but it must be kept in mind that the universality of the original fit was only meant to be at the $\pm 20\%$ level. Moreover, in a recent work Reed and collaborators [4] have reported the violation of universality at high redshifts (up to $z = 30$). To fit the mass function they have incorporated an additional free parameter, the effective spectral index $n_{\rm eff}$, with the aim of understanding and taking into account the extra redshift dependence missing from conventional mass–function–fitting formulae. Our simulation results are consistent with the trends found by Reed et al. [4] at low redshifts ($z \leq 5$), but at higher redshifts we do not observe a statistically significant violation of the universal form of the mass function.

On the other hand, results from some previous simulations have reported good agreement with the Press-Schechter mass function at high redshifts. Since the Press-Schechter fit has been found significantly discrepant with low–redshift results ($z < 5$), this would imply a strong disagreement with extending the well-validated low–redshift notion of (approximate) mass function universality to high z. Our conclusion is that the simulations on which these findings were based violated one or more of the criteria to be discussed below.

As simulations are perforce restricted to finite volumes, the obtained mass function clearly cannot represent that of an infinite box. Not only is sampling a key issue, but also the fact that simulations with periodic boundary conditions have no fluctuations on scales larger than the box size. To minimize and test for these effects we were conservative in our choices of box size and the mass range probed in each individual box. We also used nested-volume simulations

to directly test for finite-volume effects. Because we used multiple boxes and averaged mass function results over the box ensemble, extended Press-Schechter theory can be used to correct for residual finite volume–effects [36, 76]. Details on this are given in §3.4.3.

The chapter is organized as follows. In §3.1 we give a brief overview of the mass function and popular fitting formulae, discussing as well previous numerical work on the halo mass function at high redshifts. In §3.2 we give a short description of the N-body code MC^2 (**M**esh-based **C**osmology **C**ode) and a summary of the performed simulations. In §3.3 we derive and discuss some simple criteria for the starting redshift and consider systematic errors related to the numerical evolution such as mass and force resolution and time stepping. These considerations in turn specify the input parameters for the simulations in order to span the desired mass and redshift range for our investigation. In §3.4 we present results for the mass function at different redshifts as well as the halo growth function, which describes the number of halos of a given mass as a function of redshift. Here we also discuss the importance of post-processing corrections such as FOF particle sampling compensation and finite-volume effects. We discuss the results and implications in §3.5.

3.1 Definitions and Previous Work

The mass function describes the number density of halos of a given mass. In order to determine the mass function in simulations one has to first *identify* the halos and then *define* their mass. No precise theoretical basis exists for these operations. Nevertheless, depending on the situation at hand, the observational and numerical communities have adopted a few 'standard' ways of defining halos and their associated masses. For a recent review of these issues with regard to observations, see, e.g., Ref. [77], but for a more theoretically oriented review, see, Ref. [78], and Chapter 4.

3.1.1 Halo Mass

The detailed analysis of mass definition choice, as well as cosmologically interesting consequences are discussed in Chapter 4; here we will provide just a brief overview. There are basically two ways to find halos in a simulation. One, the overdensity method, is based on identifying overdense regions above a certain threshold. The threshold can be set with respect to the critical density $\rho_{\rm c} = 3H^2/8\pi G$ (or the background density $\rho_{\rm b} = \Omega_m \rho_{\rm c}$). The mass M_Δ of a halo identified this way is defined as the mass enclosed in a sphere of radius r_Δ whose mean density is $\Delta\rho_{\rm c}$. Common values for Δ range from 100 to 500 (or even higher). For clusters of galaxies observations [77], higher values for Δ are easier to work with. Properties of clusters are easier to observe in higher density regions and these regions are more relaxed than the outer parts which

are subject to the effects of inflow and incomplete mixing. The disadvantage of defining a halo in this manner is that sphericity of halos is implied, an assumption which may be easily violated, e.g., in the case of halos that formed in a recent merger event or halos at high redshifts. At higher redshifts, the nonlinear mass scale M_* decreases rapidly, and the ratio of the considered halo mass $M_{\rm halo}$ to M_* can become large. This translates into producing large-scale structures roughly analogous to supercluster structures today. While these structures are gravitationally bound, they are often not virialized, nor spherical. Even the much smaller structures (which are considered here) are not virialized at high redshifts, and therefore, assumptions about sphericity are most likely violated. Hence the spherical overdensity method does not suggest itself as an obvious way to identify halos at high redshift.

The other method, the FOF algorithm, is based on finding neighbors of particles and neighbors of neighbors as defined by a given separation distance (see, e.g., [79, 80]). The FOF algorithm leads to halos with arbitrary shapes since no prior symmetry assumptions have been made. The halo mass is defined simply as the sum of particles which are members of the halo. While this definition is easy to apply to simulations, the connection to observations is difficult to establish directly. (Chapter 4 analyses this, also see [78]).

It is important to keep in mind that the definition of a halo is essentially the adoption of some sort of convention for the halo boundary. In reality, a sharp distinction between the particles in a halo and particles in the simulation 'field' does not exist. Jenkins showed that the choice of a FOF finder with a linking length $b = 0.2$ to define halo masses provides the best fit for a universal form of the mass function. This choice has since been adopted by many numerical practitioners as a standard convention. A useful discussion of the various halo definitions can be found in ref. [81].

In the following we use the FOF algorithm with a fixed linking length of $b = 0.2$ to identify halos and determine their masses. Also, we correct masses to account for the finite number of particles; as recently pointed out by Warren et al. ([8], hereafter Warren) FOF masses suffer from a systematic problem when halos are sampled by relatively small numbers of particles. Although halos can be robustly identified with as few as 20 particles, if a given halo has too few particles, its FOF mass turns out to be systematically too high. Exact compensation for this effect is described in §3.4.2.

3.1.2 Defining the Mass Function

The exact definition of the mass function, e.g., integrated versus differential form or count versus number density, varies widely in the literature. To characterize different fits, Jenkins introduced the scaled differential mass function $f(\sigma, z)$ as

a fraction of the total mass per $\ln\sigma^{-1}$ that belongs to halos:

$$f(\sigma,z) \equiv \frac{d\rho/\rho_b}{d\ln\sigma^{-1}} = \frac{M}{\rho_{\rm b}(z)}\frac{dn(M,z)}{d\ln[\sigma^{-1}(M,z)]}. \tag{3.1}$$

Here $n(M,z)$ is the number density of halos with mass M, $\rho_{\rm b}(z)$ is the background density at redshift z, and $\sigma(M,z)$ is the variance of the linear density field. As pointed out by Jenkins, this definition of the mass function has the advantage that to a good accuracy it does not explicitly depend on redshift, power spectrum, or cosmology; all of these are encapsulated in $\sigma(M,z)$. For the most part, we will display the mass function

$$F(M,z) \equiv \frac{dn}{d\log M} \tag{3.2}$$

as a function of $\log M$ itself. [In §3.4 we include results for $f(\sigma,z)$.]

To compute $\sigma(M,z)$, the power spectrum $P(k)$ is smoothed with a spherical top-hat filter function of radius R, which on average encloses a mass M ($R = [3M/4\pi\rho_{\rm b}(z)]^{1/3}$):

$$\sigma^2(M,z) = \frac{d^2(z)}{2\pi^2}\int_0^\infty k^2 P(k) W^2(k,M) dk, \tag{3.3}$$

where $W(k,M)$ is the top-hat filter:

$$W(r) = \begin{cases} \frac{3}{4\pi R^3}, & r < R \\ 0, & r > R \end{cases} \tag{3.4}$$

$$W(k) = \frac{3}{(kR)^3}\left[\sin(kR) - kR\cos(kR)\right]. \tag{3.5}$$

The redshift dependence enters only through the growth factor $d(z)$, normalized so that $d(0) = 1$:

$$\sigma(M,z) = \sigma(M,0)d(z). \tag{3.6}$$

In the approximation of negligible difference in the CDM and baryon peculiar velocities, the growth function in a ΛCDM universe is given by (see §2.2.1)

$$d(a) = \frac{D^+(a)}{D^+(a=1)}, \tag{3.7}$$

where we consider d as a function of the cosmological scale factor $a = 1/(1+z)$, and

$$D^+(a) = \frac{5\Omega_{\rm m}}{2}\frac{H(a)}{H_0}\int_0^a \frac{da'}{[a'H(a')/H_0]^3} \tag{3.8}$$

with $H(a)/H_0 = \left[\Omega_{\rm m}/a^3 + (1-\Omega_{\rm m})\right]^{1/2}$. In particular, for $z \gg 1$, when matter dominates the cosmological constant, $D^+(a) \simeq a$.

Even in linear theory, equation (3.8) is only an approximation because baryons began their gravitational collapse with velocities different from those of

CDM particles. Until recombination at $z \sim 1100$, well into the matter era with non-negligible growth of CDM inhomogeneities, the baryons were held against collapse by the pressure of the CMB photons [82]. While thereafter the relative baryon-CDM velocity decayed as $1/a$, the residual velocity difference was sufficient to affect the growth function $d(z)$ at $z = 50$ by more than 1% and at $z = 10$ by about 0.2% [83, 84].

3.1.3 Fitting Functions

Over the last three decades several different fitting forms for the mass function have been suggested. The mass function is not only a sensitive measure of cosmological parameters by itself but also a key ingredient in analytic and semianalytic modeling of the dark matter distribution, as well as of several aspects of the formation, evolution, and distribution of galaxies. Therefore, if a reliable and accurate fit for the mass function applicable to a wide range of cosmologies and redshifts were to exist, it would be of obvious utility. In this section we briefly review the common fitting functions and compare them at different redshifts.

The first analytic model for the mass function was developed by PS. Their theory accounts for a spherical overdense region in an otherwise smooth background density field, which then evolves as a Friedmann universe with a positive curvature. Initially, the overdensity expands, but at a slower rate than the background universe (thus enhancing the density contrast), until it reaches the 'turnaround' density, after which collapse begins. Although from a purely gravitational standpoint this collapse ends with a singularity, it is assumed that in reality – due to the spherical symmetry not being exact – the overdense region will virialize. For an Einstein-de Sitter universe, the density of such an overdense region at the virialization redshift is $z \approx 180\rho_{\rm c}(z)$. At this point, the density contrast from the linear theory of perturbation growth $[\delta(\vec{x}, z) = d(z)\delta(\vec{x}, 0)]$ would be $\delta_{\rm c}(z) \approx 1.686$ in an Einstein-de Sitter cosmology. For $\Omega_{\rm m} < 1$, the value of the threshold parameter $\delta_{\rm c}$ can vary [85], but the dependence on cosmology has little quantitative significance [75]. Thus, throughout this work we adopt $\delta_{\rm c} = 1.686$.

Following the above reasoning and with the assumption that the initial density perturbations are described by a homogeneous and isotropic Gaussian random field, the PS mass function is specified by

$$f_{\rm PS}(\sigma) = \sqrt{\frac{2}{\pi}} \frac{\delta_{\rm c}}{\sigma} \exp\left(-\frac{\delta_{\rm c}^2}{2\sigma^2}\right). \tag{3.9}$$

The PS approach assumes that all mass is inside halos, as enforced by the constraint

$$\int_{-\infty}^{+\infty} f_{\rm PS}(\sigma)\, d\ln\sigma^{-1} = 1. \tag{3.10}$$

Table 3.1: Mass Function Fits for $f(\sigma)$. Shown are examples of commonly used fitting functions. ST used $a = 0.707$ and $p = 0.3$, while [2] suggest that $a = 0.75$ leads to a better fit. The Warren fit represents by far the largest uniform set of simulations based on multiple boxes with the same cosmology run with the same code. We use it as a reference standard throughout this work. [3] suggest an empirical adjustement of the ST fit, which is slightly modified in [4]. For the latter, $G_1(\sigma)$ and $G_2(\sigma)$ are given by eqs. (3.16) and (3.17), respectively, $c = 1.08$, $ca = 0.764$, and $A = 0.3222$.

Reference	Fitting Function $f(\sigma)$	Fit range
ST (2001)	$0.3222\sqrt{\frac{2a}{\pi}}\frac{\delta_c}{\sigma}\exp\left[-\frac{a\delta_c^2}{2\sigma^2}\right]\left[1+\left(\frac{\sigma^2}{a\delta_c^2}\right)^p\right]$	unspecified
Jenkins	$0.315\exp\left[-\lvert\ln\sigma^{-1}+0.61\rvert^{3.8}\right]$	$z = 0-5$
Reed (2003)	$f_{\rm ST}(\sigma)\,\exp\left\{-0.7/\left[\sigma(\cosh(2\sigma))^5\right]\right\}$	$z = 0-15$
Warren	$0.7234\left(\sigma^{-1.625}+0.2538\right)\exp\left[-\frac{1.1982}{\sigma^2}\right]$	$z = 0$
Reed (2007)	$A\sqrt{\frac{2a}{\pi}}\left[1+\left(\frac{\sigma^2}{a\delta_c^2}\right)^p+0.6G_1(\sigma)+0.4G_2(\sigma)\right]$ $\times\frac{\delta_c}{\sigma}\exp\left[-\frac{ca\delta_c^2}{2\sigma^2}-\frac{0.03}{(n_{\rm eff}+3)^2}\left(\frac{\delta_c}{\sigma}\right)^{0.6}\right]$	$z = 0-30$

While as a first rough approximation the PS mass function agrees with simulations at $z = 0$ reasonably well, it overpredicts the number of low–mass halos and underpredicts the number of massive halos at the current epoch. Furthermore, it is significantly in error at high redshifts (see, e.g., [86, 73], but also §3.4.4).

After PS, several suggestions were made in order to improve the mass function fit. These suggestions were based on more refined dynamical modeling, direct fitting to simulations, or a combination of the two.

Using empirical arguments ST proposed an improved mass function fit of the form:

$$f_{\rm ST}(\sigma) = A\sqrt{\frac{2a}{\pi}}\frac{\delta_{\rm c}}{\sigma}\exp\left(-\frac{a\delta_{\rm c}^2}{2\sigma^2}\right)\left[1+\left(\frac{\sigma^2}{a\delta_{\rm c}^2}\right)^p\right], \tag{3.11}$$

with $A = 0.3222$, $a = 0.707$ and $p = 0.3$. (Sheth & Tormen later suggest $a = 0.75$ as an improved value [2].) Note that for values $A = 0.5$, $a = 1$ and $p = 0$ above equation becomes the PS formula. Sheth et al. [87] rederived this fit theoretically by extending the PS approach to an elliptical collapse model. In this model, the collapse of a region depends not only on its initial overdensity but also on the surrounding shear field. The dependence is chosen such that it recovers the Zel'dovich approximation [56] in the linear regime. A halo is considered virialized when the third axis collapses (see also ref. [88] for an earlier, different approach to the same idea).

Jenkins combined high resolution simulations for four different CDM cosmologies (τCDM, SCDM, ΛCDM, and OCDM) spanning a mass range of over 3 orders of magnitude ($\sim (10^{12}-10^{15})\,h^{-1}M_\odot$), and including several redshifts between $z = 5$ and 0. Independent of the underlying cosmology, the following

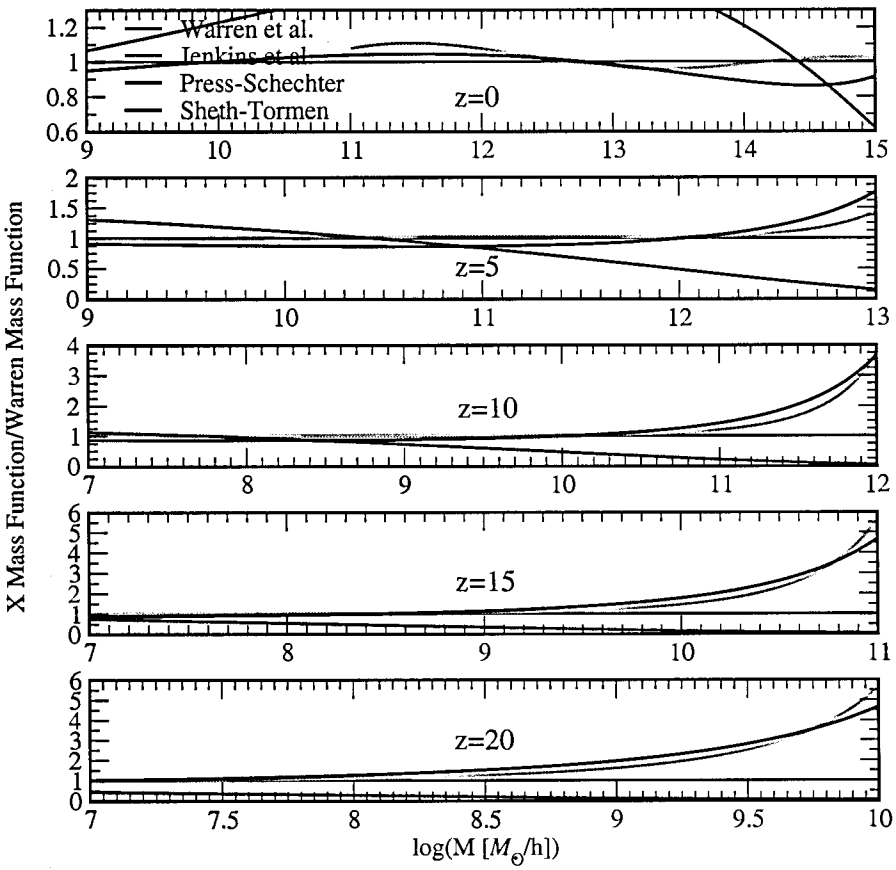

Figure 3.1: Ratio of the Jenkins, PS, and ST mass function fits with respect to the Warren fit for five different redshifts over a range of halo masses. Top to bottom: Redshifts $z = 0$, 5, 10, 15, and 20. Note that the ranges of the axes are different in the different panels. Jenkins fit is not shown below masses of $10^{11}h^{-1}M_\odot$ at $z = 0$, since it is not valid for such low masses at that redshift.

fit provided a good representation of their numerical results (within $\pm 20\%$):

$$f_{\rm Jenkins}(\sigma) = 0.315 \exp\left(-|\ln \sigma^{-1} + 0.61|^{3.8}\right). \tag{3.12}$$

The above formula is very close to the Sheth-Tormen fit, leading to some improvement at the high-mass end. The disadvantage is that it cannot be simply extrapolated beyond the range of the fit, since it was tuned to a specific mass range of their simulations.

By performing 16 nested-volume dark matter simulations, Warren was able to obtain significant halo statistics spanning a mass range of 5 orders of magnitude ($\sim (10^{10}–10^{15})\, h^{-1}M_\odot$). Because this represents by far the largest uniform set of simulations–based on multiple boxes with the same cosmology run with the same code–we use it as a reference standard throughout this work. Using a functional form similar to ST, Warren determined the best mass function fit to

be

$$f_{\rm Warren}(\sigma) = 0.7234\left(\sigma^{-1.625} + 0.2538\right)\exp\left(-\frac{1.1982}{\sigma^2}\right). \tag{3.13}$$

For a quantitative comparison of the different fits at different redshifts, we show the ratio of the PS, Jenkins, and ST fits with respect to the Warren fit in Figure 3.1. We do not show the Jenkins fit below $10^{11}\,h^{-1}M_\odot$ at $z = 0$ since it diverges in this regime. The original ST fit, the Jenkins fit, and the Warren fit all give similar predictions. The discrepancy between PS and the other fits becomes more severe for higher masses at high redshifts. PS dramatically underpredicts halos in the high-mass range at high redshifts (assuming that the other fits lead to reasonable results in this regime). For low-mass halos the disagreement becomes less severe. For $z = 0$ the Warren fit agrees, especially in the low-mass range below $10^{13}\,h^{-1}M_\odot$, to better than 5% with the ST fit. At the high-mass end the difference increases up to 20%. The Jenkins fit leads to similar results over the considered mass range. At higher redshifts and intermediate-mass ranges around $10^9\,h^{-1}M_\odot$, the Warren and ST fit disagree by roughly a factor of 2.

Several other groups have suggested modifications of the ST fit. In §3.4 we compare our results with two of them. In 2003 Reed et al. [3] suggested an empirical adjustment to the ST fit by multiplying it with an exponential function, leading to

$$f_{\rm Reed03}(\sigma) = f_{\rm ST}(\sigma)\,\exp\left\{-0.7/\left[\sigma(\cosh(2\sigma))^5\right]\right\}, \tag{3.14}$$

valid over the range $-1.7 \leq \ln\sigma^{-1} \leq 0.9$. This adjustment leads to a suppression of the ST fit at large σ^{-1}. In 2007 the adjustment to the ST fit is slightly modified again [4], leading to the following new fit:

$$\begin{aligned} f_{\rm Reed07}(\sigma) &= A\sqrt{\frac{2a}{\pi}}\left[1 + \left(\frac{\sigma^2}{a\delta_c^2}\right)^p + 0.6G_1 + 0.4G_2\right] \\ &\quad \times\frac{\delta_c}{\sigma}\exp\left[-\frac{ca\delta_c^2}{2\sigma^2} - \frac{0.03}{(n_{\rm eff}+3)^2}\left(\frac{\delta_c}{\sigma}\right)^{0.6}\right], \end{aligned} \tag{3.15}$$

$$G_1 = \exp\left[-\frac{\ln(\sigma^{-1} - 0.4)^2}{2(0.6)^2}\right], \tag{3.16}$$

$$G_2 = \exp\left[-\frac{\ln(\sigma^{-1} - 0.75)^2}{2(0.2)^2}\right], \tag{3.17}$$

with $c = 1.08$, $ca = 0.764$, and $A = 0.3222$. The adjustment has very similar effects to that of 2003 [3], as we show in §3.4. Reed et al. (2007) [4] also note that the (small) suppression of the mass function relative to ST as a function of redshift seen in simulations can be treated by adding an extra parameter, the power spectral slope at the scale of the halo radius, $n_{\rm eff}$ (formally defined by equation (3.42) below). We return to this issue when we discuss our numerical results in §5. Most commonly used fitting functions are summarized in Table 3.1.

Although fitting functions may be a useful way to approximately encapsulate results from simulations, meaningful comparisons to observations require overcoming many hurdles, e.g., an operational understanding of the definition of halo mass, how it relates to various observations, and error control in N-body codes [89, 90, 91]. Here, our focus is first on identifying possible systematic problems in the N-body simulations themselves and how they can be avoided and controlled.

3.1.4 Halo Growth Function

A useful way to study the statistical evolution of halo masses in simulations is to transform the mass function into the halo growth function, $n(M_1, M_2, z) \equiv \int_{M_1}^{M_2} F \, d\log M$ [73], which measures the mass-binned number density of halos as a function of redshift. The halo growth function, plotted versus redshift in Figure 3.2, shows at a glance how many halos in a particular mass bin and box volume are expected to exist at a certain redshift. This helps set the required mass and force resolution in a simulation which aims to capture halos at high redshifts. For a given simulation volume, the halo growth function directly predicts the formation time of the first halos in a given mass range.

In order to derive this quantity approximately, we first convert an accurate mass function fit (we use the Warren fit here) into a function of redshift z. As will be shown in the following sections, mass function fits work reliably enough out to at least $z = 20$, and can therefore be used to estimate the halo growth function. Figure 3.2 shows the evolution of eight different mass bins, covering the mass range investigated here, as a function of redshift z. As expected from the paradigm of hierarchical structure formation in a ΛCDM cosmology, small halos form much earlier than larger ones. An interesting feature in the lower mass bins is that they have a maximum at different redshifts. The number of the smallest halos grows until a redshift of $z = 2$ and then declines when halos start merging and forming much more massive halos. This feature is reflected in a crossing of the mass functions at different redshifts for small halos.

3.1.5 Mass Function at High Redshift: Previous Work

Most of the effort to characterize, fit, and evaluate the mass function from simulations has been focused on or near the current cosmological epoch, $z \sim 0$. This is mainly for two reasons: (1) so far most observational constraints have been derived from low-redshift objects ($z < 1$); (2) the accurate numerical evaluation of the mass function at high redshifts is a nontrivial task.

The increasing reach of telescopes on the ground and in space, such as the upcoming James Webb Space Telescope, allows us to study the Universe at higher and higher redshifts. Recent discoveries include 970 galaxies at redshifts between $z = 1.5$ and $z = 5$ from the VIMOS VLT Deep Survey [92], and the recent observation of a galaxy at $z = 6.5$ [93]. The epoch of reionization (EOR) is

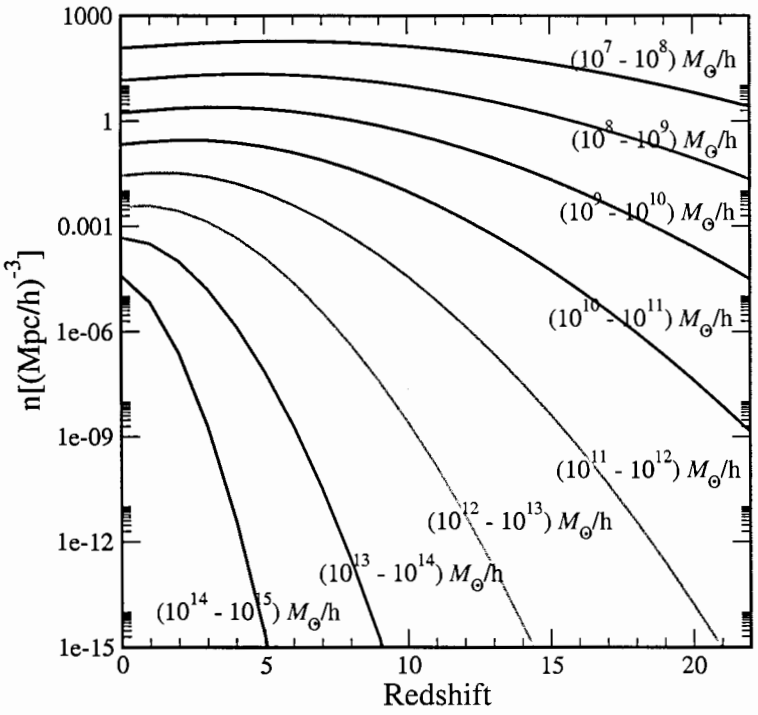

Figure 3.2: Halo growth function based on the Warren mass function fit for different mass bins. The curves for the lower mass bins have a maximum at $z > 0$ which reflects a crossover of the mass functions at different redshifts.

of central importance to the formation of cosmic structure. Although our current observational knowledge of the EOR is rather limited, future 21 cm experiments have the potential for revolutionizing the field. Proposed low-frequency radio telescopes include LOFAR (Low Frequency Array) [1], the Mileura Wide Field Array (MWA) [94][2], and the next-generation SKA (Square Kilometer Array) [3]. The observational progress is an important driver for high-redshift mass function studies.

Theoretical studies of the mass function at high redshifts are challenging due to the small masses of the halos at early times. In order to capture these small-mass halos, high mass and force resolution are both required. For the large simulation volumes typical in cosmological studies, this necessitates a very large number of particles, as well as very high force resolution. Such simulations are very costly, and only a very limited number can be performed, disallowing exploration of a wide range of possible simulation parameters. Alternatively, many smaller volume simulation boxes, each with moderate particle loading, can be employed. This leads automatically to high force and mass resolution in grid codes (such as particle-mesh [PM]) and also reduces the costs for achieving sufficient resolution for particle codes (such as tree codes) or hybrid codes (such as

[1] See http://www.lofar.org
[2] See http://haystack.mit.edu/arrays/MWA/
[3] See http://www.skatelescope.org

TreePM). The disadvantages of this strategy are the limited statistics in individual realizations (because fewer halos form in a smaller box) and the unreliability of simulations below an intermediate redshift at which the largest mode in the box is still (accurately) linear. In addition, results from small boxes may be biased, since they only focus on a small region and volume. Therefore, one must show that the simulations are free from finite-volume artifacts, e.g. missing tidal forces, and run a sufficient number of statistically independent simulations to reduce the sample variance. Both strategies, employing large volume or multiple small-volume simulations, have been followed in the past in order to obtain results at high redshifts. In the following we provide a detailed discussion on previous findings as organized by simulation volume.

Small-Volume Simulations

Small-box simulations of side $\sim 1\,h^{-1}$Mpc have been performed by several groups. Using a treecode with softening length $0.4\,h^{-1}$kpc, and a $1\,h^{-1}$Mpc box with 128^3 particles, Jang-Condell & Hernquist [95] evolved their simulation from $z_{\rm in} = 100$ to $z = 10$. With a halo finder that combined overdensity criteria with an FOF algorithm, the mass function was determined over the range $10^{5.5} - 10^{8.1}\,h^{-1}M_\odot$, keeping halos with as few as eight particles. At $z = 10$ they found 'remarkably close agreement' with the PS fit but did not quantify the agreement explicitly.

In a series of papers, Yoshida et al. ran simulations with similar box sizes as above, most including the effects of gas dynamics. The simulations were performed with the TreePM/smoothed particle hydrodynamics code GADGET-II [96] and followed the evolution of $2\times\ 324^3$ particles (324^3 in the case of dark matter only), covering a halo mass range of 10^5-$10^{7.5}M_\odot$. All simulations were started at $z_{\rm in} = 100$ from 'glass' initial conditions [97, 98], in contrast to the grid-based initial conditions used here. The focus of their first paper [99] was the origin of primordial star-forming clouds. As part of that investigation, a dark-matter-only simulation in a $1.6\,h^{-1}$Mpc box was carried out. The halo density results for $z = 20$ to 32 lay systematically below the PS prediction, with the discrepancy being worse at high redshifts. The authors argued that this low abundance of halos was (possibly) due to finite-box-size effects. In the second paper [100], the mass function at $z = 20$ for a warm dark matter model was compared with CDM, with the simulation set up being very similar to their previous work [99], a 1 Mpc box started at $z = 100$. The results obtained were also similar; at $z = 20$ the CDM mass function was in good agreement with the PS fit. In a third paper, [101], a running spectral index was considered. Here results for a standard CDM mass function for a 1 Mpc box were given, this time at $z = 17$ and 22. Consistent with their previous results, they found good agreement with PS at these redshifts. (The FOF linking length used in the last paper was $b = 0.2$, while in the first two papers $b = 0.164$ was chosen. This

did not appear to make much of a difference, however.) These papers do not quantitatively compare the numerical mass function to the PS fit. (In contrast to these findings, a recent 1 Mpc box GADGET-II simulation with $z_{\rm in} \sim 120$ has been performed by Maio and collaborators [102] who find good agreement with the Warren fit as extrapolated by linear theory – in clear disagreement with PS.)

A similar strategy was followed by Cen et al. [103] who investigated dark matter halos in a mass range of $10^{6.5}$ to $10^9\,h^{-1}M_\odot$, using a TreePM code [104, 105]. The box size was taken to be $4\,h^{-1}$Mpc, the softening length was set at $0.14\,h^{-1}$kpc, 512^3 particles were used, and the simulations had a starting redshift of $z_{\rm in} = 53$. Halos were identified using the overdensity scheme DENMAX [106]. Among other quantities, they studied the mass function between $z = 11$ and 6 and found that the PS function 'provides a good fit' but without explicit quantification.

Overall, these small-box simulations, run with different codes and different halo finders, all found a 'depressed' mass function (see Fig. 3.1), consistent with PS and deviating very significantly from the predictions of the more modern fitting forms. In contrast, other simulations also using small boxes have come to quite different conclusions. For example, in Reed et al. [4], a large suite of different box sizes and simulations was used to cover the mass range between 10^5 and $10^{11.5}\,h^{-1}M_\odot$ at high redshift. The smallest boxes considered in this study were 1 h^{-1}Mpc on a side. The authors studied the halo mass function at redshifts out to $z = 30$, implementing a correction scheme to account for finite-box effects, as discussed in more detail below. Overall, their conclusion is that PS underestimates the mass function considerably (by at least a factor of 5 at high redshift and high masses), and ST overpredicts the halo abundance at high redshift.

Large-Volume Simulations

The large-box strategy is exemplified by a recent dark matter simulation with the GADGET-II code [86]. The evolution of 2160^3 particles in a $500\,h^{-1}$Mpc box was followed from $z_{\rm in} = 127$ until $z = 0$. The softening length was $5\,h^{-1}$kpc. The high mass and force resolution was sufficient to study the mass function reliably down to a redshift of $z = 10$, covering a mass range of 10^{10} to $10^{16}\,h^{-1}M_\odot$, with halos being identified by a standard FOF algorithm with $b = 0.2$. The results are consistent with the Jenkins fit, even though the mass function points at redshifts $z = 1.5$, 3.06, and 5.72 are slightly higher than the Jenkins fit and slightly lower for $z = 10$. No residuals were shown nor quantitative statements made.

Recently, two groups independently investigated cosmic reionization, providing mass function results at high redshift as part of this work. Iliev et al. [107] ran a PM simulation with PMFAST [108] in a $100\,h^{-1}$Mpc box with 1624^3 par-

ticles on a 3248^3 mesh. They present results for the mass function at redshifts between $z = 6$ and 18.5, using a spherical overdensity halo finder. At lower redshifts they find good agreement with ST, and at high redshift ($z > 10$) the results are closer to PS (because of their limited mass range, a more quantitative statement is difficult to make). Zahn et al. [109] ran a 1024^3 particle simulation (dark matter only) in a $65.6\,h^{-1}$Mpc box with GADGET-II and analyzed the FOF, $b = 0.2$ mass function out to $z = 20$. Between $z = 6$ and 14 they found good agreement with ST in the mass range of 10^9 to $10^{12} M_\odot$. At $z = 20$ they found that the simulation results were below ST but above PS, in relatively good agreement with the findings of [73] and [4].

Medium Volume Simulations

In 2003 paper, Reed et al. [3] chose a compromise between the large- and small-box strategies by picking a $50\,h^{-1}$Mpc box sampled with 432^3 particles. The tree code PKDGRAV was used to evolve the simulation from different starting redshifts between $z_{\rm in} = 139$ and 69 until $z = 0$. The smallest halo contained 75 particles, leading to a mass range of roughly 10^{10} to $10^{14.5}\,h^{-1} M_\odot$. Good agreement (better than 10%) was found with the ST fit up to $z \simeq 10$. For higher redshifts, the ST fit overpredicted the number of halos, up to 50% at $z = 15$. At this high redshift, statistics were lacking, and the resolution was not sufficient to resolve very small halos. A more recent $50\,h^{-1}$Mpc simulation with PMFAST with $z_{\rm in} = 60$ has been carried out by Trac & Cen [110] using a spherical overdensity definition of halo mass. In this work, the mass function, in the redshift range $6 < z < 15$, is found to be in very good agreement with PS, in gross contradiction with the results of most of the other simulations mentioned above. (This contradiction has recently been resolved by rerunning their simulation with $z_{in} = 300$ and identifying halos with a $b = 0.2$ FOF finder.)

Previous Work: Summary

In summary, there is considerable variation in the high-redshift ($z > 10$) mass function as found by different groups, independent of box size and simulation algorithm. Broadly speaking, the results fall into two classes: either consistent with linear theory scaling of a universal form (Jenkins, Reed, ST, or Warren) at low redshift [3, 4, 86, 73, 102, 109] or more consistent with the PS fit [95, 99, 100, 101, 103, 107, 110].

Our aim here is to determine the evolution of the mass function accurately, at the few percent level, and at the same time understand and characterize many of the numerical and physical factors that control the error in the mass function, and bring rise to discrepancies in reported results between different groups. In this work is analyzed a large suite of N-body simulations with varying box sizes between 4 and $256\,h^{-1}$Mpc, including many realizations of the small boxes, to study the mass function at redshifts up to $z = 20$ and to cover a large mass range

Table 3.2: Summary of the Performed Runs. The smallest halos we consider contain 40 particles. All simulations have 256^3 particles, evolved on 1024^3 grid.

Box Size (h^{-1}Mpc)	Resolution (h^{-1}kpc)	$z_{\rm in}$	$z_{\rm final}$	Particle Mass ($h^{-1}M_\odot$)	No. of Realizations
256	250	100	0	8.35×10^{10}	5
128	125	200	0	1.04×10^{10}	5
64	62.5	200	0	1.31×10^{9}	5
32	31.25	150	5	1.63×10^{8}	5
16	15.63	200	5	2.04×10^{7}	5
8	7.81	250	10	2.55×10^{6}	20
4	3.91	500	10	3.19×10^{5}	15

between 10^7 and $10^{13.5}\,h^{-1}M_\odot$. The number of small-box realizations is large in order to improve the statistics at high redshifts. Our results categorically rule out the PS fit as being more accurate than any of the more modern forms at *any* redshift up to $z = 20$, the discrepancy increasing with redshift.

3.2 The Code and the Simulations

All simulations in this work are carried out with the parallel PM code MC^2. This code solves the Vlasov-Poisson equations for an expanding universe. It uses standard mass deposition and force interpolation methods allowing periodic or open boundary conditions with second-order (global) symplectic time stepping and fast fourier transform based Poisson solves. Particles are deposited on the grid using the cloud-in-cell method. Code is presented in more details in the Appendix A. The overall computational scheme has proven to be accurate and efficient: relatively large time steps are possible with exceptional energy conservation being achieved. Finally, MC^2 has been extensively tested against state-of-the-art cosmological simulation codes (see chapter 6).

We use the following cosmology for all simulations:

$$\begin{aligned} &\Omega = 1.0, \quad \Omega_{\rm CDM} = 0.253, \quad \Omega_b = 0.048, \\ &\sigma_8 = 0.9, \quad H_0 = 70 \text{ km s}^{-1}\,\text{Mpc}^{-1}, \quad n = 1, \end{aligned} \tag{3.18}$$

in concordance with cosmic microwave background and large scale structure observations [111] (the third-year Wilkinson Microwave Anisotropy Probe observations suggest a lower value of σ_8; [112]). The transfer functions are generated with CMBFAST [113]. We summarize the different runs, including their force and mass resolution, in Table 3.2. As mentioned earlier, we identify halos with a standard FOF halo finder with a linking length of $b = 0.2$. Despite several shortcomings of the FOF halo finder, e.g., the tendency to link up two halos which are close to each other [114, 115] or statistical biases [8], the FOF algorithm itself is well defined and very fast. As discussed in §3.1.1, we adopt the

correction for sampling bias given by Warren when presenting our results.

3.3 Initial Conditions and Time Evolution

In a near-ideal simulation with very high mass and force resolution, the first halos would form very early. By $z = 50$, a redshift commonly used to start cosmological simulations, a large number of small halos would already be present (see, e.g. ref. [116] for a discussion of the first generation of star-forming halos). In a more realistic situation, however, the initial conditions at $z = 50$ have of course no halos, the particles having moved only the relatively small distance assigned by the initial Zel'dovich step. Only after the particles have traveled a sufficient distance and come close together can they interact locally to form the first halos. In the following we estimate the redshift when the Zel'dovich grid distortion equals the interparticle spacing, leading to the most conservative estimate for the redshift of possible first halo formation. From this estimate, we derive the necessary criterion for the starting redshift for a given box size and particle number.

3.3.1 Initial Redshift

In order to capture halos at high redshifts, we have found that it is very important to start the simulation sufficiently early. We consider two criteria for setting the starting redshift: (1) ensuring the linearity of all the modes in the box used to sample the initial matter power spectrum, and (2) restricting the initial particle move to prevent interparticle crossing and to keep the particle grid distortion relatively small. The first criterion is commonly used to identify the starting redshift in simulations. However, as shown below, it fails to provide sufficient accuracy of the mass functions, accuracy which can be obtained when a second (much more restrictive) control is applied. Furthermore, it is important to allow a sufficient number of expansion factors between the starting redshift $z_{\rm in}$ and the highest redshift of physical significance. This is needed to make sure that artifacts from the Zel'dovich approximation are negligible and that the memory of the artificial particle distribution imposed at $z_{\rm in}$ (grid or glass) is lost by the time any halo physics is to be extracted from the simulation results.

Although not studied here, it is important to note that high-redshift starts do require the correct treatment of baryons as noted in §3.1.2. In addition, redshift starts that are too high can lead to force errors for a variety of reasons, e.g., interpolation systematics, round-off, and correlated errors in tree codes.

Initial Perturbation Amplitude

The initial redshift in simulations is often determined from the requirement that all mode amplitudes in the box below the particle Nyquist wavenumber

Table 3.3: Initial Redshift Estimates from the Linearity of $\Delta^2(k_{\rm Ny})$. The number of particles is 256^3, the same in all simulations.

Box Size (h^{-1}Mpc)	$k_{\rm Ny}$ ($h\,{\rm Mpc}^{-1}$)	$T(z=0,\,k_{\rm Ny})$	$z_{\rm in}$
126	6.3	0.0002	33
32	25	$1.7{\cdot}10^{-5}$	45
16	50	$4.8{\cdot}10^{-6}$	50
8	100	$1.3{\cdot}10^{-6}$	55

characterized by $k_{\rm Ny}/2$ with $k_{\rm Ny} = 2\pi/\Delta_{\rm p}$, where $\Delta_{\rm p}$ is the mean interparticle spacing, be sufficiently linear. The smaller the box size chosen (keeping the number of particles fixed), the larger the largest k-value. Therefore, in order to ensure that the smallest initial mode in the box is well in the linear regime, the starting redshift must increase as the box size decreases. In the following we give an estimate based on this criterion for the initial redshift for different simulation boxes. We (conservatively) require the dimensionless power spectrum $\Delta^2 = k^3P(k)/2\pi^2$ to be smaller than 0.01 at the initial redshift. The initial power spectrum is given by

$$\Delta^2(k_{\rm Ny}, z_{\rm in}) = \frac{k^3P(k_{\rm Ny}, z_{\rm in})}{2\pi^2} \sim \frac{B\; k^{n+3}T^2(k_{\rm Ny}, z=0)}{2\pi^2(z_{\rm in}+1)^2}, \qquad (3.19)$$

where B is the normalization of the primordial power spectrum (see ref. [117] for a fitting function for B including COBE results) and $T(k)$ is the transfer function. We assume the spectral index to be $n = 1$, which is sufficient to obtain an estimate for the initial redshift. For a ΛCDM universe the normalization is roughly $B \sim 3.4 \times 10^6 (h^{-1}{\rm Mpc})^4$. Therefore, $z_{\rm in}$ is simply determined by

$$z_{\rm in} \simeq 4150\; k_{\rm Ny}^2 T(z=0, k_{\rm Ny}). \qquad (3.20)$$

We present some estimates for different box sizes in Table 3.3. For the smaller boxes ($< 8\; h^{-1}$Mpc), the estimates for the initial redshifts are at around $z_{\rm in} = 50$.

It is clear that this criterion simply sets a minimal requirement for $z_{\rm in}$ and neglects the fact that the initial particle move should be small enough to maintain the dynamical accuracy of perturbation theory (linear or higher order) used to set the initial conditions. Also, this criterion certainly does not tell us that if, e.g., $z_{\rm in} = 50$, then we may already trust the mass function at, say, $z = 30$. An example of this is provided by the results of Reed et al. [3], who find that their high-redshift results between $z = 7$ and 15 have not converge if they start their simulations at $z_{\rm in} = 69$. (A value of $z_{\rm in} = 139$ was claimed to be sufficient in their case.)

We now consider another criterion – ostensibly similar in spirit – that par-

ticles should not move more than a certain fraction of the interparticle spacing in the initialization step. This second criterion demands much higher redshift starts.

First Crossing Time

In cosmological simulations, initial conditions are most often generated using the Zel'dovich approximation [56]. Initially each particle is placed on a uniform grid or in a glass configuration and is then given a displacement determined by the relation

$$\mathbf{x} = \mathbf{q} - d(z)\nabla\Phi, \tag{3.21}$$

according to the discussion in §2.4.1, where the gradient of the potential Φ is independent of the redshift z. The Zel'dovich approximation holds in the mildly nonlinear regime, as long as particle trajectories do not cross each other (no caustics have formed). Studying the magnitude of $|\nabla\Phi|$ allows us to estimate two important redshift values: first, the initial redshift $z_{\rm in}$ at which the particles should not have moved on average more than a fraction of the interparticle spacing $\Delta_{\rm p} = L_{\rm box}/n_{\rm p}$, where $L_{\rm box}$ is the physical box size and $n_{\rm p}$ the number of particles in the simulation; second, the redshift at which particles first move more than the interparticle spacing, $z_{\rm cross}$, i.e., at which they have traveled on average a distance greater than $\Delta_{\rm p}$.

For a given realization of the power spectrum, the magnitude of $|\nabla\phi|$ depends on two parameters: the physical box size and the interparticle spacing. Together these parameters determine the range of scales under consideration. The smaller the box, the smaller the scales; therefore, $|\nabla\Phi|$ increases and both $z_{\rm in}$ and $z_{\rm cross}$ increase. Increasing the resolution has the same effect. In Figure 3.3 we show the probability distribution function for $|\nabla\phi|$ for three different box sizes, 8, 32, and $126\,h^{-1}$Mpc, representing values studied by other groups, as well as in this chapter. To make the comparison between the different box sizes more straightforward, we have scaled $|\nabla\Phi|$ with respect to the interparticle spacing $\Delta_{\rm p}$. All curves are drawn from simulations with 256^3 particles on a 256^3 grid, in accordance with the set up of our initial conditions. The behavior of the probability function follows our expectations: the smaller the box, or the higher the force resolution, the larger the initial displacements of the particles on average. From the mean and maximum values of such a distribution we can determine appropriate values for $z_{\rm in}$ and $z_{\rm cross}$. For our estimates we assume $d(z) \simeq 1/(1+z)$, which is valid for high redshifts. The maximum and rms initial displacements of the particles can then be easily calculated:

$$\delta_{\rm in}^{\rm max} \simeq \frac{\max(|\nabla\Phi|/\Delta_{\rm p})}{1+z_{\rm in}}, \tag{3.22}$$

$$\delta_{\rm in}^{\rm rms} \simeq \frac{\mathrm{rms}(|\nabla\Phi|/\Delta_{\rm p})}{1+z_{\rm in}}. \tag{3.23}$$

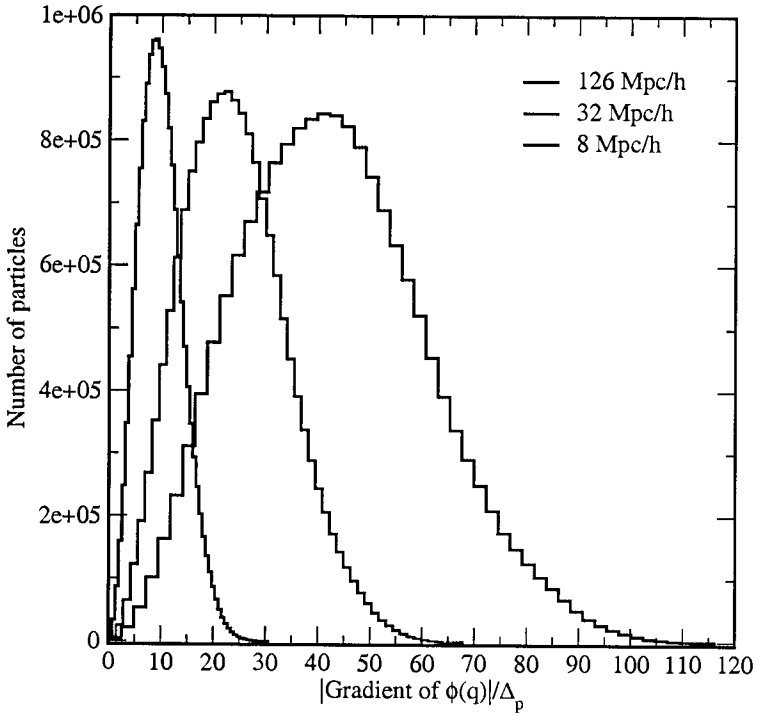

Figure 3.3: Probability distribution of $|\nabla\phi|$ in units of the interparticle spacing $\Delta_{\rm p}$. All curves shown are drawn from 256^3 particle simulations from an initial density grid of 256^3 zones. The physical box sizes are $126\,h^{-1}$Mpc (black line), $32\,h^{-1}$Mpc (red line), and $8\,h^{-1}$Mpc (green line). As expected, $\langle|\nabla\phi|\rangle$ increases with decreasing box size (which is equivalent to increasing force resolution). Therefore, $z_{\rm in}$ and $z_{\rm cross}$ are higher for the smaller boxes.

The very first 'grid crossing' of a particle occurs when $\delta_{\rm in}^{\rm max} = 1$; on average the particles have moved more than one particle spacing when $\delta_{\rm in}^{\rm rms} = 1$. This leads to the following estimates:

$$z_{\rm cross}^{\rm first} \simeq \max(\nabla\Phi/\Delta_{\rm p}) - 1, \tag{3.24}$$

$$z_{\rm cross}^{\rm rms} \simeq {\rm rms}(\nabla\Phi/\Delta_{\rm p}) - 1. \tag{3.25}$$

We show these two redshifts in Figure 3.4 for 10 different box sizes ranging from 1 to $512\,h^{-1}$Mpc and for 256^3 and 128^3 particles. The left panel shows the average redshift of the first crossing as a function of box size (which corresponds to the maximum in Fig. 3.3). The right panel shows the redshift where the first 'grid crossing' occurs (corresponding to the right tail in Fig. 3.3). To estimate the scatter in the results, we have generated five different realizations for each box. As expected, the small boxes show much more scatter. The average redshift of the first crossing in the $1\,h^{-1}$Mpc box varies between $z = 63$ and 83, while there is almost no scatter in the $512\,h^{-1}$Mpc box. Since $|\nabla\Phi|/\Delta_{\rm p}$ is independent of redshift in the Zel'dovich approximation, a simple scaling determines the appropriate initial redshift from these plots. For example, if a particle should

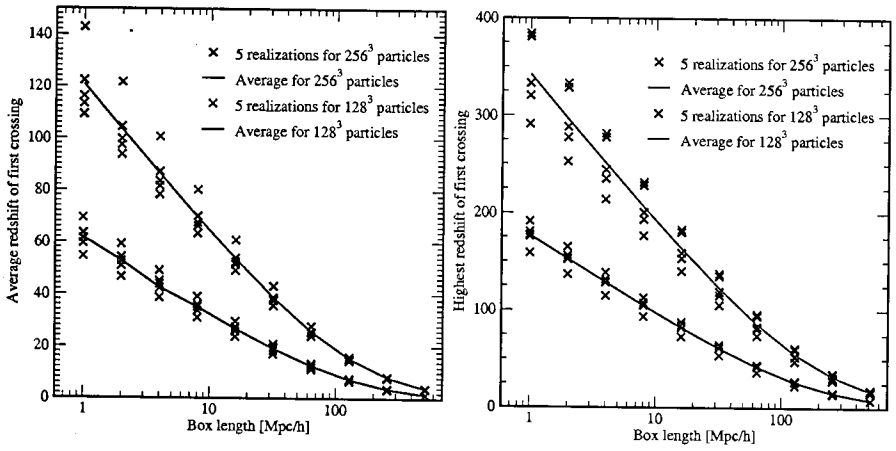

Figure 3.4: Average redshift of first crossing (left panel) and highest redshift of first crossing (right panel) as a function of box size. The initial conditions (five different realizations) are shown for boxes between 1 and 512 h^{-1}Mpc with 128^3 and 256^3 particles. For each initial condition, $z_{\rm cross}^{\rm first}$ and $z_{\rm cross}^{\rm rms}$ are shown by the crosses. The solid lines show the average from the five realizations. As expected, scatter from the different realizations is larger for smaller boxes. These plots provide estimates of the required initial redshift for a simulation since $|\nabla\Phi|/\Delta_{\rm p}$ is z-independent in the Zel'dovich approximation.

not have moved more than $0.3\Delta_{\rm p}$ on average at the initial redshift, the average redshift of first crossing has to be multiplied by a factor $1/0.3 = 3.\bar{3}$. For an $8\,h^{-1}$Mpc box this leads to a minimum starting redshift of $z = 230$, while for a $126\,h^{-1}$Mpc box this suggests a starting redshift of $z_{\rm in} = 50$. The 128^3 particle curve can be scaled to the 256^3 particle curve by multiplying by a factor of 2. Curves for different particle loadings can be obtained similarly.

3.3.2 Transients and Mixing

The Zel'dovich approximation matches the exact density and velocity fields to linear order in Lagrangian perturbation theory. Therefore, there is in principle an error arising from the resulting discrepancy with the density and velocity fields given by the exact growing mode initialized in the far past.

This error is linear in the number of expansion factors between $z_{\rm in}$ and the redshift of interest $z_{\rm phys}$. It has been explored in the context of simulation error by [5] and by [118]. Depending on the quantity being calculated, the number of expansion factors between $z_{\rm in}$ and $z_{\rm phys}$ required to limit the error to some given value may or may not be easy to estimate. For example, unlike quantities such as the skewness of the density field, there is no analytical result for how this error impacts the determination of the mass function. Neither does there exist any independent means of validating the result aside from convergence studies. Nevertheless, it is clear that to be conservative, one should aim for a factor of ~ 20 in expansion factor in order to anticipate errors at the several percent level,

a rule of thumb that has been followed by many N-body practitioners (and often violated by others!). This rule of thumb gives redshift starts that are roughly in agreement with the estimates in the previous subsection. Convergence tests done for our simulations show that the suppression in the mass function is very small (less than 1%) for simulations whose evolution covers a factor of 15 in the expansion factor and can be up to 20% for simulations that evolved by only 5 expansion factors. However, due to modest particle loads, we were unable to distinguish between the error induced by too few expansion factors and the breakdown of the Zel'dovich approximation.

Another possible problem, independent of the accuracy of the Zel'dovich approximation, is the initial particle distribution itself. Whether based on a grid or a glass, the small-distance ($k > k_{\rm Ny}$) mass distribution is clearly not sampled at all by the initial condition. Therefore, unlike the situation that would arise if a fully dynamically correct initial condition were given, some time must elapse before the correct small-separation statistics can be established in the simulation. Thus, all other things being equal, for the correct mass function to exist in the box, one must run the simulation forward by an amount sufficiently greater than the time taken to establish the correct small-scale power on first-halo scales while erasing memory on these scales of the initial conditions. If this is not done, structure formation will be suppressed, leading to a lowering of the halo mass function.

Because there is no fully satisfactory way to calculate $z_{\rm in}$ in order to compute the mass function at a given accuracy, we subjected *every* simulation box to convergence tests in the mass function while varying $z_{\rm in}$. The results shown here are all converged to the sub-percent level in the mass function. In the follofing is given an example of one such convergence test.

Initial Redshift Convergence Study

As mentioned above, we have tested and validated our estimates for the initial redshift for all the boxes used in the simulation suite via convergence studies. Here, we show results for an $8\,h^{-1}$Mpc box with initial redshifts $z_{\rm in} = 50$, 150, and 250 in Figure 3.5, where the mass functions at $z = 10$ are displayed. For the lowest initial redshift, $z_{\rm in} = 50$, the average initial particle movement is $1.87\Delta_{\rm p}$, while some particles travel as much as $5.03\Delta_{\rm p}$. This clearly violates the requirement that the initial particle grid distortion be kept sufficiently below 1 grid cell. The starting redshift $z_{\rm in} = 150$ leads to an average displacement of $0.63\Delta_{\rm p}$ and a maximum displacement of $1.71\Delta_{\rm p}$, and therefore just barely fulfills the requirements. For $z_{\rm in} = 250$ we find an average displacement in this particular realization of $0.37\Delta_{\rm p}$ and a maximum displacement of $1.00\Delta_{\rm p}$.

The bottom plot in each of the three panels of Figure 3.5 shows the ratio of the mass functions with respect to the Warren fit. In the middle and right panels the ratio for the largest halo is outside the displayed range. The mass

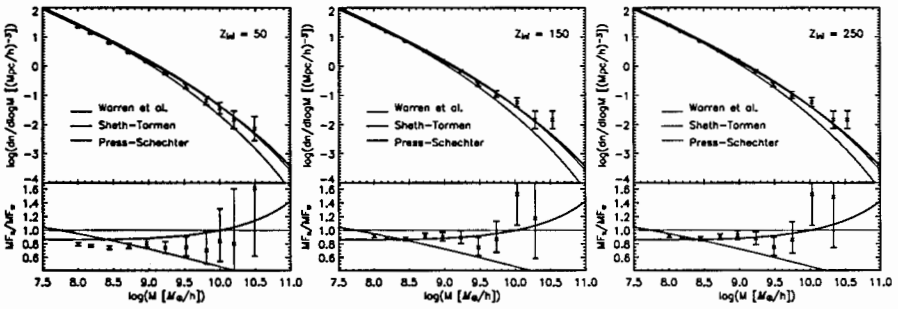

Figure 3.5: Dependence of the mass function on the initial redshift. The results are at $z = 10$ from three $8\,h^{-1}$Mpc box simulations with $z_{\rm in} = 50$ (left), $z_{\rm in} = 150$ (middle), and at $z_{\rm in} = 250$ (right). The mass function in the left panel is systematically lower than the other two by roughly 15%. Poisson error bars are shown.

function from the simulation started at $z_{\rm in} = 50$ (left panel) is noticeably lower, $\sim 15\%$, than for the other two simulations. The mass functions from the two higher redshift starts are in good agreement, showing that the choice for average grid distortion of approximately $0.3\Delta_{\rm p}$ is conservative, and that one can safely use $(0.5\text{–}0.6)\Delta_{\rm p}$. The general conclusion illustrated by Figure 3.5 is that if a simulation is started too late, halos are found to be missing over the entire mass range. With the late start, there is less time to form bound objects. Also, some particles that are still streaming towards a halo do not have enough time to join it. Both of these artifacts lead to an overall downshift of the mass function.

To summarize, requiring a limit on initial displacements sets the starting redshift much higher than simply demanding that all modes in the box stay linear. Indeed, the commonly used latter criterion (with $\delta^{\rm rms} \sim 0.1$) is not adequate for computing the halo mass function at high redshifts. One must verify that the chosen $z_{\rm in}$ sets an early enough start as shown here. Implications this has on the results from other groups with will be commented in §3.5.

3.3.3 Force and Mass Resolution

We now take up an investigation of the mass and force resolution requirements. The first useful piece of information is the size of the simulation box: from Figure 3.2 we can easily translate the number density into when the first halo is expected to appear in a box of volume V. For example, a horizontal line at $n = 10^{-6}$ would tell us at what redshift we would expect on average to find 1 halo of a certain mass in a $(100\,h^{-1}{\rm Mpc})^3$ box. The first halo of mass $10^{11} - 10^{12}\,h^{-1}M_\odot$ will appear at $z \simeq 15.5$, and the first cluster-like object of mass $10^{14} - 10^{15}\,h^{-1}M_\odot$ at $z \simeq 2$. Of course, these statements only hold if the mass and force resolution are sufficient to resolve these halos. The mass of a particle in a simulation, and hence the halo mass, is determined by three parameters: the matter content of the Universe Ω_m, including baryons and dark

matter, the physical box size $L_{\rm box}$, and the number of simulation particles $n_{\rm p}^3$:

$$m_{\rm particle} = 2.775 \times 10^{11} \Omega_m \left(\frac{L_{\rm box}}{n_{\rm p}\, h^{-1}{\rm Mpc}} \right)^3 h^{-1} M_\odot. \tag{3.26}$$

The required force resolution to resolve the chosen smallest halos can be estimated very simply. Suppose we aim to resolve a virialized halo with comoving radius r_Δ at a given redshift z, where Δ is the overdensity parameter with respect to the critical density $\rho_{\rm c}$. The comoving radius r_Δ is given by

$$r_\Delta = 9.51 \times 10^{-5} \left[\frac{\Omega(z)}{\Omega_m} \right]^{1/3} \left(\frac{1}{\Delta} \frac{M_{\Delta {\rm c}}}{h^{-1}{\rm M}_\odot} \right)^{1/3} h^{-1}{\rm Mpc}, \tag{3.27}$$

where $\Omega(z) = \Omega_m (1+z)^3 / [\Omega_m (1+z)^3 + \Omega_\Lambda]$ and the halo mass $M_{\Delta {\rm c}} = m_{\rm part} n_{\rm h}$, where $n_{\rm h}$ is the number of particles in the halo. We measure the force resolution in terms of

$$\delta_{\rm f} = \frac{L_{\rm box}}{n_{\rm g}}. \tag{3.28}$$

In the case of a grid code, $n_{\rm g}$ is literally the number of grid points per linear dimension; for any other code, $n_{\rm g}$ stands for the number of 'effective softening lengths' per linear dimension. To resolve halos of mass $M_{\Delta {\rm c}}$, a minimal requirement is that the code resolution be smaller than the radius of the halo we wish to resolve:

$$\delta_{\rm f} < r_\Delta. \tag{3.29}$$

Note that this minimal resolution requirement is aimed only at capturing halos of a certain mass, not at resolving their interior profile. Next, inserting the expression for the particle mass (eq. 3.26) and the comoving radius (eq. 3.27) into the requirement (eq. 3.29) and employing the relation between the interparticle spacing $\Delta_{\rm p}$ and the box size $\Delta_{\rm p} = L_{\rm box}/n_{\rm p}$, the resolution requirement reads

$$\frac{\delta_{\rm f}}{\Delta_{\rm p}} < 0.62 \left[\frac{n_h \Omega(z)}{\Delta} \right]^{1/3}. \tag{3.30}$$

We now illustrate the use of this simple relation with an example. Let $\Delta = 200$ and consider a ΛCDM cosmology with $\Omega_m = 0.3$. Then for PM codes for which $\delta_{\rm f}/\Delta_{\rm p} = n_{\rm p}/n_{\rm g}$, we have the following conclusions. If the number of mesh points is the same as the number of particles ($n_{\rm p} = n_{\rm g}$), halos with less than 2500 particles cannot be accurately resolved. If the number of mesh points is increased to 8 times the particle number ($n_{\rm p} = 1/2 n_{\rm g}$), commonly used for cosmological simulations with PM codes, the smallest halo reliably resolved has roughly 300 particles, and if the resolution is increased to a ratio of 1 particle per 64 grid cells, which we use in the main PM simulations in this work, halos with roughly 40 particles can be resolved. It has been shown [90] that this ratio (1:64) does not cause collisional effects and that it leads to consistent results in comparison to high-resolution codes. Note that increasing the resolution beyond

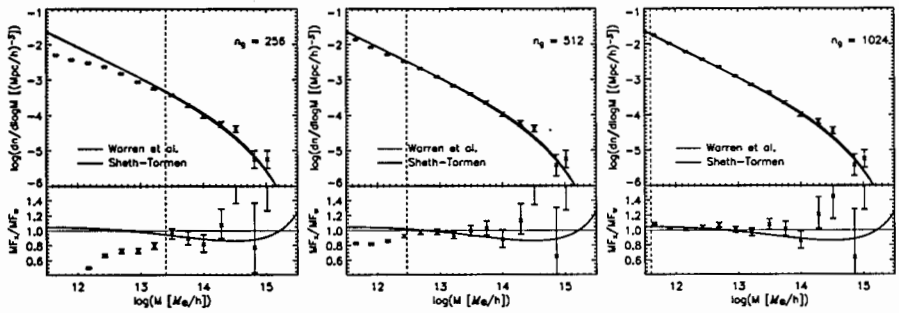

Figure 3.6: Convergence of the mass function as a function of force resolution. All results are shown at $z = 0$, for 256^3 particles and a $126\,h^{-1}$Mpc box with Poisson error bars. The resolution varies between 256^3 (left), 512^3 (middle), and 1024^3 grid points (right). The vertical line denotes the predicted theoretical resolution limit: halos on the right of the line should not be lost. The resolution limit is 2500 particles per halo for the 256^3 grid, 300 particles per halo for the 512^3 grid, and 40 particles per halo for the 1024^3 grid.

this point will not help, since it is unreliable to sample halos with too few particles. Note also that a similar conclusion holds for any simulation algorithm and not just for PM codes.

In Figure 3.6 we show results from a resolution convergence test at $z = 0$. We run 256^3 particles in a $126\,h^{-1}$Mpc box with three different resolutions: 0.5, 0.25, and $0.125\,h^{-1}$Mpc. The vertical line in each figure shows the mass below which the resolution is insufficient to capture all halos following condition (3.30). In all three cases, the agreement with the theoretical prediction is excellent.

3.3.4 Time Stepping

Next, we consider the question of time-step size and estimate the minimal number of time steps required to resolve the halos of interest. We begin with a rough estimate of the characteristic particle velocities in halos. For massive halos, the halo mass M_{200} and its velocity dispersion are connected by the approximate relation [119]:

$$M_{200} \simeq \frac{10^{15}\,h^{-1}M_\odot}{H/H_0}\left(\frac{\sigma_v}{1080\,\mathrm{km/s}}\right)^3. \tag{3.31}$$

There is even more accurate expression [120], but the above is more than sufficient for our purposes. At high redshift, Ω_Λ can be neglected, and we can express the velocity dispersion as a function of redshift:

$$\sigma_\mathrm{v} \simeq 10^{-2}\sqrt{1+z}\left(\frac{M_{200}}{h^{-1}M_\odot}\right)^{1/3}\mathrm{kms}^{-1}. \tag{3.32}$$

In a time δt, the characteristic scale length δl is given by $\delta l \simeq \sigma_\mathrm{v}\delta t$ or

$$\delta t \simeq \frac{\delta l}{\sigma_v} = \frac{100\,\delta l/\mathrm{km}}{\sqrt{1+z}}\left(\frac{M_{200}}{h^{-1}M_\odot}\right)^{-1/3}\mathrm{s}. \tag{3.33}$$

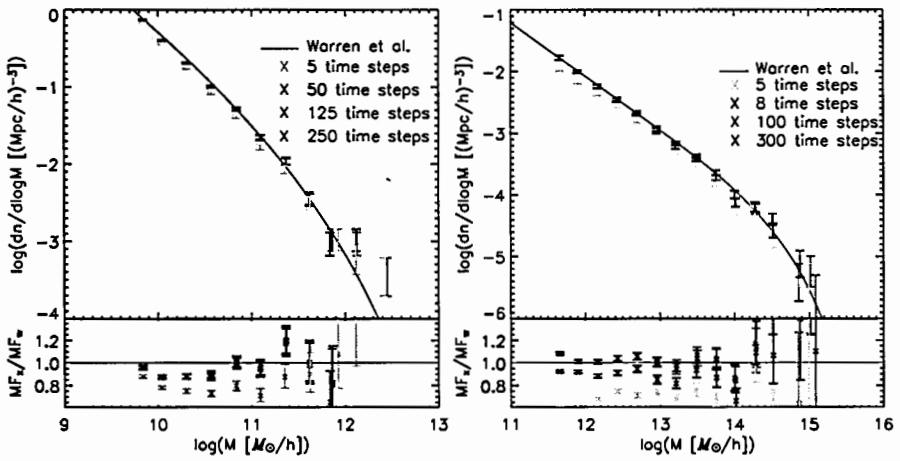

Figure 3.7: Left panel: One of the $32\,h^{-1}$Mpc box realizations run with 250, 125, 50 and 5 time steps between $z_{\rm in} = 150$ and $z_{\rm final} = 5$. The mass function is shown at the final redshift $z = 5$. Data points for all runs except the one with five time steps are so close that they are difficult to distinguish. Right panel: A $126\,h^{-1}$Mpc box with 300, 100, 8, and 5 time steps between $z_{\rm in} = 50$ and $z_{\rm final} = 0$. The agreement for the very large halos for 100 and 300 time steps is essentially perfect. Poisson error bars are shown.

Expressed in terms of the scale factor, equation (3.33) reads:

$$\delta a \simeq 10^4 \frac{\delta l}{h^{-1}\mathrm{Mpc}} \left(\frac{M_{200}}{h^{-1}M_\odot}\right)^{-1/3}. \tag{3.34}$$

We are interested in the situation where δl is actually the force resolution, $\delta_{\rm f}$. In a single time step, the distance moved should be small compared to $\delta_{\rm f}$; i.e., the actual time step should be smaller than δa estimated from the above equation when δl is replaced on the right–hand side with $\delta_{\rm f}$. Let us consider a concrete example for the case of a PM code where $\delta_{\rm f} = L_{\rm box}/n_{\rm g}$ as explained earlier. For a 'medium' box size of $L_{\rm box} = 256\,h^{-1}$Mpc and a grid size of $n_{\rm g} = 1024$, $\delta_{\rm f} = 0.25\,h^{-1}$Mpc. For a given box, the highest mass halos present have the largest $\sigma_{\rm v}$ and give the tightest constraints on the time step. For the chosen box size, a good candidate halo mass scale is $M_{200} \sim 10^{15}\,h^{-1}M_\odot$ (this could easily be less, but it does not change the result much). In this case,

$$\delta a \simeq 0.025. \tag{3.35}$$

If, for illustration, we start a simulation at $z = 50$ and evolve it down to $z = 0$, this translates to roughly 40 time steps. It has to be stressed that this estimate is aimed only at avoiding disruption of the halos themselves, and is certainly not sufficient to resolve the *inner* structure of the halo.

In Figure 3.7 we show two tests of the time step criterion. The left panel shows the result from a $32\,h^{-1}$Mpc box at redshift $z = 5$. The simulation starts

at $z_{\rm in} = 150$ and is evolved with 50, 125, and 250 time steps down to $z = 5$. Following the argument above for this box size, one would expect all three choices to be acceptable, and the excellent agreement across these runs testifies that this is indeed the case. We also carried out a run with only five time steps, which yields a clearly lower ($\sim 20\%$) mass function than the others, but not as much as one would probably expect from such an imprecise simulation.

The right panel shows the results from a $126\,h^{-1}$Mpc box at $z = 0$. This simulation was started at $z$in $= 50$ and run to $z = 0$ with 5, 8, 100, and 300 time steps. Again, as we would predict, the agreement is very good for the last two simulations, and the convergence is very fast, confirming our estimate that only $\mathcal{O}(10)$ time steps is enough to get the correct halo mass function. Overall, the halo mass function appears to be a very robust measure, not very sensitive to the number of time steps. Nevertheless, we used a conservatively large number of time steps: 500 for the simulations stopping at $z = 0$ and 300 for those stopping at $z = 10$.

In the previous subsections we have discussed and tested different error control criteria for obtaining the correct simulated mass function at all redshifts. These criteria are (1) a sufficiently early starting redshift to guarantee the accuracy of the Zel'dovich approximation at that redshift and provide enough time for the halos to form; (2) sufficient force and mass resolution to resolve the halos of interest at any given redshift; and (3) sufficient numbers of time steps. Violating any of these criteria *always* leads to a suppression of the mass function. Most significantly, these tests show that a late start (i.e., starting redshift too low) leads to a suppression over the entire mass range under consideration, and is a likely explanation of the low mass function results in the literature. As intuitively expected, insufficient force resolution leads to a suppression of the mass function at the low-mass end, while errors associated with time stepping are clearly subdominant and should not be an issue in the vast majority of simulations.

3.4 Results and Interpretation

In this section we present the results from our simulation suite. We describe how the data are obtained as well as the post-processing corrections applied. The latter include compensation for FOF halo mass bias induced by finite (particle number) sampling, and the (small) systematic suppression of the mass function induced by the finite volume of the simulation boxes.

3.4.1 Binning of Simulation Data

Before venturing into the simulation results, we first describe how they were obtained and reported from individual simulations. We used narrow mass bins while conservatively keeping the statistical shot noise of the binned points no

worse than some given value. Bin widths $\Delta \log M$ were chosen such that the bins contain an equal number of halos $N_{\rm h}$. The worst-case situation occurs at $z = 20$ for the $8\,h^{-1}$Mpc box, which has $N_{\rm h} = 80$; the $4\,h^{-1}$Mpc box at the same redshift has $N_{\rm h} = 400$. At $z = 15$ we have $N_{\rm h} = 150$, 1600, and 3000 for box sizes 16, 8, and $4\,h^{-1}$Mpc, respectively. At $z = 10$ the smallest value $N_{\rm h} = 450$ is for the $32\,h^{-1}$Mpc box, while at $z = 5$ and 0 we essentially always have $N_{\rm h} > 10000$.

With a mass function decreasing monotonically with M, this binning strategy results in bin widths increasing monotonically with M. The increasing bin size may cause a systematic deviation – growing towards larger masses – from an underlying 'true' continuous mass function. The data points for the binned mass function give the average number of halos per volume in a bin,

$$\bar{F} \equiv N_{\rm h}/(V \Delta \log M), \tag{3.36}$$

plotted versus an average halo mass, averaged by the number of halos in the bin:

$$\bar{M} \equiv \sum_{\rm bin} M/N_{\rm h}. \tag{3.37}$$

Assuming that the true mass function $dn/d\log M$ has some analytic form $F(M)$, a systematic deviation due to the binning prescription

$$\epsilon_{\rm bin} \equiv \frac{\bar{F} - F(\bar{M})}{F(\bar{M})} \tag{3.38}$$

can be evaluated by computing $\bar{F}$ and $\bar{M}$ as

$$\bar{F} = \frac{\int_{\Delta M} dn}{\Delta \log M}, \quad \bar{M} = \frac{\int_{\Delta M} M dn}{\int_{\Delta M} dn}, \tag{3.39}$$

where $dn \equiv F(M)\, d\log M$ and the integrations are over a mass range $[M, M + \Delta M]$. For the leading-order term of the Taylor expansion of $\epsilon_{\rm bin}(\Delta M)$, we find

$$\epsilon_{\rm bin} \simeq \frac{F'' - 2(F')^2/F}{24F} (\Delta M)^2, \tag{3.40}$$

where the primes denote $\partial/\partial M$. A characteristic magnitude of this $\epsilon_{\rm bin}$ for a general $F(M)$ is $(\Delta M/M)^2/24$. However, in our case, where the relevant scales $k \gg k_{\rm eq} \sim 0.01 h{\rm Mpc}^{-1}$, $\epsilon_{\rm bin}$ has a much stronger suppression, as explained below.

We know that the mass function is close to the universal form,

$$F(M) = \frac{\rho_{\rm b}}{M} f(\sigma) \frac{d\ln \sigma^{-1}}{d\log M} \tag{3.41}$$

(see, eq. 3.1). Note that for $k \gg k_{\rm eq}$, $\sigma^{-1}(M)$ is a slowly varying function, i.e.,

$$\frac{d\log\sigma^{-1}}{d\log M} \equiv \frac{n_{\rm eff}+3}{6} \tag{3.42}$$

is much smaller than unity, and the derivative $d\log\sigma^{-1}/d\log M$ also changes slowly with M. Then, despite the steepness of $F(\sigma)$ at small σ, the factor $f(\sigma)\,d\ln\sigma^{-1}/d\log M$ in equation (3.41) depends weakly on M. Therefore, the mass function $F(M)$ is close to being inversely proportional to M. In the limit of exact inverse proportionality, $F \propto M^{-1}$, equation (3.40) tells us that $\epsilon_{\rm bin} \to 0$. This effective cancellation of the two terms on the right-hand side of equation (3.40) makes the binning error negligible to the accuracy of our $F(M)$ reconstruction whenever a bin width $\Delta\log M$ does not exceed 0.5. To confirm the absence of any systematic offsets due to the binning, we binned the data into $\log M$ intervals 5 times narrower and wider, with no apparent change in the inferred $F(M)$ dependence.

It should be remarked that the situation could be quite different with another binning choice. For example, if the binned masses $\bar{M}$ were chosen at the centers of the corresponding $\log M$ intervals, $\log\bar{M} = [\log M + \log(M+\Delta M)]/2$, the systematic binning deviation

$$\epsilon_{\rm bin}^{\rm (center)} \simeq \frac{F''+F'/M}{24F}(\Delta M)^2 \tag{3.43}$$

would have no special cancellation for the studied type of mass function. A corresponding binning error would be about 2 orders of magnitude larger than that of equations (3.36) and (3.37).

The statistical error bars used are Poisson errors, following the improved definition of Heinrich [121]:

$$\sigma_\pm = \sqrt{N_{\rm h}+\frac{1}{4}} \pm \frac{1}{2}. \tag{3.44}$$

At large values of N_h, these error bars asymptote to the familiar form $\sqrt{N_{\rm h}}$. At smaller values of $N_{\rm h}$ – which are of minor concern here – equation 3.44 has several advantages over the standard Poisson error definition, some being (1) it is nonzero for $N_{\rm h} = 0$; (2) the lower edge of the error bar does not go all the way to zero when $N_{\rm h} = 1$; (3) the asymmetry of the error bars reflects the asymmetry of the Poisson distribution.

Finally, as noted earlier and discussed in the next section, all the results shown in the following include a correction for the sampling bias of FOF halos according to equation (3.45). This mass correction brings down the low-mass end of the mass function.

3.4.2 FOF Mass Correction

The mass of a halo as determined by the FOF algorithm displays a systematic bias with the number of particles used to sample the halo. Too few particles lead to an increase in the estimated halo mass. By systematically subsampling a large halo population from N-body simulations (at $z = 0$), Warren determined an empirical correction for this undersampling bias. For a halo with $n_{\rm h}$ particles, his correction factor for the FOF mass is given by

$$n_{\rm h}^{\rm corr} = n_{\rm h} \left(1 - n_{\rm h}^{-0.6}\right) . \tag{3.45}$$

We have carried out an independent exercise to check the systematic bias of the FOF halo mass as a function of particle number based on Monte Carlo sampling of an NFW halo mass profile with varying concentration and particle number, as well as by direct checks against simulations (e.g., Fig. 3.8); our results are broadly consistent with equation (3.45). Details are presented in the chapter 4.

The effect of the FOF sampling correction can be quickly gauged by considering a few examples: for a halo with 50 particles, the mass reduction is almost 10%, for a halo with 500 particles, it is $\sim$ 2.4%, and for a well-sampled halo with 5000 particles, it is only 0.6%. As a cautionary remark, this correction formula does not represent a general recipe but can depend on variables such as the halo concentration. Since the conditions under which different simulations are carried out can differ widely, corrections of this type should be checked for applicability on a case-by-case basis. Note also that the correction for the mass function itself depends on how halos move across mass bins once the FOF correction is taken into account.

The choice of the mass function range in a given simulation box always involves a compromise: too wide a dynamic range leads to poor statistics at the high-mass end and possible volume-dependent systematic errors, and too narrow a range leads to possible undersampling biases. Our choice here reflects the desire to keep good statistical control over each mass bin at the expense of wide mass coverage, compensating for this by using multiple box sizes. Therefore, in our case it is important to demonstrate control over the FOF mass bias. An example of this is shown in Figure 3.8, where results from four box sizes demonstrate the successful application of the Warren correction to simulation results at $z = 10$.

3.4.3 Simulation Mass and Growth Function

The complete set of simulations, summarized in Table 3.2, allows us to study the mass function spanning the redshift range from $z = 20$ to 0. The mass range covers dwarf to massive galaxy halos at $z = 0$ (cluster scales are best covered by bigger boxes as in Warren and chapter 4), and at higher redshifts goes down to $10^7\, h^{-1} M_\odot$, the mass scale above which gas in halos can cool via atomic line

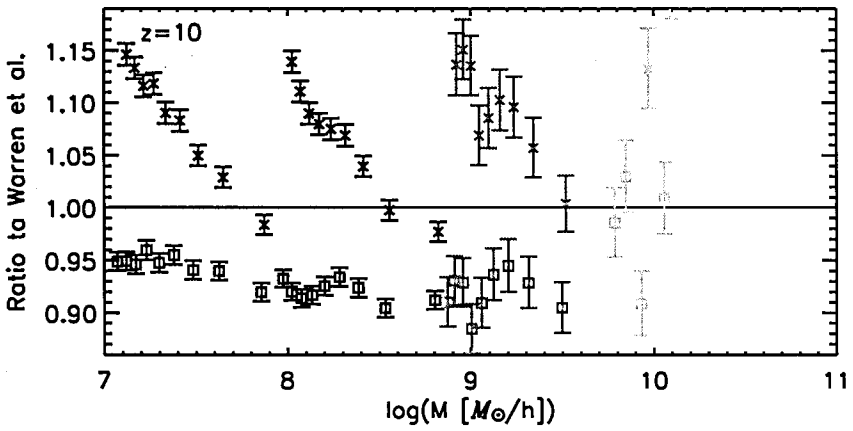

Figure 3.8: FOF mass correction for halos in 4 (dark blue), 8 (black), 16 (light blue), and 32 (yellow) h^{-1}Mpc boxes. To show the effect clearly, we plot the ratio of our data to the Warren fit. Crosses show the uncorrected mass function and squares the mass function after correction, following eq. (3.45). Note the smooth behavior of the corrected mass function as opposed to the mass-function jumps across box sizes for the uncorrected data.

cooling [122].

3.4.4 Time Evolution of the Mass Function

Halo mass functions from the multiple-box simulations are shown in Figure 3.9, with results being reported at five different redshifts with no volume corrections applied. The combination of box sizes is necessary because larger boxes do not have the mass resolution to resolve very small halos at early redshifts, while smaller boxes cannot be run to low redshifts. The bottom plot of each panel shows the ratio of the numerically obtained mass function, and various other fits, to the Warren fit as scaled by linear theory (for volume-corrected results, see Fig. 3.11). Displaying the ratio has the advantage over showing relative residuals that large discrepancies (more than 100%) appear more clearly. For all redshifts, the agreement with the Warren fit is at the 20% level. The ST fit matches the simulations for small masses very well but overpredicts the number of halos at large masses. This overprediction becomes worse at higher redshifts. For example, at $z = 15$ ST overpredicts halos of $10^9\,h^{-1}M_\odot$ by a factor of 2. Agreement with both Reed et al. fits is also good, within the 10% level. Discussion focused around the question of universality is given in §3.4.7. The PS fit in general is not satisfactory over a larger mass range at any redshift. It crosses the other fits at different redshifts for different masses. Away from this crossing region, however, the disagreement can be as large as an order of magnitude, e.g. for $z = 20$ over the entire mass range we consider here.

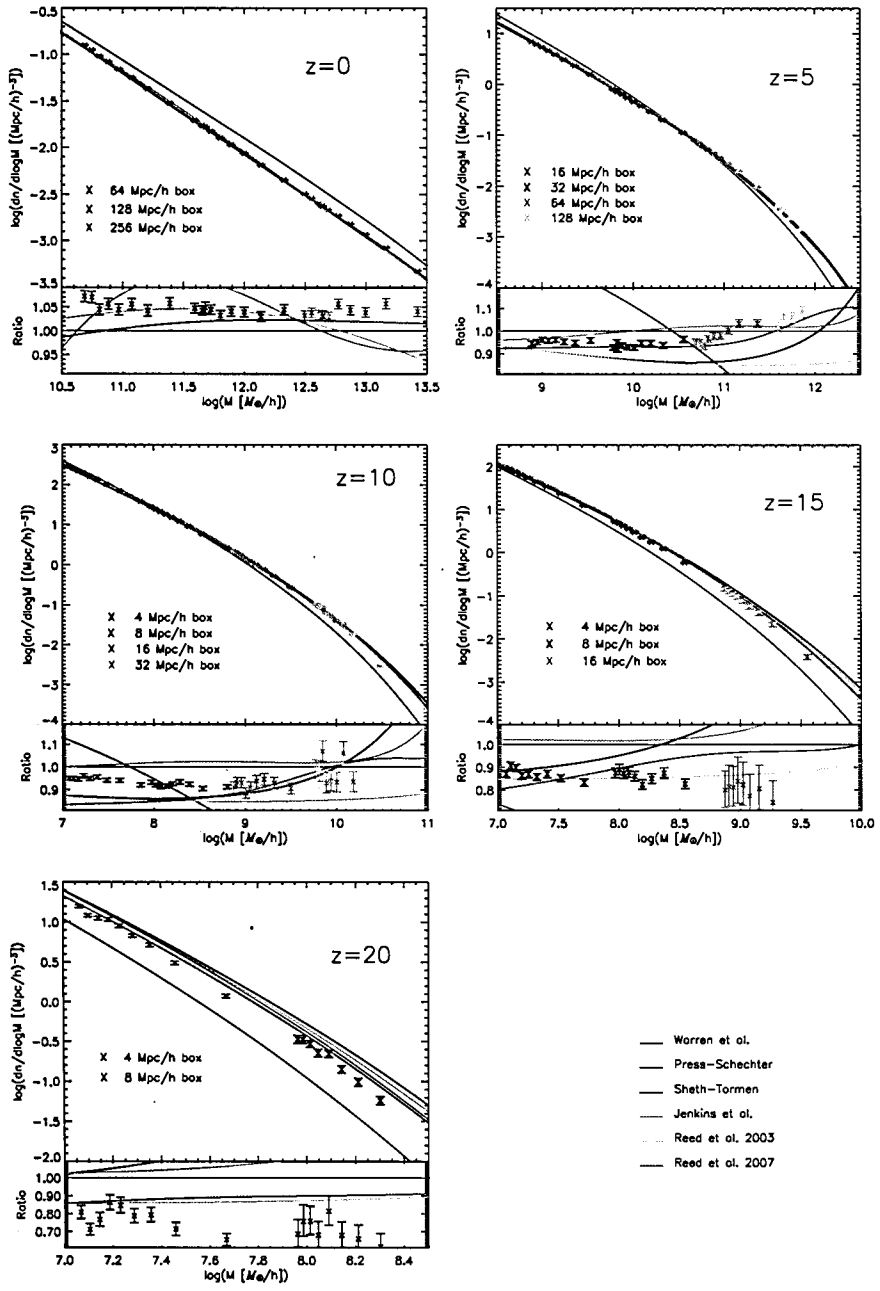

Figure 3.9: Mass function at five different redshifts ($z = 0$, 5, 10, 15, and 20; top to bottom) compared to different fitting formulae. Note that the mass ranges are different at different redshifts. The simulation results have been corrected for FOF bias following Warren but not for finite-volume effects (for these, see Fig. 12). The bottom panel shows the ratio with respect to the Warren fit. Our simulations agree with the Warren fit at the 10% level for redshifts smaller than 10, although there is a systematic offset of 5% at $z = 0$, where our numerical results are higher than the fit. At higher redshifts, the agreement is still very good (at the 20% level) and becomes very close once finite-volume corrections are applied (Fig. 3.11). PS is a bad fit at all redshifts, and especially at high redshifts, where the difference between PS and the simulation results is an order of magnitude.

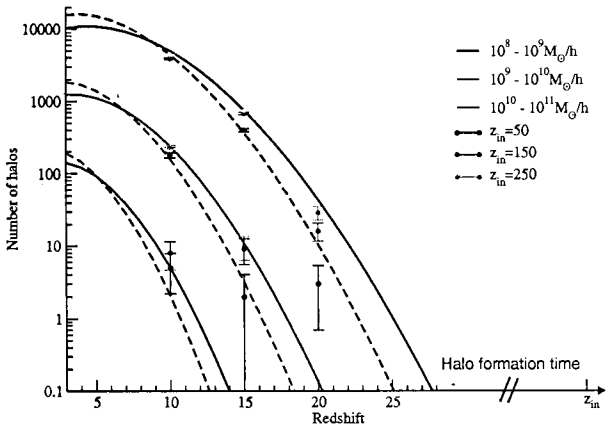

Figure 3.10: Halo growth function for an $8\,h^{-1}$Mpc box started from three different redshifts. The blue data points results from the $z = 50$ start, the turquoise data points from the $z = 150$ start, and orange from the $z = 250$ start, which is the redshift satisfying our starting criteria. The two fits shown are the Warren fit (solid line) and the PS fit (dashed line). Three different mass bins are shown. It is interesting to note that the late start seems to follow the PS fit at high redshift.

3.4.5 Halo Growth Function

As discussed in §3.1.4 the halo growth function (the number density of halos in mass bins as a function of redshift) offers an alternative avenue to study the time evolution of the mass function. Figure 3.10 shows the halo growth function for an $8\,h^{-1}$Mpc box for three different starting redshifts, $z_{\rm in} = 50$, 150, and 250 (these are the same simulations as in Fig. 3.5). The results are displayed at three redshifts, $z = 20$, 15, and 10 and for three mass bins, $10^8 - 10^9\,h^{-1}M_\odot$, $10^9 - 10^{10}\,h^{-1}M_\odot$, and $10^{10} - 10^{11}\,h^{-1}M_\odot$.

Assuming that the Warren fit scales at least approximately to high redshifts, the first halos in the lowest mass bin are predicted to form at $z_{\rm form} \sim 25$ (see Fig. 3.5). We have found that if $z_{\rm form}$ is not sufficiently far removed from $z_{\rm in}$, formation of the first halos is significantly delayed/suppressed. In turn, this leads to suppressions of the halo growth function and the mass function at high redshifts. As shown in Figure 3.10, the suppression can be quite severe at high redshifts: the simulation result at $z = 20$ from the late start at $z_{\rm in} = 50$ is an *order of magnitude* lower than that from $z_{\rm in} = 250$. At lower redshifts, the discrepancy decreases, and results from late-start simulations begin to catch up with the results from earlier starts. Coincidentally, the suppression due to the late start at $z_{\rm in} = 50$ is rather close to the PS prediction which is very significantly below the Warren fit in the mass and redshift range of interest (see Fig. 3.10).

3.4.6 Finite-Volume Corrections

The finite size of simulation boxes can compromise results for the mass function in multiple ways. It is important to keep in mind that finite-volume boxes cannot be run to lower than some redshift, $z_{\rm final}$, the stopping point being determined by when nonlinear scales approach close enough to the box size. Approaching too near this point delays the ride-up of nonlinear power towards the low-k end, with a possible suppression of the mass function.

As a consequence of this delay, the evolution (incorrectly) appears more linear at large scales than it actually should, as compared to the $P(k)$ obtained in a much bigger box. Therefore, verifying linear evolution of the lowest k-mode is by itself *not* sufficient to establish that the box volume chosen was sufficiently large. For all of our overlapping-volume simulations we have checked that the power spectra were consistent across boxes up to the lowest redshift from which results have been reported (Table 3.2 lists the stopping redshifts).

Aside from testing for numerical convergence, it is important to show that finite-volume effects are also under control, especially any suppression of the mass function with decreasing box size (due to lack of large-scale power on scales greater than the box size). Several heuristic analyses of this effect have appeared in the literature. Rather than rely solely on the unknown accuracy of these results, however, here are also numerically investigated possible systematic differences in the mass function with box size.

Over the redshifts and mass ranges probed in each of our simulation boxes, we find no direct evidence for an error caused by finite volume (at more than the $\sim 20\%$ level), as can be seen in Figure 3.9. This is not to say that there are no finite-volume effects (the very high-mass tail in a given box must be biased low simply from sampling considerations) but that their relative amplitude is small. Below we discuss how to correct the mass function for finite box size.

Volume Corrections from Universality

Let us first assume that mass function universality holds strictly, in other words, that for any initial condition the number of halos can be described by a certain scaled mass function (eq. 3.1) in which $\sigma(M)$ is the variance of the top-hat-smoothed linear density field. In the case of infinite simulation volume, $\sigma(M)$ is determined by equation (3.3), and the mass function $F(M)$ of equation (3.2) is

$$F(M) \equiv \frac{dn}{d\log M} = \frac{\rho_b}{M} f(\sigma) \frac{d\ln\sigma^{-1}}{d\log M} \ . \tag{3.46}$$

In an ensemble of finite-volume boxes, however, one necessarily measures a different quantity:

$$F'(M') \equiv \frac{dn'}{d\log M'} = \frac{\rho_b}{M'} f(\sigma') \frac{d\ln\sigma'^{-1}}{d\log M'} \ . \tag{3.47}$$

Here $\sigma'(M')$ is determined by the (discrete) power spectrum of the simulation ensemble, although if universality holds as assumed, f in equations (3.46) and (3.47) is the same function.

Since we are, in general, interested in the mass function which corresponds to an infinite volume, we can then correct the data obtained from our simulations as follows: for each box size we can define a function $M'(M)$ such that

$$\sigma(M) \equiv \sigma'(M'(M)). \tag{3.48}$$

Using equations (3.46) – (3.48), we determine $F(M)$ as

$$F(M) = F'(M') \frac{dM'(M)}{dM}. \tag{3.49}$$

Thus, the corrected number of halos in each bin is calculated as

$$dn = dn' \frac{M'}{M} \ . \tag{3.50}$$

The universality must eventually break down for sufficiently small boxes or high accuracy because the nonlinear coupling of modes is more complicated than that described by the smoothed variance. This violation can be partly corrected for by modifying the functional form of $\sigma'(M')$. Therefore, we also explore other choices of $\sigma'(M')$ which may better represent the mass function in the box. To address this question we provide a short summary of the Press-Schechter approach.

Motivation from Isotropic Collapse

We first consider the idealized case of a random isotropic perturbation of pressureless matter and assume that the primordial overdensity at the center of this perturbation has a Gaussian probability distribution. The probability of local matter collapse at the center is then fully determined by the local variance of the primordial overdensity σ^2. Consequently, for the isotropic case the contribution of Fourier modes of various scales to the collapse probability is fully quantified by their contribution to σ^2.

To see this, consider the evolution of matter density $\rho_{\rm loc}$ at the center of the spherically symmetric density perturbation. For transparency of argument, let us focus on the evolution during the matter-dominated era; it is straightforward to generalize the argument to include a dark energy component $\rho_{\rm de}(z)$, homogeneous on the length scales of interest, by a substitution $\rho_{\rm loc} \to \rho_{\rm m,\,loc} + \rho_{\rm de}$ in equations (3.51) and (3.53). By Birkhoff's law, the evolution of $\rho_{\rm loc}$ and the central Hubble flow $H_{\rm loc} \equiv \frac{1}{3}\nabla \cdot \mathbf{v}_{\rm loc}$ are governed by the closed set of the Friedmann and conservation equations,

$$H_{\rm loc}^2 \ = \ \frac{8\pi G \rho_{\rm loc}}{3} - \frac{\kappa}{a_{\rm loc}^2}, \tag{3.51}$$

$$a_{\rm loc} \equiv \left(\frac{\rho_0}{\rho_{\rm loc}}\right)^{1/3}, \quad \frac{da_{\rm loc}}{dt} = H_{\rm loc} a_{\rm loc}, \tag{3.52}$$

where κ is a constant determined by the initial conditions, ρ_0 is arbitrary (e.g., $\rho_0 = \rho_b|_{z=0}$), and t is the proper time.

The degree of nonlinear collapse at the center can be quantified by a dimensionless parameter

$$q \equiv 1 - \frac{3H_{\rm loc}^2}{8\pi G \rho_{\rm loc}}. \tag{3.53}$$

First consider early times, when the evolution is linear, and let $\rho_{\rm loc} = \rho_b(1+\delta)$. Then for the growing perturbation modes during matter domination $H_{\rm loc} = \bar{H}(1 - \delta/3)$. Given these initial conditions, which set the initial $\rho_{\rm loc}$ and the constant κ in equation (3.51), the subsequent evolutions of $\rho_{\rm loc}$, $H_{\rm loc}$, and therefore q are determined unambiguously.

During the linear evolution in the matter era $q = 5\delta/3$ is small and grows proportionally to the cosmological scale factor a. For positive overdensity, nonlinear collapse begins when q becomes of order unity, reaching its maximal value $q = 1$ when $H_{\rm loc} = 0$, and decreasing rapidly afterwards. (We can observe the latter by rewriting eq. (3.53) as

$$q = \frac{3\kappa}{8\pi G a_{\rm loc}^2 \rho_{\rm loc}} \propto \rho_{\rm loc}^{-1/3}, \tag{3.54}$$

having applied eqs. (3.51) and (3.52).) Nonlinear collapse of matter at the center of the considered region can be said to occur either when $q \to 0$ or when q reaches a critical 'virialization' value q_c.

Now it is easy to argue that in the isotropic case the Press-Schechter approach gives the true probability of the collapse, $P(q > q_c, z)$, for a redshift z. Indeed, the evolution of q is set deterministically by the primordial density perturbation at the center; for adiabatic initial conditions specifically, it is set by the curvature perturbation ζ at the center. Since higher values of ζ lead to earlier collapse,

$$P(q > q_c, z) = P(\zeta > \zeta_c(z)) = \frac{1}{2}\,\mathrm{erfc}\left[\frac{\zeta_c(z)}{\sqrt{2}\,\sigma}\right], \tag{3.55}$$

where the last equality uses the explicit form of $P(\zeta)$ as a Gaussian distribution with a variance σ^2.

If the considered isotropic distribution is confined by a (spherical) boundary and σ at the center is reduced by removal of large-scale power, then equation (3.55) should accurately describe the corresponding change of the collapse probability. In numerical simulations, due to the imposition of periodic boundary conditions, there is no power on scales larger than the box size. In this case the variance σ should be specified by the analogue of equation (3.3) with the integral replaced by a sum over discrete modes.

For the mass function (eq. [3.46]), a constant reduction of the variance $\sigma^2(M)$ due to the removal of large-scale power leads to a suppression of the mass function at the high-mass end and, counterintuitively, a boost at the low-mass end. The latter is easily understood as follows: The σ-dependent terms of equation (3.46),

$$f(\sigma)\frac{d\ln\sigma^{-1}}{d\log M} = \frac{d\rho(M)/\rho_b}{d\log M}, \tag{3.56}$$

give the fraction of the total matter density that belongs to the halos of mass M. When the variance is decreased by the box boundaries, this fraction is boosted at low masses due to a shift of halo formation to an earlier stage, where a larger fraction of matter is bound into low-mass objects.

Numerical Results and Comparisons

Following the above intuition, we employ the extended Press-Schechter formalism [72] to correct for the missing fluctuation variance on box scales. This formalism, while clearly inadequate at various levels in describing halo formation in realistic simulations [72, 123, 98], has nevertheless been very successful as a central engine in describing the statistics of cosmological structure formation. As shown by Mo & White [36] using N-body simulations, the biasing of halos in a spherical region with respect to the average mass overdensity in that region is very well described by the extended Press-Schechter approach. Barkana & Loeb [76] discussed the suppression of the halo mass function in terms of this bias, and suggested a prescription for adjusting large-volume mass function fits such as Warren or ST to small boxes. Here we do not follow this path but directly work with the numerical data by correcting the number of halos in each bin as in equation (3.50).

In the extended Press-Schechter scenario of halo formation, σ' on the right-hand side of equation (3.47) would be approximately connected with σ via $\sigma'^2 = \sigma^2 - \sigma^2_{R(\mathrm{box})}$ [72], where $\sigma^2_{R(\mathrm{box})}$ is the variance of fluctuations in spheres that contain the simulation volume. Since extended Press-Schechter theory is derived for spherical regions, while our simulation boxes are cubes, we define $R(\mathrm{box})$ as the radius of a sphere enclosing the same volume as in the simulations.

The action of this correction is shown in Figure 3.11. Finite-volume corrections are subdominant to statistical error at $z = 0$ and 5. At higher redshifts, the corrections produce results that are consistent across box sizes, i.e., that have no systematic shape changes or 'jumps' across box boundaries. Moreover, the action of the corrections is to bring the simulation results closer to a universal behavior.

For completeness, we mention two other approaches aimed at box-adjusting the mass function. The first [101, 124] simply replaces the original mass variance

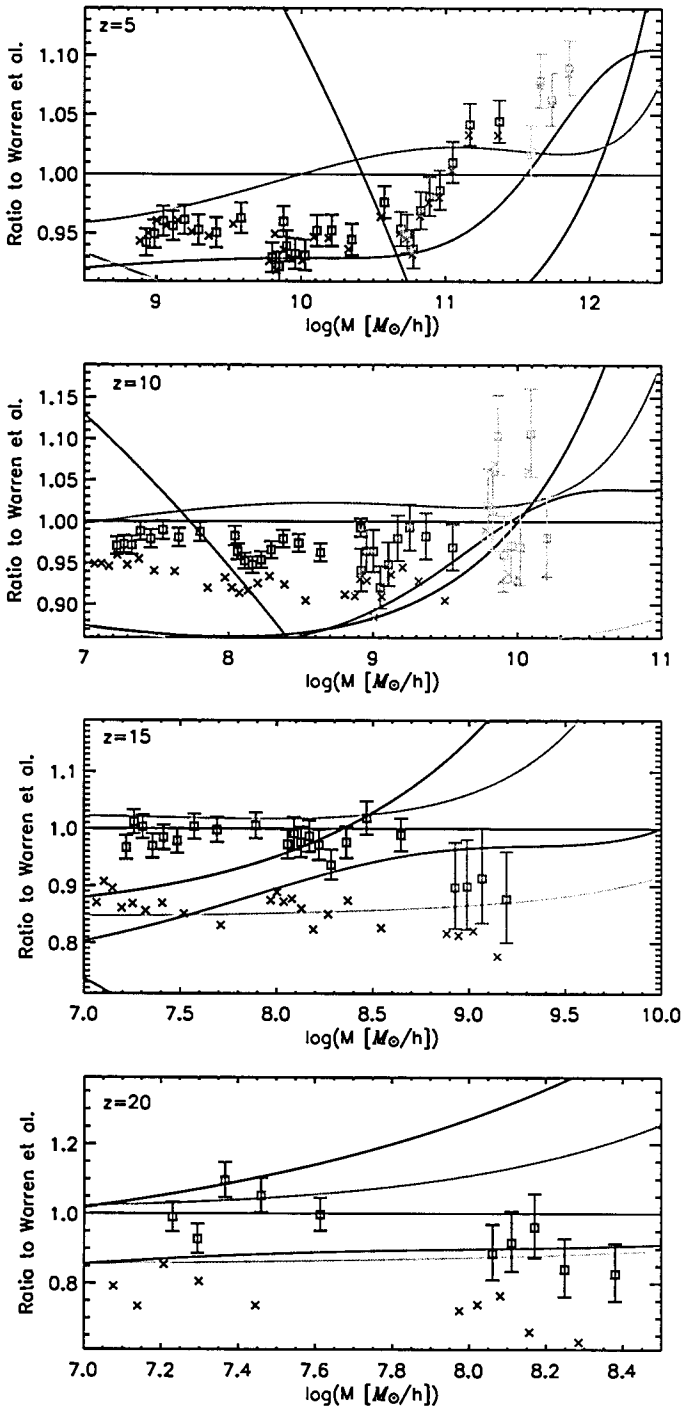

Figure 3.11: Mass function data corrected for finite box volume by the extended Press-Schechter prescription of §3.4.6 (squares). We show the results as a ratio with respect to the Warren fit and follow the conventions of Fig. 3.9. We also display the volume-uncorrected data (crosses). Note that the volume-corrected data join smoothly across the box-size boundaries. This box correction brings the results very close to universal behavior at high redshifts (see Fig.3.14).

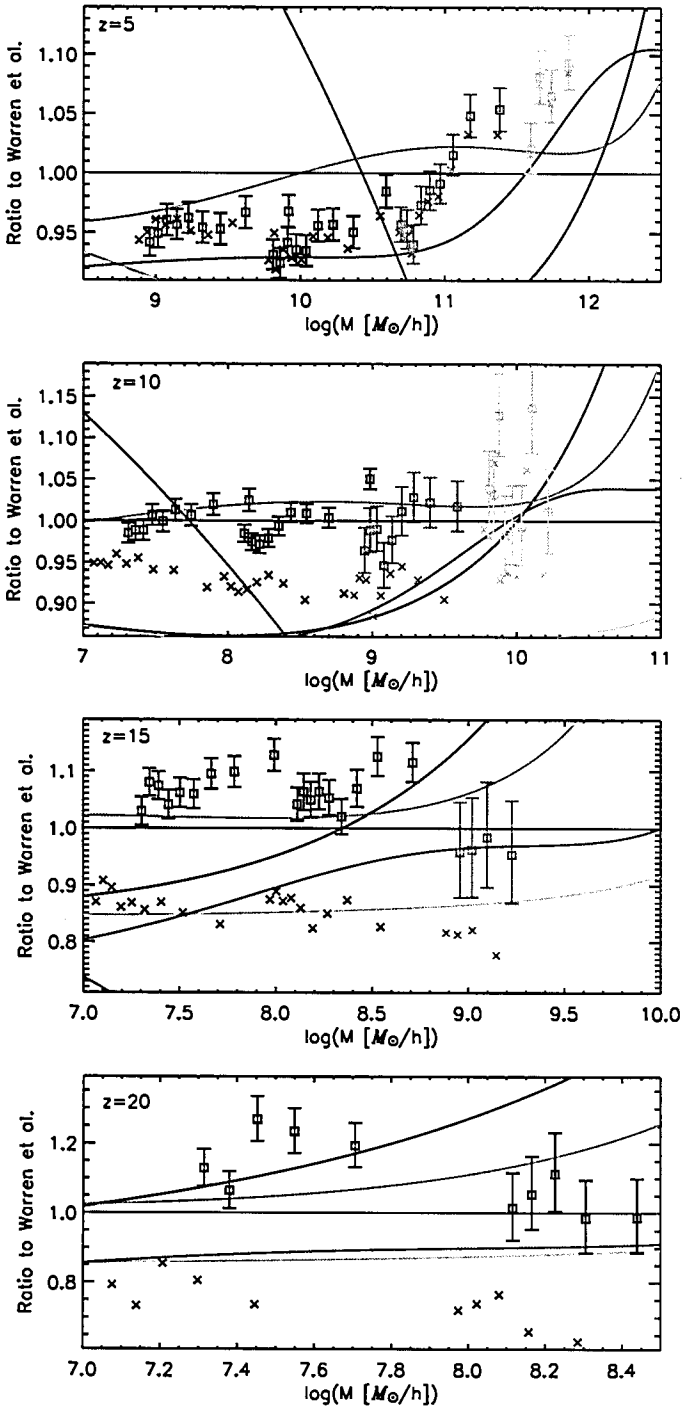

Figure 3.12: Mass function corrected for a finite box using the assumption of strict universality, as described in §3.4.6 (squares). Again, we show uncorrected data as well (crosses), and follow the conventions of Fig. 3.9. This correction produces a clear systematic shift in the results across box boundaries.

(eq. (3.3)) with

$$\sigma^2_{\rm box}(M,z) = \frac{d^2(z)}{2\pi^2}\int_{2\pi/L}^{\infty} k^2 P(k) W^2(k,M) dk \ , \qquad (3.57)$$

the lower cut-off arising from imposing periodic boundary conditions (L is the box-size). (For enhanced fidelity with simulations, the integral in eq. [3.57] goes to a sum over the simulation box modes.) This approach basically assumes that σ defined via an infrared cutoff is the appropriate replacement for the infinite-volume mass variance. Figure 3.12 shows the effect of this suggested correction: At $z = 0$ and 5 it is not noticeable, but at higher redshifts the correction is significant relative to the accuracy with which the binned mass function is determined. Furthermore, it exhibits systematic shape changes and offsets across boxes, in contrast to the results shown in Figure 3.11. For example, at $z = 10$ the corrected data at the crossover point between the 4 and $8\,h^{-1}$Mpc boxes ($\sim 10^8\,h^{-1}M_\odot$) have an offset of 5%. We conclude that this approach is disfavored by our simulation results.

An alternative strategy is to estimate the mass variance from each realization of $P(k)$ in the individual simulation boxes and to treat every box individually, as done in ref. [4]. This has in fact two purposes: to compensate for the realization-to-realization variation in density fluctuations (which could be a problem for small boxes) and also to compensate for an overall suppression in the mass function as discussed above. The disadvantage is that each of many realizations now has a different $\sigma(M)$ for a given value of M.

3.4.7 Mass Function Universality

Finally, we investigate the universality of the mass function found by Jenkins. Approximate universality is expected from the analytic arguments of PS and its extended, excursion-set formulation [72]. The universal behavior of halo formation persists even in the model of ellipsoidal collapse of ST, in which the predicted mass function is no longer of the PS form. On the other hand, the universality cannot be exact if the nonlinear interactions of different scales are fully accounted for: The nonlinear evolution that leads to the formation of halos of mass M must involve multiple degrees of freedom that are described by more parameters than the overall variance of the primordial overdensity smoothed by a top-hat filter $W(r,M)$. The universality is expected to be violated at sufficiently high resolution of the mass function even in the PS-type spherical collapse model: It is more reasonable to represent the probability of the collapse not by a fraction of particles at the center of spheres enclosing a mass M but by any fraction of particles belonging to such spheres [125]. The improved mass-function derived from this argument deviates somewhat from a universal form [126].

To investigate the extent our numerical simulations are consistent with uni-

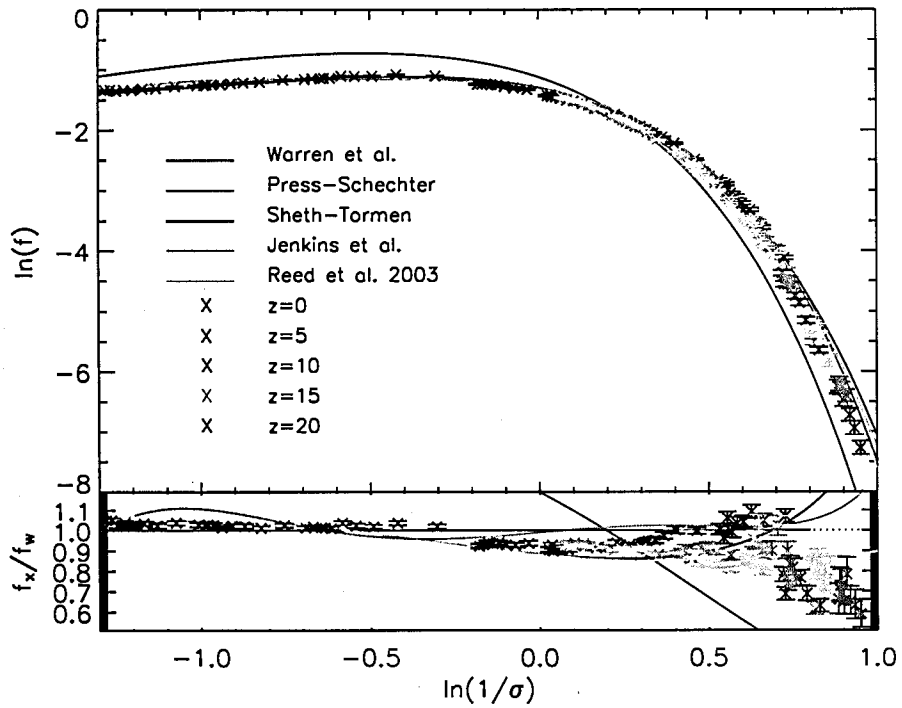

Figure 3.13: Scaled differential mass function from all simulations, prior to applying finite-volume corrections. Fits shown are Warren (red), PS (dark blue), ST (black), Jenkins (light blue), and Reed et al. (2003) (yellow). Dashed lines denote an extrapolation beyond the original fitting range. The bottom panel shows the ratio relative to the Warren fit. The failure of the different redshift results to lie on top of each other at small values of σ indicate a possible violation of universality.

versality, we combine our results for $f(\sigma, z)$ as a function of the variance σ^{-1} from the entire simulation set in one single curve at various redshifts. This curve is expected to be independent of redshift if universality holds. We display the results in Figure 3.13 for the raw data and in Figure 3.14 for the same data after applying the volume corrections discussed earlier.

In the raw data of Figure 3.13, the agreement with the various fits is quite tight (except for PS) until $\ln \sigma^{-1} > 0.3$. Beyond this point, the multiple-redshift simulation results do not lie on top of each other; in the absence of any possible systematic deviation, this would denote a failure of the universality of the FOF, $b = 0.2$ mass function at small σ. Note also that beyond this point the ST and Jenkins fits have a steeply rising asymptotic behavior (relative to the Warren fit). The Reed et al. (2003) fit, meant to be valid over the range $-1.7 \leq \ln \sigma^{-1} \leq 0.9$, is in better agreement with our results, to the extent that a single fit can be overlaid on the data.

The ostensible violation of universality seen above is small, however, and subject to a systematic correction due to the finite simulation volume(s). On applying the volume correction, we obtain the results shown in Figure 3.14, the

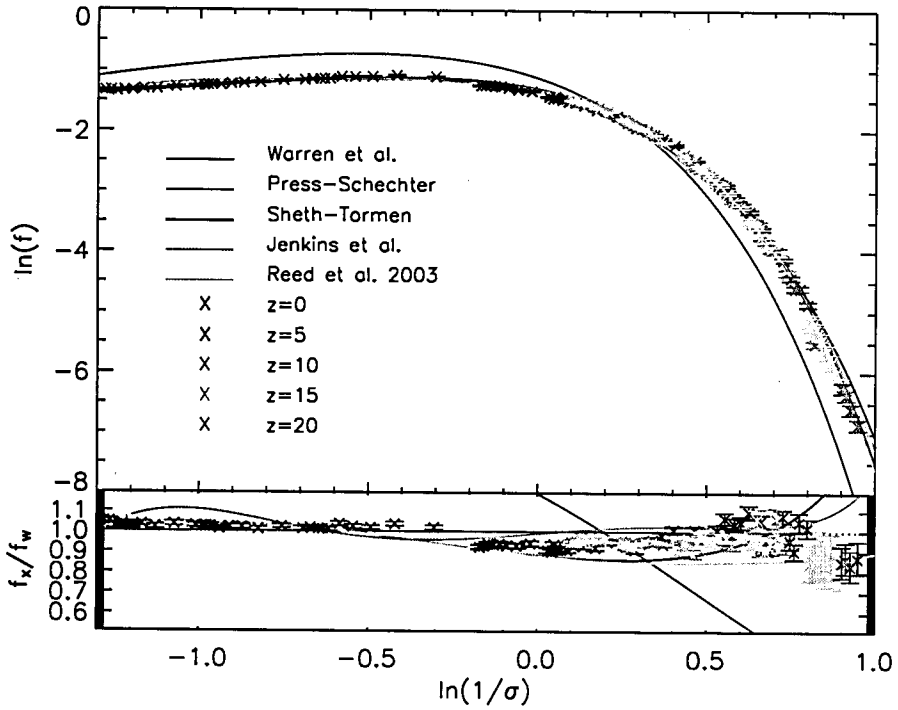

Figure 3.14: Volume-corrected scaled differential mass function following Fig.3.13. Note the significantly improved agreement with universal behavior (overlapping results beyond $\ln \sigma^{-1} \sim 0.3$).

key difference being that beyond $\ln \sigma^{-1} > 0.3$ the multiple-redshift simulation results now lie on top of each other and, within the statistical resolution of our simulations, are consistent with universal behavior. Specifically, we do not observe the sort of violation reported by Reed et al. [4] at high redshifts. This could be due to several factors. Here are used different finite-sampling FOF mass correction, as well as the finite-volume corrections. Moreover, the boxes we use at high redshifts are significantly larger. We note also that the difference between the Warren fit and the z-dependent Reed 2007 fit does not appear to be statistically very significant given either our or Reed et al. data.

3.5 Implications

We have investigated the halo mass function from N-body simulations over a large mass and redshift range. A suite of 60 overlapping-volume simulations with box sizes ranging from 4 to 256 h^{-1}Mpc allowed us to cover the halo mass range from 10^7 to $10^{13.5}\, h^{-1} M_\odot$ and an effective redshift range from $z = 0$ to 20.

In order to reconcile conflicting results for the mass function at high redshifts, as well as to investigate the reality of the breakdown of the universality of the mass function, we have studied various sources of error in N-body computations of the mass function. A set of error control criteria need to be satisfied in order

to obtain accurate mass functions. These simple criteria include an estimate for the necessary starting redshift, for the required mass and force resolution to resolve the halos of interest at a certain mass and redshift, and for the number of time steps.

The criteria for the initial redshift appear to be particularly restrictive. For small boxes, commonly used in the study of the formation of the first objects in the Universe, significantly higher initial redshifts are required than is the normal practice. A violation of this criterion leads to a strong suppression of the mass function, most severe at high redshifts. Recent results by other groups may be contaminated due to a violation of this requirement; a careful re-analysis of small-box simulations is apparently indicated.

The force resolution criterion is especially useful for grid codes, PM as well as adaptive mesh. The mass function can be obtained reliably from PM codes down to small-mass and up to high-mass halos provided the halos are adequately resolved. The resolution criterion is also very useful in setting refinement levels for adaptive mesh refinement (AMR) codes. As will be shown in chapter 6 the mass function from AMR codes is suppressed at the low-mass end if the base refinement level is too coarse.

The results for the required number of time steps to resolve the mass functions is somewhat surprising. The halo mass function appears to be very robust with respect to the number of time steps chosen to follow the evolution, even though the inner structure of the halos will certainly not be correct. Even a small number of time steps is sufficient to obtain a close-to-correct mass function at $z = 0$. This considerably simplifies the study of the mass function and its evolution.

Since finite-volume effects can also lead to a suppression of the mass function, we have tried to minimize the importance of these effects by avoiding too-small box sizes, by using overlapping boxes, and by restricting the mass range investigated in a given box size. In addition, we have found that a box-size correction motivated by the extended Press-Schechter formalism for the mass variance appears to give consistent results when applied to our multiple-box simulation ensembles.

We now briefly comment on results found previously by other groups. Jang-Condell & Hernquist [95] find good agreement with the PS fit at $z = 10$ for a mass range $4 \times 10^5 - 4 \times 10^8\, h^{-1} M_\odot$. The crossover of PS with the more accurate fits at $z = 10$ takes place in exactly this region (see Figs. 3.1 and 3.9). Therefore, all fits are very close, and the mass function from a single 1 Mpc box at a single redshift as shown in their paper cannot distinguish between them.

As mentioned earlier in §3.1.5, good agreement with the PS result has been reported at high redshifts (some results being even lower than PS) by several other groups [99, 100, 101, 103, 110]. The simulations of [103] and [110] were started at $z_{\rm in} \sim 50$, substantially below the starting redshift that would be suggested by our work. Furthermore, the very large number of particles in

the [110] simulation requires a high starting redshift (Fig. 3.4). Therefore, the depressed mass function results of these simulations are very consistent with a too-low initial redshift. ([110] have recently rerun their simulations with a much higher initial redshift – $z = 300$, and now find results consistent with results here.) The initial particle density of the [107] simulations is very close to that of our $16\,h^{-1}$Mpc box, in which case also a high redshift start is indicated (we used $z_{\rm in} = 200$). Finally, the initial redshift of the Yoshida et al. papers, $z_{\rm in} = 100$, for boxes of size $\sim 1\,h^{-1}$Mpc, also appears to be significantly on the low side.

We have compared our simulation results for the mass function with various fitting functions commonly used in the literature. The recently introduced ($z = 0$) fit of Warren leads to good agreement (at the 20% level with no volume correction, and at the 5% level with volume correction) at all masses and all redshifts we considered. Other modern fits, such as Reed et al. (2003, 2007), also lie within this range. These fits do not suffer from the overprediction of large halos at high redshifts observed for the ST fit. The PS fit performs poorly over almost all the considered mass and redshift ranges, at certain points falling below the simulations by as much as an order of magnitude.

The evolution of the mass function can be used to test the (approximate) universality of the FOF, $b = 0.2$ mass function. At low redshifts our data are in good agreement with those of [4] (at $z = 5$), finding a (possible) mild redshift dependence (at the 10% level). At higher redshifts, however, we find that volume corrections are important to the extent that little statistically significant evidence for breakdown of universality remains in our mass function data. A full theoretical understanding of this very interesting result remains to be elucidated.

We have made no attempt to provide a fitting function for our data due to several reasons. First, the current simulation state of the art has not reached the point that one can be confident of percent-level agreement between results from different simulations even in regimes that are not statistics-dominated [91]. Second, simulations have not sufficiently explored the extent to which universal forms for the mass function are indeed applicable as cosmological parameters are systematically varied. Third, absent even a compelling phenomenological motivation for the choice of fitting functions, there is an inherent arbitrariness in the entire procedure. Finally, it is not clear how to connect the FOF mass function to observations. In general, tying together mass-observable relations requires close coupling of simulations and observational strategies. In studies of cosmological parameter estimation, we support working directly with simulations rather than with derived quantities, which would add another layer of possible systematic error. Because observations already significantly constrain the parametric range, and are a smooth function of the parameters, this approach is quite viable in practice [127, 128].

4 Structure of Halos

Dark matter halos occupy a central place in the paradigm of structure formation: Gas condensation, resultant star formation, and eventual galaxy formation occur within halos. The distribution of halo masses – the halo mass function – and its time evolution, are sensitive probes of cosmology, particularly so at low redshifts, $z < 2$, and high masses. This last feature allows cluster observations to constrain the dark energy content, Ω_Λ, and the equation of state parameter, w [32]. In addition, phenomenological modeling of the dark matter in terms of the halo model (reviewed in ref. [129]) requires knowledge of the halo mass distribution and density profiles, as does the halo occupancy distribution (HOD) approach to modeling galaxy bias.

Because accurate theoretical results for the mass function (and other halo properties) do not exist, many numerical studies of halos and their properties, and of the mass function, have been carried out over widely separated mass and redshift ranges. Despite the intuitive simplicity and practical importance of the halo paradigm, halo definitions and characterizations have been somewhat *ad hoc*, mostly because of the lack of an adequate theoretical framework. For the purposes of this work, there are two crucial results that have been well-established by the numerical studies. The first is that spherically averaged halo profiles are well-described by the two-parameter NFW profile [130, 74] (this shape is consistent with observational studies of clusters), and second, that a simple 'universal' form for the FOF halo mass function (with link length, $b = 0.2$) holds for standard cold dark matter cosmologies [75]. A detailed understanding of both of these numerically established results remains elusive.

The universality of the FOF mass function has been recently verified to the level of $\sim 10\%$ accuracy for essentially all observationally relevant redshifts ($z < 10$) by several simulation efforts [73, 3, 4, 131], as described in chapter 3. The result is potentially very useful, because at this level of accuracy there is no longer any reason to simulate individual cosmologies, as the universal form already covers the parametric region of interest. There is one serious problem, however: the universal form of the mass function does not hold for the SO mass as defined and used by observers when determining the masses of galaxy groups and clusters [78, 77]. Unlike the SO criterion, the FOF method [79, 80] does not determine a (spherically-averaged) overdensity structure, but instead defines an object bound by some isodensity contour (Fig. 4.1). In principle, isodensity-based methods can be used in observations, but require significantly more work

than the SO approach.

At this point, one could ask the question whether the SO and FOF masses could be mapped to each other if more information regarding halo properties were avalilable. Alternatively, one could forsake universality and attack the SO mass function problem directly via simulations, [132, 133]. The work here will proceed along the first path and investigate whether an effective solution to the problem can be found [1]. We first show that even for perfect NFW halos, there is no simple direct mapping between FOF and SO masses, because of a significant dependence on the halo concentration. The mapping depends as well on the number of particles sampling a given halo, something that needs to be taken into account when interpreting results from simulations. However, we establish the useful result that for NFW halos sampled by a given number of particles, a two-parameter map utilizing concentration and particle number indeed connects the two masses with a small Gaussian scatter, quantified below in §4.2).

The key question is whether these relationships for idealized NFW halos survive when applied to the more realistic case of halos within cosmological N-body simulations. We find that this is indeed the case for halos that can be considered to be relatively isolated (a notion to be made more concrete in §4.2), and not possess significant substructure; i.e., approximately $80-85\%$ of all halos in the mass-range $10^{12.5}-10^{15.5}h^{-1}M_{\odot}$ explored by the simulations. This fraction of isolated halos is close to the conclusion of Evrard et al. [134] who anlayzed results from a large suite of simulations. For these halos, the two-parameter map derived above succeeds remarkably well in accurately converting the FOF mass function to the corresponding SO mass function, at the $\sim 5\%$ level – the current level of descriptive accuracy as limited by the robustness of halo definitions and numerical results from simulations (see chapter 3). We show that the concentration dependence of the FOF-SO mass relation is significant at the current levels of accuracy for the determination of halo masses. Conversion between FOF and SO masses will incur significant error if halo concentration is not considered. To transform between the FOF and the SO mass function, the scatter in concentration must also be considered. The work here has implications for observationally determined mass functions, and for HOD and other methods of deriving mock galaxy catalogs.

An additional point is that, in the N-body simulations, there not only exists a simple relationship between the halo concentration and the SO (or FOF) mass with a (relatively) large scatter, but that the scatter can be very well fit by a Gaussian distribution at a given mass. Using this simple concentration-mass relation and its Gaussian variance, one may go directly from the FOF mass function to the SO mass function or vice-versa. This procedure solves the mass function mapping problem for the subset of isolated halos, which comprise the bulk of the halo population. It does not, however, enable one to transform from

[1] For an earlier discussion, see [81], who noted that FOF and SO masses are correlated, but with a significant scatter.

the universal FOF mass function to a chosen SO mass function because of the $15-20\%$ fraction of FOF halos with irregular morphologies, most of which are 'bridged' halos (density peaks connected by high density filaments or ridges). A potential way around this difficulty is to treat explicitly the 'multiplicity' of apparently discrete SO halos within FOF halos in the transformation between FOF and SO mass functions.

Based on our runs for two different cosmologies, we have good evidence that the fraction of bridged halos rises as a function of mass, and that this fraction is also 'universal', i.e., more or less independent of the cosmology when written in units of M/M_*, where M_* is the characteristic halo mass-scale set by matching the RMS linear density fluctuation to the threshold density for collapse. We also find that the fraction of halos with major satellites as a function of the satellite mass fraction (with respect to the main halo) is cosmology dependent. This may pave the way for constraining cosmology from clusters of galaxies in a new way, essentially independent of the sampling volume, and therefore with enhanced immunity against selection effects. At the very least, using the major satellite halo fraction should provide a valuable cross-check for cosmological constraints derived from the mass function in the conventional manner.

4.1 Mass Definitions

The spherical overdensity and friends-of-friends methods are the two main approaches to defining halos and their associated masses in simulations. SO identifies halos by identifying spherical regions with prescribed spherical overdensities Δ:

$$M_\Delta = \frac{4\pi}{3} R_\Delta^3 \, \Delta \rho_c \ , \tag{4.1}$$

where ρ_c is the critical density [2]. An often-used value for the overdensity is $\Delta = 200$, roughly the theoretically predicted value given by the spherical collapse model, $18\pi^2$, for virialized halos in an Einstein-de Sitter universe (section §2.4.2). For the currently favored ΛCDM model ($\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$), spherical collapse actually predicts a smaller overdensity at virialization: $\Delta \approx 100$. X-ray observers, on the other hand, prefer higher density contrasts, $\Delta = 500$ or 1000, because strucutures on those scales are much brighter, and more relaxed compared to the outer regions.

The main drawback of the SO mass definition is that it is somewhat artificial, enforcing spherical symmetry on all objects, while in reality halos often have an irregular structure [81]. For some applications, such an approach may be well founded (for example X-ray cluster analysis for relaxed clusters), but may not be universally applicable. Furthermore, defining an SO mass can be ambiguous, since for two close density peaks, the corresponding SO spheres might overlap,

[2] Again, overdensities are in some works stated with respect to the background density: $\rho_b = \Omega_m \rho_c$, but we restrict ourselves to defining them with respect to ρ_c.

and one has to decide how to distribute particles between them (or assign them to both, breaking mass conservation).

The FOF algorithm, on the other hand, is not based on the notion of a certain overdensity structure, but defines instead an object bound by some isodensity contour. The mass of a halo is then simply the sum of all particles inside a given contour. By linking particles which are separated at most by the distance $ll = bn^{-1/3}$ (where n is the number density of particles in the simulation, and b is the linking length), the FOF method, in effect locates an isodensity surface of

$$\rho_{\rm iso} \approx kb^{-3}\rho_b \ , \tag{4.2}$$

where k is a constant of order 2 [135]. For $b = 0.2$, and the concordance ΛCDM cosmology, this leads to $\rho_{\rm iso} = 75\rho_c$. Given their percolation-centric nature FOF halos can have complicated shapes and topologies (Fig. 4.1).

4.2 Mass Mapping from Mock Halos

In order to address the relation of FOF and SO masses, we first turn to a controlled test using idealized 'mock' halos. These are taken to be spherical dark matter halos with the NFW density profile:

$$\rho(r) = \frac{\rho_s}{r/r_s\,(1+r/r_s)^2} \ , \tag{4.3}$$

where ρ_s and r_s are the core density and scale radius respectively. Instead of ρ_s and r_s, it is often convenient to use physically more transparent quantities: the SO mass M_Δ and the concentration $c = r_s/R_\Delta$:

$$\rho_s = \frac{\Delta\ \rho_c\ c^3}{3\ [\ln(1+c) - c/(1+c)]} \ ; \tag{4.4}$$

$$r_s = \frac{1}{c}\left[\frac{3\ M_\Delta}{4\ \pi\ \Delta\ \rho_c}\right]^{1/3} \ . \tag{4.5}$$

The cumulative mass within a radius r can be calculated as:

$$\begin{aligned} M(r) &= \int_0^r 4\pi r^2 \frac{\rho_s}{r/r_s\,(1+r/r_s)^2} dr \\ &= 4\pi\rho_s r_s^3\,[\ln(1+r/r_s) - (r/r_s)/(1+r/r_s)] \ . \end{aligned} \tag{4.6}$$

While it is still unclear whether the very inner parts of the halos ($\sim 1\%$ of R_{200}) have density profiles steeper than NFW [136, 137, 138, 139, 140], the inner asymptotic slope is not of concern here, and does not affect results presented here.

The mock NFW halos are generated in the following way: first we fix the SO mass ($M_\Delta \equiv M_{200}$) of a halo and choose the number of particles which will reside in it (N_{200}). We then populate the halo with particles according to

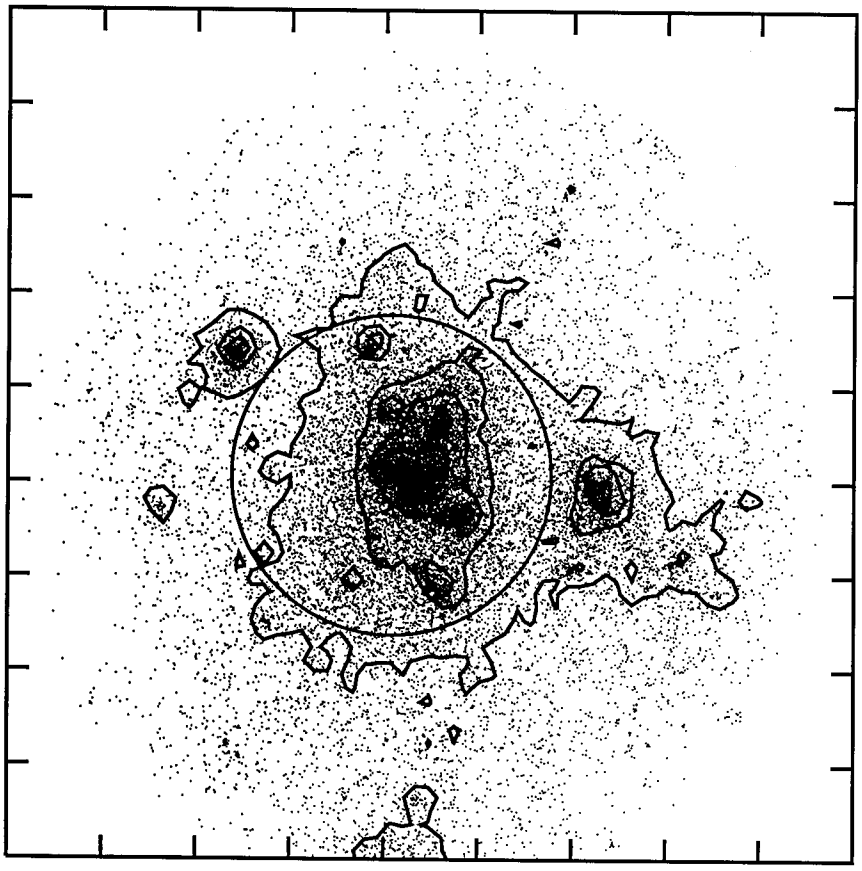

Figure 4.1: Different halo definitions for the same particle distribution in a simulation. The green points show all particles in a sphere centered around the minimum potential FOF particle and with radius 1.1 times the distance to the farthest FOF member ($b = 0.2$). The black contours are for the two dimensional density field projected onto the $z-$direction as calculated from all the particles. The blue particles show the actual FOF halo members. The red circle shows the SO halo centered around the same point as the FOF halo. The box spans approximately $3.15h^{-1}$Mpc in x- and y-direction, R_{200} is approximately $0.6h^{-1}$Mpc. The FOF mass of the halo is $6.70\times10^{13}h^{-1}\mathrm{M}_{\odot}$, the SO mass of the main halo is $4.91\times10^{13}h^{-1}\mathrm{M}_{\odot}$ and the SO mass of the major subclump on the right (which belongs to the FOF halo) is $8.50\times10^{12}h^{-1}\mathrm{M}_{\odot}$. The small subclump on the left (which was neither included in the FOF halo nor in the SO halo) is $2.97\times10^{12}h^{-1}\mathrm{M}_{\odot}$. This plot demonstrates how closely the FOF halo boundary tracks an isodensity contour.

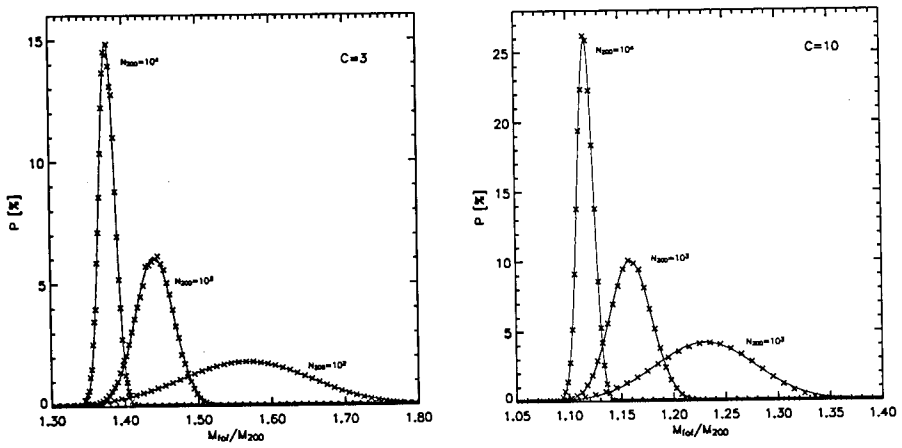

Figure 4.2: Distribution of $b = 0.2$ FOF masses for NFW halos with concentrations $c = 3$ (left panel), and $c = 10$ (right panel), sampled with different particle numbers: 100 (blue), 1000 (green), 10000 (red). The number of Monte Carlo samples are 10^6, 10^5, and 10^4 for $N_{200} = 100$, 1000, and 10000, respectively. The solid curves are Gaussian fits. Note that the two panels have different units along both axes.

the NFW distribution such that we enforce the desired mass to be M_{200} within the radius R_{200}. We then extend the NFW distribution further out – adding particles to a 'halo tail'. The choice of $\Delta = 200$ can easily be changed to any other desired value such as $\Delta = 500$ or 1000 as more appropriate for cluster studies. In any case, for a given NFW profile choice, all overdensity masses are immediately fixed, so there is no lack of generality in our specific choice which corresponds to an approximate notion of the 'virial mass' [130, 74].

Having fixed M_{200} for all the mock halos, we determine the FOF mass for every halo. Because the particles are randomly sampled inside a halo (following the NFW density profile), one cannot expect that for every realization of a mock halo, the FOF finder will return exactly the same mass. Given a large number of mock halos with the same density profile and statistical independence of the realizations, the central limit theorem predicts a Gaussian distribution for the FOF masses. Indeed, just as expected, a normal distribution gives an excellent description for M_{FOF}/M_{200}. Thus, one can not only determine to what SO mass a certain M_{FOF} corresponds (on average), but can also quantify the systematic deviation of an FOF halo finder through a standard deviation (Figs. 4.2). The Gaussian spread of FOF masses is centered around a mean value that shifts systematically with the number of sampling particles, N, as empirically noted by Warren et al. [8] (Fig. 4.3).

Besides this N-dependence, we also wish to examine how M_{FOF}/M_{Δ} depends on the underlying profile. We have found that this dependence leads to another source of bias for FOF masses relative to SO masses. In Fig. 4.3, we show average values of M_{FOF} for a range of particle numbers and concentrations. It is clear that one cannot accurately match a given M_{200} to a corresponding

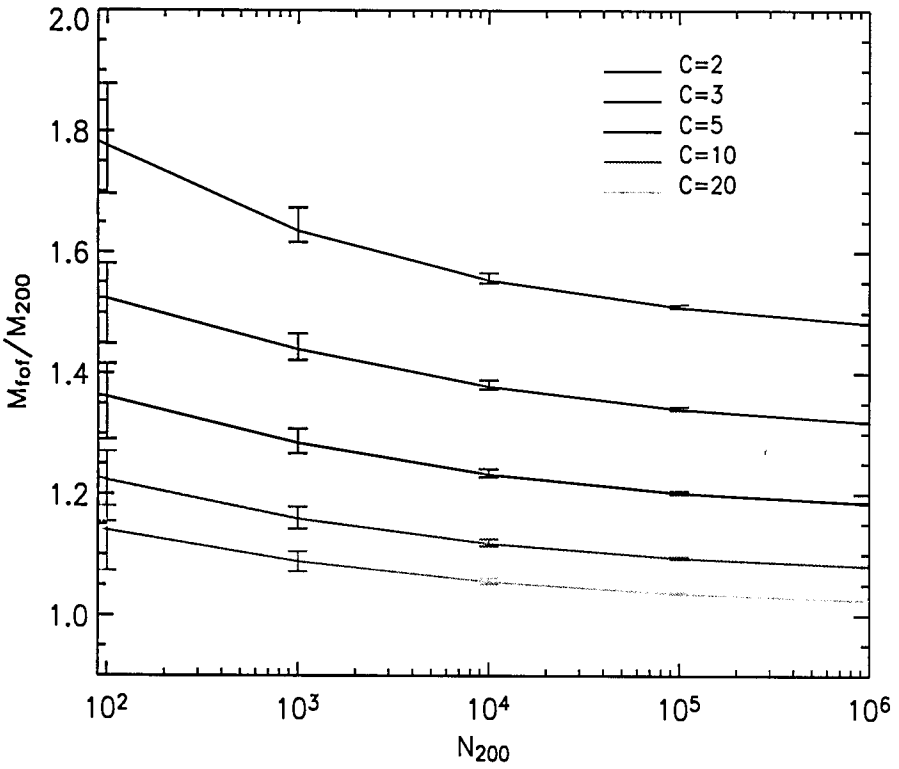

Figure 4.3: Ratio of the ($b = 0.2$) FOF mass to M_{200} for NFW mock halos with different concentrations and particle number, N, but the same value of M_{200}. Low concentration halos have up to a factor of two higher FOF mass than M_{200}. For high concentration halos, the ratio of the two mass definitions is closer to unity, the FOF mass being always higher.

M_{FOF} without the concentration being specified. Concentration variation from $c \sim 20$ (typical for galaxies) to $c \sim 5$ (typical for clusters) [141, 142] corresponds to systematic FOF mass shifts of $\sim 30\%$, much larger than can be tolerated by the accuracy to which the FOF mass function can currently be determined numerically ($\sim 5\%$). For any given N_{200}, this concentration dependence follows the functional form:

$$\frac{M_{FOF}}{M_{200}} = \frac{a_1}{c^2} + \frac{a_2}{c} + a_3 \ , \tag{4.7}$$

where the coefficients a_1, a_2, a_3, depend on N_{200} only (Table 4.1).

Well-sampled halos, with $N > 1000$, are characterized by a small variance in the M_{FOF}/M_{200} ratio, with a maximum value of $\sigma \sim 0.02 - 0.03$, depending on the concentration. With such a low intrinsic scatter in the mass relationship for a given concentration, the logical next step is to see whether the mean $M_{FOF}(M_{200}, c)$ relationship obtained from the mock NFW halos actually applies to the real halos in N-body simulations. Here, it should be noted that actual simulated halos are not expected to be spherical due to the episodic and anisotropic nature of mass accretion, and in fact are much better described as ellipsoids [143, 144]. Nevertheless, as we are interested in an averaged quan-

Table 4.1: Best Fit Coefficients. Best fit coefficients for different N_{200}, as obtained from the mock halo analysis. For all values of N_{200}, the functional form of the fit is given by Eqn. (4.7).

Coeff.	N_{200}						
	100	600	10^3	6×10^3	10^4	10^5	10^6
a_1	-0.3887	-0.3063	-0.2790	-0.2210	-0.1970	-0.1642	-0.1374
a_2	1.6195	1.4130	1.3669	1.2459	1.2157	1.1392	1.0900
a_3	1.0715	1.0313	1.0226	1.0008	0.9960	0.9800	0.9714

tity, the halo mass, an approach based on idealized halos may well provide an adequate description.

4.3 Mass Mapping in N-Body Simulations

In order to investigate the validity of the mock halo mass relationships, we use results from four cosmological simulations for two flat ΛCDM cosmologies, each simulated with 174 and 512 h^{-1}Mpc boxes. The pre-WMAP, high-σ_8 cosmology has the following parameters: matter density, $\Omega_m = 0.3$; dark energy density, $\Omega_\Lambda = 0.7$; fluctuation amplitude, $\sigma_8 = 1.0$; Hubble constant $h = 0.7$ (in units of 100 km s^{-1} Mpc^{-1}); primordial spectral index, $n_s = 1$; and the Bardeen et al. [145] transfer function with $\gamma = \Omega_m h$. For the WMAP 3 compatible cosmology runs, the parameters are: $\Omega_m = 0.26, \Omega_\Lambda = 0.74, \sigma_8 = 0.75, h = 0.71$, $n_s = 0.938$, and a transfer function generated using CMBFAST [113]. We use the parallel gravity solver GADGET2 [96] to follow the evolution of 512^3 dark matter particles starting from a redshift $z = 99$, high enough to satisfy the initial redshift requirements given in the chapter 3. The particle masses are 3.3×10^9 and $8.3 \times 10^{10} h^{-1} M_\odot$ for the high-σ_8 run, and 2.8×10^9 and $7.2 \times 10^{10} h^{-1} M_\odot$ for the WMAP 3 cosmology. These masses are small enough to comfortably resolve groups and clusters to the level required for this study [146, 140, 147]. The FOF mass functions from these simulations are in very close agreement with the results of previous chapter, well within a few percent. By using cosmologies with normalizations that bracket the currently favored cosmologies, we are able to show that our results are applicable to any likely cosmology, once (cosmology dependent) halo concentrations are specified.

To carry out a realistic test of the mass relationships, we adopt the following procedure: (i) First run an FOF halo finder on the final particle distribution, and select all halos with $N > 1000$. (ii) Define halo centers by identifying the local gravitational potential minima. This corresponds closely to the most bound particle, as well as the density peak of the halo in most cases [78]. (iii) Construct individual SO profiles around these minima, thereby determining M_{200}. The halo density is computed in 32 logarithmically equidistant bins, and we fit the NFW profile treating both r_s and ρ_s as free parameters. As

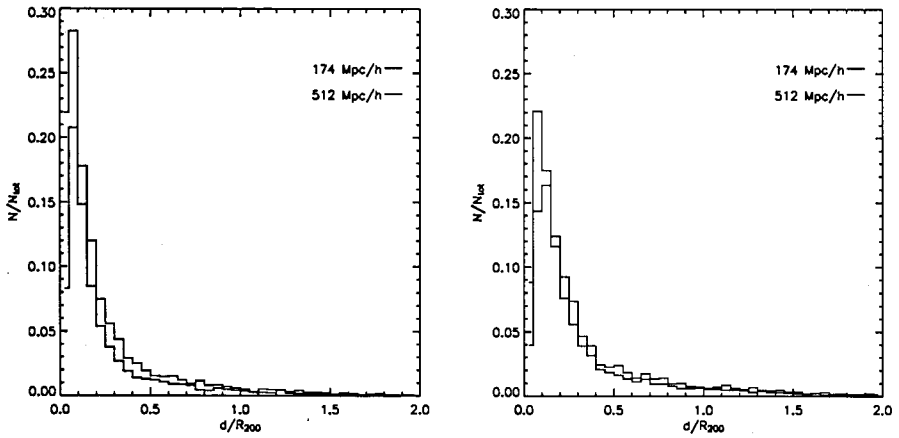

Figure 4.4: Distribution of distances between FOF center of mass, and potential minimum for 512 h^{-1}Mpc box (red) and 174 h^{-1}Mpc box (blue), scaled by R_{200}. Left panel is for High σ_8 cosmology, right panel for WMAP 3.

a consistency check, we use an alternative approach, where M_{200} is measured directly from the mass within a sphere, and NFW is treated as a one-parameter function (by fixing ρ_s such that the enclosed overdensity is $200\rho_c$). No significant differences were found between the two approaches.

The $N > 1000$ halo particle cut keeps the variance in the mass ratios small (Cf. Figs. 4.2-4.3) and also allows stable calculations of the individual halo concentrations. For each FOF halo we find its center of mass from all the particles linked together by the halo finder. On occasion, the FOF finder connects apparently distinct halos (bridging); these halos may well be in some stage of merging. Since it makes little sense to define an SO profile and an associated concentration for very close halos and those undergoing major mergers, we use the distance between the center of mass and the potential minima to exclude such halos. In figure 4.4 we show the distribution of that distance (d) for all halos with $N > 1000$ from both of the simulation boxes. While most of the halos appear to be isolated objects where the difference between the two center definitions is due to substructure, there are severe outliers, and even objects where the FOF center of mass is more than R_{200} away from the potential minimum!

To proceed further, we first set aside all halos with $d/R_{200} > 0.4$. Although this cut is somewhat arbitrary, the results are relatively insensitive to the particular choice, as discussed below. Furthermore, the mock halo analysis on regular NFW halos shows that, even at low concentrations, one expects approximately $M_{FOF}/M_{200} \sim 1.5$ (Cf. Fig. 4.3). Larger values therefore are a signal of a potential merger, as was verified directly by confirming with the simulation results. In figure 4.5 where we plot both 'isolated' (blue) and 'bridged' (red) halos, the strong correlation between our cut, based on the difference between halo mass and potential centers, and the high values of M_{FOF}/M_{200} (with respect to the mock halo expectation) can be easliy verified.

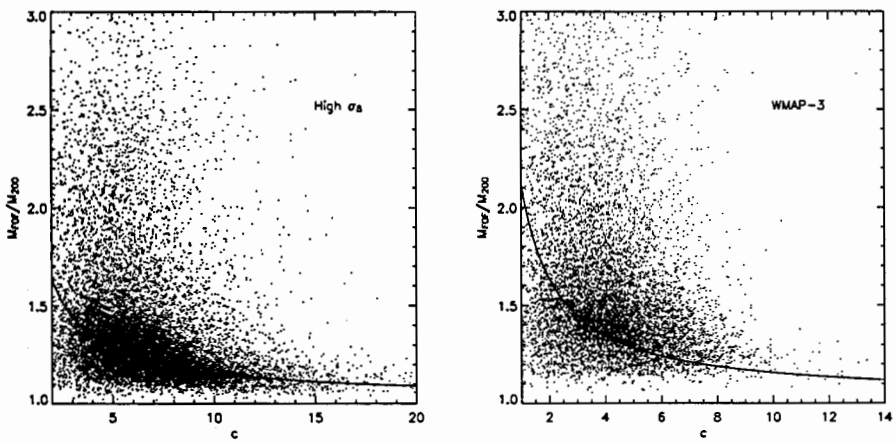

Figure 4.5: Left panel: Scatterplot of the ratio of FOF and SO(200) masses from the High σ_8 simulations as a function of the measured concentration for (i) halos passing the criterion $d/R_{200} < 0.4$ (blue), where d is the distance between the center of mass and the potential minima (see discussion in the text), and (ii) halos not passing this criterion (red). The solid line shows the mock halo prediction for halos with particle number, $N_{200} = 10^3$, which dominate the sample. Right panel: The same for WMAP 3 cosmology.

Finally, we compare our cutoff with an SO analysis of FOF halos: for each halo we find the potential minimum particle and R_{200} around it, and than move to the next particle in the potential hierarchy which resides outside R_{200} (if inside, we define it as a piece of substructure rather than a 'satellite halo' bridged by the FOF procedure), find R_{200} and M_{200} for the second halo, and iterate this procedure until all FOF particles are exhausted. When separate SO halos overlap we assign particles in the overlapping region to all SO halos, keeping the overdensity idea straightforward, but breaking mass conservation. Of course, if one goes down to a few particles, then virtually all FOF halos will be resolved into multiple SO objects. But if the threshold of the satellite mass is raised to 20% of the main halo mass, most of the FOF halos appear as a single SO halo. The two methods: $d/R_{200} > 0.4$, and $M_{\rm satellite}/M_{\rm main} > 0.2$ correlate extremely well, agreeing in 85-90% of all cases (the agreement is worse for larger masses, and better for smaller halo masses). This gives us additional confidence that our cutoff criterion separates isolated from bridged halos.

For the both halo samples, we now apply the $M_{FOF}(M_{200}, c)$ relationship determined by the mock halo results of Fig. 4.3, as encapsulated in the fits specified in Table 4.1. From the figure 4.6 we see that by using our recipe we are able to sucessfully estimate M_{200} knowing $M_{F\delta F}$ for isolated halos. On the other hand, for excluded (bridged) halo sample, analysis done on NFW halos is clearly not applicable. Note also that these halos show significantly bigger scatter in mass estimate. We will return to an analysis of the excluded halos in Section 4.4, and for the moment we shell focus on isolated halos.

The halo exclusion cut eliminates only about 15-20% of all halos, so while

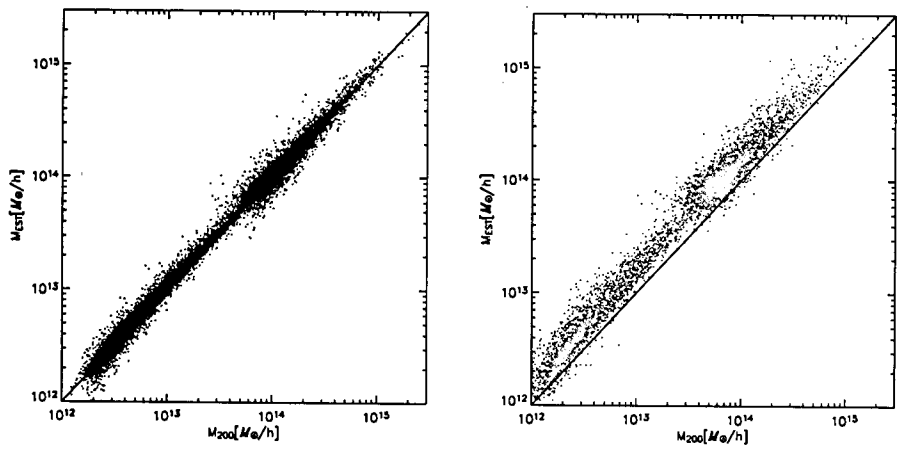

Figure 4.6: Scatterplot showing mass estimate (M_{EST}) for M_{200} using our mock halo recipe, and M_{200} as measured in simulations. Left panel is showing isolated halo sample, passing $d/R_{200} < 0.4$ criterion, while right panel is for halos excluded by the same criterion.

not statistically very significant, it is certainly not negligible. One of the main tests of halo mass mapping is the mass function; the results are shown in the figure 4.7, where the measured mass functions are displayed in terms of a ratio to a Warren fitting form for the FOF mass function [3] The undernormalization of the FOF mass function relative to the fit is simply due to the exclusion procedure described above. Note that the FOF and SO mass functions, as numerically determined, differ by as much as $20-40\%$ depending on the mass bin. However, application of the mock halo mass relationship to every individual FOF halo *correctly reproduces the SO mass function at the* 5% *level*, the current (numerical) limiting accuracy of mass function determination. The success of this simple mapping idea is a testimony to the accuracy of the NFW description for (spherically averaged) realistic halos in simulations, and consistent with the overall conclusion [134] that the majority of cluster-scale halos are structurally regular.

Using the expression for the cumulative NFW mass [Eqn. (4.6)], we can find the mass for any desired overdensity Δ in terms of M_{200}; defining $M_c = M_\Delta/M_{200}$, we have:

$$M_c = A(c)\left[\ln\left(1+\sqrt[3]{\frac{200}{\Delta}}M_c c\right) - \frac{\sqrt[3]{\frac{200}{\Delta}}M_c c}{1+\sqrt[3]{\frac{200}{\Delta}}M_c c}\right] , \tag{4.8}$$

where $A(c)$ is a prefactor which depends on c only:

$$A(c) = \frac{1}{\ln(1+c) - c/(1+c)} . \tag{4.9}$$

[3]This ratio is taken only for ease of interpretation, as any other mass function fit would have done just as well.

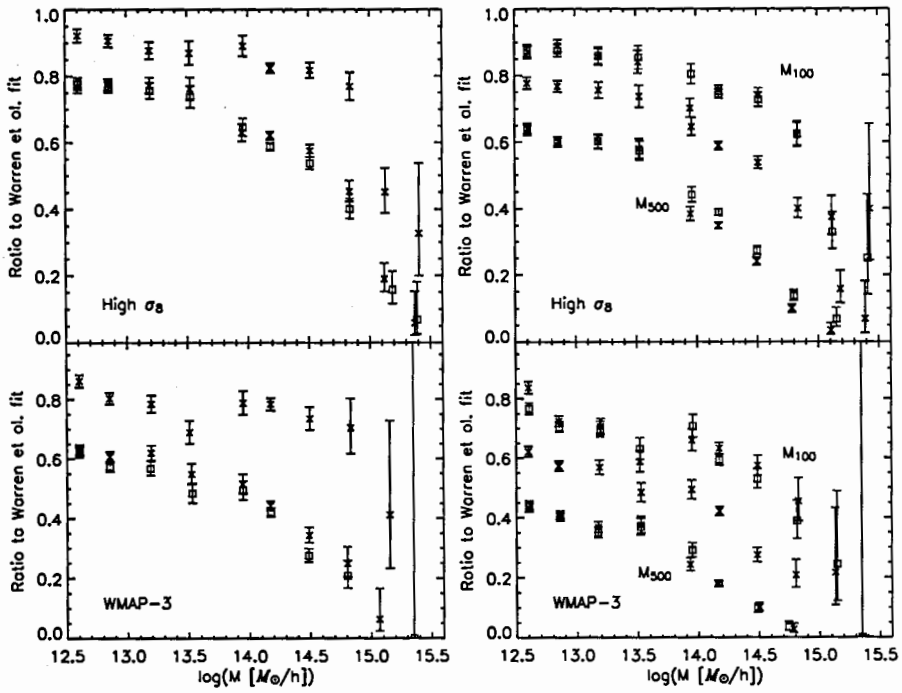

Figure 4.7: Right: mapping FOF to SO mass function. Measured mass functions normalized to the Warren et al. (2006) fit as an (arbitrary) reference, for High σ_8 (upper panel) and WMAP-3 cosmology (lower panel). Black: FOF halo masses with $b = 0.2$ and bridged halos removed as shown in Fig. 4.5. Red: M_{200} masses measured from the simulation for the same set of halos. Blue: The mass function for M_{200} halos using the idealized mock halo prediction (Fig. 4.3 and Table. 4.1), the measured FOF masses for each halo as mapped to the predicted SO mass. The agreement between measured (red) and predicted (blue) mass functions is excellent, better than 5%. Left: moving from one SO definition to another. Black: M_{200} masses measured from the simulation. Red: M_{100} and M_{500} masses measured from the simulation using the same halo centers. Blue: Idealized NFW predictions for M_{100} and M_{500} using the measured M_{200} mass for each halo. Measured and predicted quantities (red vs. blue) are again in very good agreement.

Employing this approach one can easily move from one SO mass function to another, and in Fig. 4.7 we show that this mass transformation gives accurate results for halos in simulations. Furthermore, this shows that if one is interested in any overdensity other than 200 (as considered in our mock halo analysis), our best fit for M_{FOF}/M_{200} [Eqn. (4.7) and Table 4.1] can simply be rescaled for any M_Δ using Eqn. (4.8).

The results shown in the figure 4.7 depend only weakly on the cut imposed by a particular value of d/R_{200}. Choosing a value below $d/R_{200} = 0.4$ such as 0.3 is more conservative; one loses more halos (another 5%), but the mass function mapping results remain excellent. Increasing the cut threshold to 0.5 adds 5% more halos while the mapping accuracy remains more or less the same. Beyond this point the results slowly degrade, as is to be expected.

With this important result at the level of individual halos in hand, the global mass function can be realized without knowing individual halo concentrations, and independent of cosmology, provided one has a form for the (mean) concentration-mass relation for SO (or FOF) halos as well as the PDF for the scatter in this relation. The latter cannnot be ignored since the scatter in the concentration-mass relation is known to be significant [148, 141, 142, 149, 147]. In the mass regime typical for clusters, i.e., halo masses above $\sim 3 \times 10^{14}$ $h^{-1}M_{\odot}$, the variation in concentration with mass is in fact much smaller than the concentration scatter for halos of similar mass. While our simulations are not well posed to determine concentration-mass realtion (due to the modest particle loading), we have carried out a basic analysis to establish the cosmology dependence of the concentration-mass relation, $c(M_{200})$, and its associated scatter, $\sigma_c(M_{200})$ [or $\sigma_c(M_{FOF})$]. The full analysis is left to a future work, but our preliminary research shows strong evidence that $c(M)$ relation and its scatter provides all the required information for mapping mass functions. The scatter itself is very well described by a Gaussian PDF at each mass bin (for both SO and FOF masses) and has little variation over the limited mass range relevant for clusters.

4.4 The Bridged Halos

We now turn to understanding the FOF halos that cannot be simply mapped as individual NFW profiles. Broadly speaking, we find that these halos are of two types: (i) Halos with density bridges across major substructures, and (ii) halos with complex substructure ('unrelaxed'). Halos of the first type are the ones largely excluded by our halo mass and potential centers-based cut and correspond mostly to the high mass-ratio region in figure 4.5. While our cut is very efficient in terms of identifying bridged halos, there is a very small contamination fraction due to chance symmetric bridging which does not lead to significant differences between the mass and potential minima. The second type of halos corresponds largely to the low concentration/low mass ratio region. Representative halo types are shown in Fig. 4.8: typical isolated halo (upper panel), bridged halo (middle panel), and complex substructure (lower panel).

It is clear that the idea of a single concentration or a simple mass ratio M_{FOF}/M_{SO} makes little sense for either the bridged halos or the unrelaxed halos. For the unrelaxed halos, absent a sub-halo analysis, it is not even clear what an appropriate M_{SO} might be. Nevertheless, our exclusion was designed mostly to eliminate the bridged halos; our results show that the unrelaxed population is apparently subdominant at least in terms of biasing the mass function results. Even so, it is clear that the existence of these types of substructured halos has ramifications for the simple HOD program, although the quantitative impact needs to be studied.

The halos that are bridged by the FOF procedure are typically close neigh-

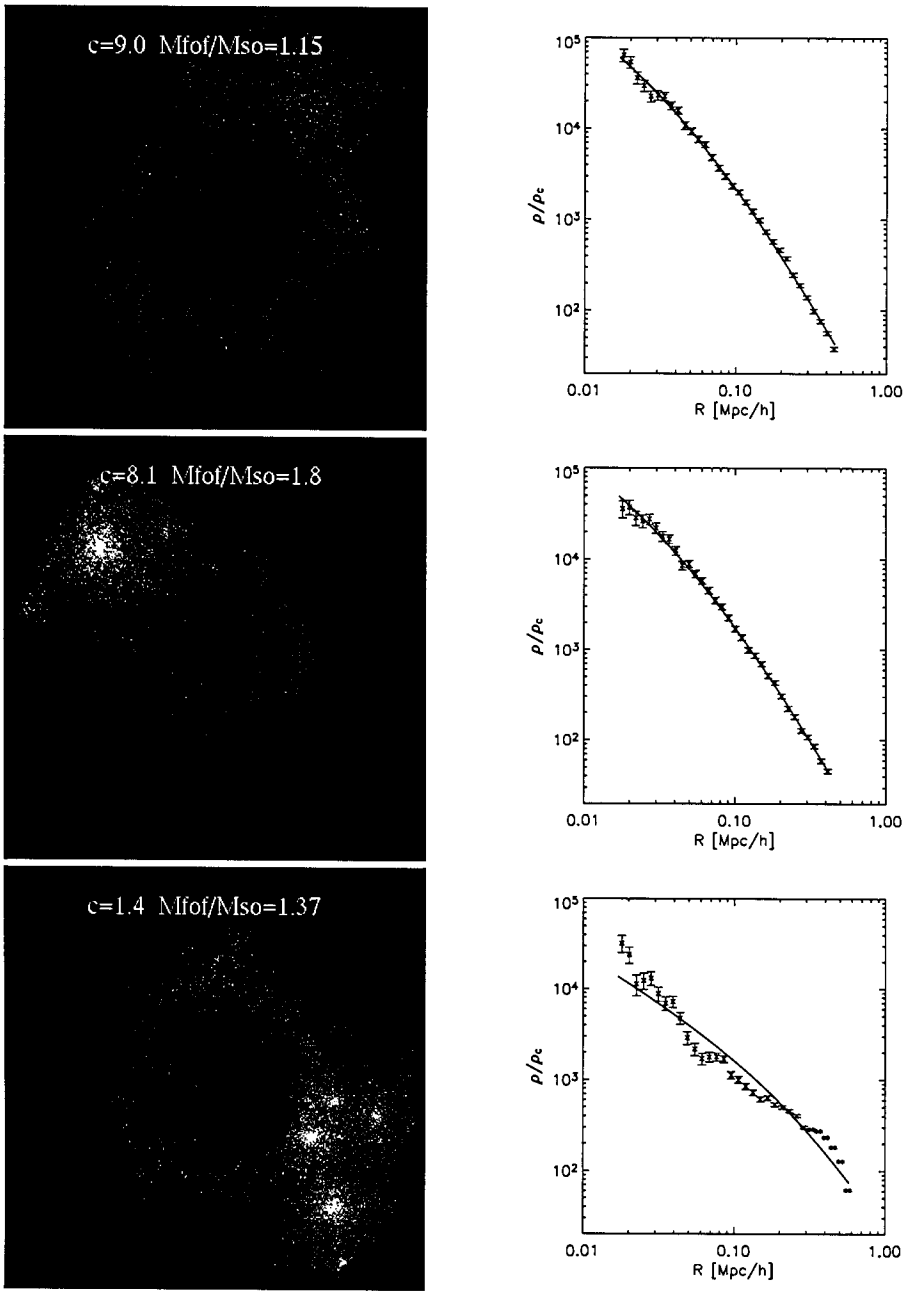

Figure 4.8: Top panel: A typical isolated FOF halo (FOF-linked particles shown as white dots) with NFW concentration, $c = 9.0$, and $M_{FOF}/M_{200} = 1.15$ (profile fit to the right). Green dots are particles within R_{200} of the corresponding SO halo. Middle panel: An example of a bridged halo. The SO halo found at the FOF center has concentration $c = 8.1$ (the NFW profile fit is a good fit), however the mass ratio $M_{FOF}/M_{200} = 1.8$ is high due to the bridged minor halo in the left upper corner. Bottom panel: A halo with major substucture, for which the NFW profile is not a good fit.

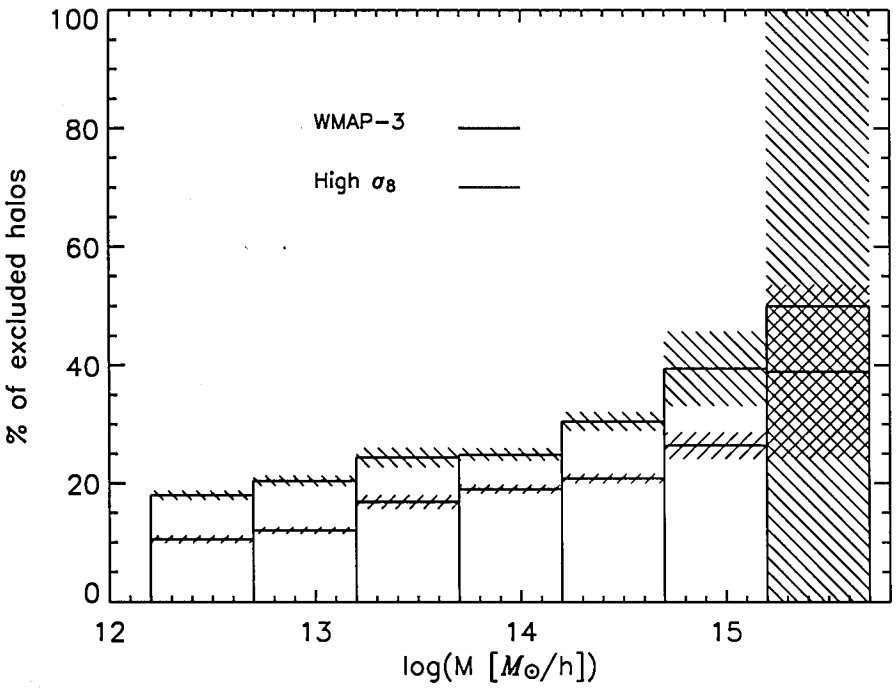

Figure 4.9: Distribution of bridged halos as a function of mass for the high σ_8 and *WMAP* 3 cosmologies. In both cosmologies, the relative fraction of such halos tends to increase with increasing mass. The shaded regions are Poisson error bars.

bors, the majority being partners in the hierarchical process of structure formation via halo merging [150]. Some of these close neighbors might be 'backsplash halos' that have previously been within R_{200} [151, 152]. In both the high-σ_8 and WMAP 3 cosmologies, we find that the fraction of bridged halos has a tendency to increase with increase in mass. This is as expected from the hierarchical merging picture since very massive halos are still forming at the current epoch. This effect is clearly shown in Fig. 4.9. We have checked that the two different-sized boxes (for each cosmology) agree well in the region of overlap, supporting the argument that numerical effects (finite mass and force resolution) are negligible for this consideration. (For the two box sizes, the mass resolution differs by a factor of approximately 25, and the force resolution by a factor of 3.)

The overall effect depends on cosmology: the results from the WMAP 3 simulation are clearly separated from the high σ_8 cosmology (Fig. 4.9). Since the structure grows differently in the two different cosmologies (due to different σ_8 and Ω_m), we can try to parametrize our exclusion as a function of M/M_*, where M_* is the characteristic collapse mass at the current epoch, defined through:

$$\sigma[M_*(z)] = 1.686\ , \tag{4.10}$$

where σ is the variance of the linear density fluctuation field $P(k)$, as defined in equation (3.3).

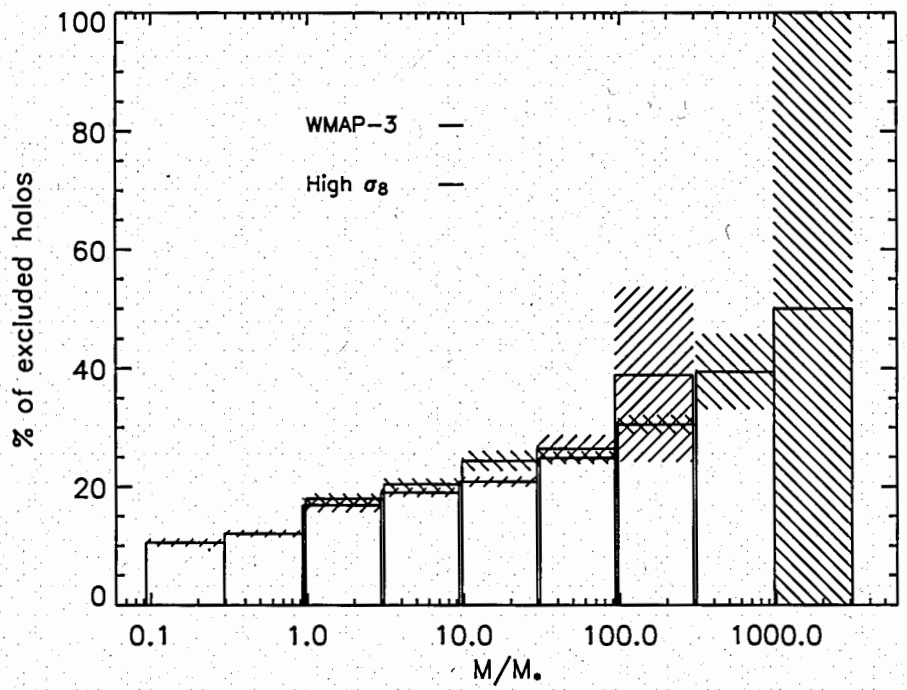

Figure 4.10: Possible universality of the bridged halo fraction: The same data as in Fig. 4.9, but with the mass now scaled by characteristic collapse mass – M_*.

As shown in Fig. 4.10, with the mass rescaled in terms of M_*, the fraction of bridged halos agrees for the two cosmologies and may very well be universal. This intriguing fact indicates, first, that our method of excising bridged halos (the principle, not necessarily the specific choice of $d/R_{200} > 0.4$) is physically well-motivated. Second, if the universality is borne out, the bridged halo fraction can be combined with the cosmology independent mock halo analysis, to yield a method for translating the universal FOF mass function to any desired SO mass function. Moreover, these results suggest that the bridged halo fraction can also provide a separate probe of cosmology, being particularly sensitive to the same parameters as the mass function itself (Fig. 4.9).

An additional way to probe the growth of structure in the Universe using clusters, aside from the mass function, would be to measure the fraction of isolated clusters versus those that have (major) satellites. In our simulations, we measure the fraction of multiple SO dark matter halos in the mass range of interest for clusters: $M_{200} \geq 10^{14} M_{\odot}/h$ (see also Ref. [134]). If we plot this fraction as a function of f, where f is defined through $M_{\rm satellite} \geq f M_{\rm main}$ we find again that the two cosmologies considered are clearly separated, as shown in Fig. 4.11. The advantage of this analysis compared to the mass function method is that it does not require measurements in a controlled volume, and will work for a random sample of observed galaxy clusters. Depending on observational possibilities [153, 154, 155, 156, 157, 158, 159], this might provide a new way of

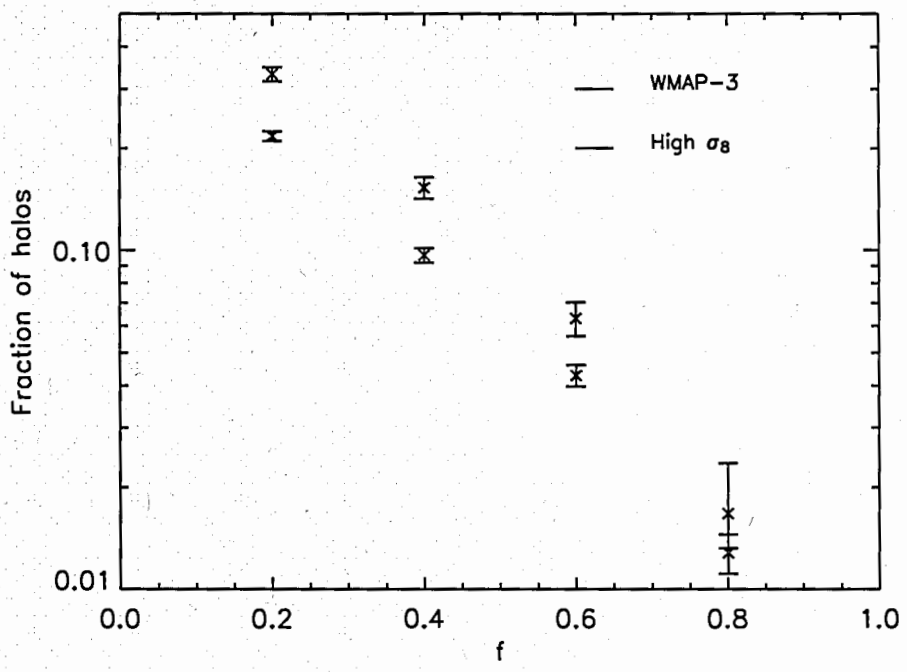

Figure 4.11: Fraction of the total number of halos in the mass range relevant to clusters, ($M_{200} \geq 10^{14} M_{\odot}/h$), as a function of the halo satellite mass fraction.

characterizing cosmologies using clusters of galaxies, or at least be a valuable method to cross-check results from mass function constraints.

The halo outliers with values of $M_{FOF}/M_{200} > 1.5$ are also a possible source of systematic bias for certain HOD applications. Given some halo mass bin above the fiducial mass cutoff for a given HOD, a bridged halo would be assigned a central galaxy with the same probability as an isolated halo. The probability of a satellite galaxy in a bridged halo [with the main halo having high mass companion(s)] is likely significantly higher than in an isolated halo. Therefore, applying the same HOD to both halo types would downweight the number of satellite galaxies, the precise amount depending on the mass range considered.

4.5 Conclusions

This chapter presents results from an analysis of idealized NFW halos and N-body simulations with the aim of clarifying the connection between FOF and SO halos, focusing mainly on the issue of halo masses and attempting to account for some of the unavoidable difficulties in simplifying a multi-scale problem in terms of primitive halo concepts. We found that a large fraction of FOF halos in N-body simulations ($80-85\%$) are relatively isolated and well-fitted by NFW profiles. This allows them to have SO counterparts, albeit the mass mapping is a two-parameter function $M_{SO} = M_{SO}(M_{FOF}, c)$, inferred from the properties of idealized NFW halos (c is the NFW halo concentration). In principle, this mock halo technique can be trivially extended to M_{Δ} with Δ values more directly

useful for cluster analyses (e.g., $\Delta = 500, 800, 1000$), or indeed to any other useful definition of the observable mass.

The rest of the halos, a fraction of $15-20\%$, appear to be dominated mainly by bridged halos. These halos consist of apparently localized structures (visually, or according to the SO halo definition) linked via density ridges into a common FOF halo, as discussed in Section 4.4. This degree of bridging is roughly consistent with X-ray observations of clusters, where in approximately $10-20\%$ of all cases there is a significant second component roughly within R_{100}, corresponding to the scale length of a $b = 0.2$ FOF halo (Vikhlinin, private communication). We have found that the bridged halo fraction rises as a function of mass, and when rescaled by the collapse mass scale M_*, also appears to be universal. We also find that in the cluster mass regime the fraction of halos with major satellites as function of the satellite's mass fraction is cosmology dependent.

The bridged FOF halo fraction complicates the procedure for transforming the global mass function. Accurate mapping between the global FOF and SO mass function must take into account SO multiplicity within FOF halos due to the bridging (which should be distinguished from the substructure mass function). Fortunately, if the bridged halo fraction is universal, then this problem can be (approximately) solved by one more iteration of the procedure described here. A simple prescription for handling the bridging problem, for example, may be the simultaneous use of two different linking lengths as a way of identifying substructure in the FOF halo identified with the longer ($b = 0.2$) linking length. Then, with mock halo mappings available for the shorter linking length, one would construct a new mass function which should be almost free of bridging artifacts to at least the 5% level.

In this work, systematic and statistical uncertainties were held to $\sim 5\%$, which represents the current state of the art in determining the halo mass function. The sensitivity of halo masses to simulation parameters such as force and mass resolution has not yet been satisfactorily controlled below this level. While further improvement is not ruled out, the universality of the FOF mass function is not known to be valid at or better than this level either.

The finite bridged halo fraction points to the existence of some level of bias when applying simple HOD schemes for the distribution of galaxies in halos, due to the existence of (minor/major) halo substructure. In standard HOD methods, halos are often selected, or assumed to be selected, by the FOF algorithm. However, this standard method then assumes a spherically-symmetric (usually NFW) distribution of satellite galaxies within halos, which is possibly at odds with a significant fraction of real halos [160, 161]. The fraction of problematic, irregular morphology FOF halos is mass-dependent, creating thereby a mass dependent source of error. Furthermore, any concentration dependence of the fraction of bridged FOF halos makes it difficult to parameterize halo properties purely as a function of halo mass, which is standard within HOD methods.

Despite these difficulties, the availability of sufficiently high resolution simulations should yield a completely satisfactory HOD more or less independent of the particular halo definition used (FOF or SO), provided that a realistic satellite distribution is implemented. The point is that, even with such a simulation, a simplified description of halos such as an NFW profile for populating halos with galaxies, would certainly fail for a not insignificant fraction of halos, and be a cause of systematic errors.

As an alternative to mapping SO mass functions beginning with the universal form of the FOF mass function, and utilizing the cosmology-dependent concentration-mass relation and its scatter, one could instead take the more computationally expensive approach of computing SO mass functions from simulations that sample a range of plausible cosmologies [133]. The additional expense of such an approach can be drastically reduced by the use of efficient statistical sampling and interpolation techniques that have been successfully demonstrated for cosmic microwave background temperature anisotropy and for the mass power spectrum [127, 128].

Finally, we remain agnostic as to the value of particular choices of halo definitions and masses in cosmological applications. For X-ray observations of relaxed clusters, the SO approach appears to be more natural since one fits directly to a spherically averaged profile as is observational practice. High-resolution views of the gas distribution in clusters [158] are hardly consistent with spherical symmetry, however, and the physics of the underlying robustness of the mass-observable relations remains to be fully established. Turning to other applications such as optical group and cluster and subcluster member identification, there may be no option but the use of (some version of) FOF techniques. Analagous to our bridged FOF halos, Sunyaev-Zel'dovich observations are likely to suffer from bridging of closely-neighboring clusters. Mock catalogs for ongoing and future cluster observations carried out via the Sunyaev-Zel'dovich effect have been built using FOF definitions for clusters (albeit with shorter linking lengths than $b = 0.2$), as the possible systematics from using spherical halo definitions are not clear [162].

5 Halo Clustering and Bias

Dark matter halos form through hierarchical structure formation, a characteristic feature of cold dark matter cosmology. These halos are not uniformly distributed throughout the universe, but form a filamentary network, or 'the cosmic web' [163]. Determination of spatial correlation between halos, as well as between mass fluctuations in general, is thus one of the main goals of cosmology. One way of measuring statistics of density fluctuations is via weak lensing observations. While these are sensitive to the unbiased distribution of all matter, they have several limitations: they measure matter in angular projection and have large sampling errors on large scales. On small scales, modelling of the weak lensing signal is difficult due to non-linear corrections. The alternative to this approach is to measure a well defined class of objects (galaxies, clusters of galaxies...) in redshift space [164, 165]. Of course, in general they will be *biased* tracers of mass, and this bias must depend on the scale over which it is measured. However, the requirement of homogeneity ensures that on large enough scales the bias will asymptote to a constant value. Besides this scale dependence of the bias, there is also a mass dependence: more massive, rare density peaks cluster more strongly than average density perturbations [36, 37, 166, 7].

As bias increases with increasing mass of correlated objects, clustering can be used as an additional mass estimate. For example, clusters of galaxies present a powerful cosmological probe, especially via their mass function (chapter 3). However, determining masses from observations is a notoriously difficult problem, as the scatter in the mass–observable relation is usually significant. Two point correlation function measurements can reduce that scatter with self–calibration techniques [167, 168, 169, 170].

5.1 Halo Clustering

The spectrum of perturbations can be quantified by the two-point correlation function ξ, which represents the probability of finding two overdensities at a given separation:

$$\xi = \langle \delta(\mathbf{r}_1, t)\delta(\mathbf{r}_1 + \mathbf{r}_2, t) \rangle \ , \tag{5.1}$$

where the ensamble average is taken over realizations of the universe. Then the probability of finding two object in volumes δV_1 and δV_2 separated by a distance d is

$$dP = n^2 \left[1 + \xi(d)\right] \delta V_1 \delta V_2 \ , \tag{5.2}$$

n being the average number of objects per unit volume. Analogously, one can define correlation functions of higher order, but they will not be completely independent, as they are connected by an infinite system of equations obtained by taking moments of the Boltzmann equation. This infinite series of equations is called Bogolyubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy, and can be solved with some condition for closure [171].

5.1.1 Calculation of the Two-point Correlation Function

The two point correlation function can be obtained from simulations by counting the number of halo pairs as a function of separation. It is commonly normalized to the same quantity calculated for a Poisson random field:

$$\xi(d) = \frac{N_{\rm hh}}{N_{\rm poisson}} - 1 \ . \tag{5.3}$$

Similarly, one can calculate the matter two point correlation function.

5.2 Asymptotic Bias

As in analytical approaches to modeling the halo mass function, the starting point in bias analyses is the feasibility to associate peaks in the linear density field to dark matter halos which arise from non-linear evolution. Recent, high resolution numerical analysis [172] show good correspondance of halos to peaks, especially for high-mass halos, although early works in the field claimed exactly the opposite [123]. It is further assumed, that on larger scales a linear bias relates matter density and halo density field:

$$\delta_h = b\,\delta \ , \tag{5.4}$$

where b is the asymptotic bias[1]. Similarly, we can write for power spectra:

$$P_h(k) = b^2 P_m(k) \ . \tag{5.5}$$

Although it is assumed that each halo corresponds to a peak, it is clear that peaks in general are not motionless, as they will go along large-scale flows. Thus, while the Lagrangian overdensity δ_L is always the same, the Eulerian overdensity will evolve as:

$$\delta_h(a) = \delta_L + \delta_m(a) \ , \tag{5.6}$$

where the second term describes the motion of density peaks. On large scales it is assumed that peaks simply follow the global flows, in which case the second

[1] In many papers, the bias at large scales is just denoted as bias, but since we will also look at its scale dependence, we use the term asymptotic bias to avoid possible confusion.

term is simply the matter overdensity. This procedure is known as the peak-background split [36, 37]. From the equation (5.6), we can write the Eulerian bias as:

$$b(a) = \frac{\delta_h(a)}{\delta_m(a)} = \frac{\delta_L + \delta_m(a)}{\delta_m(a)} = 1 + b_L(a) \ . \tag{5.7}$$

As the matter fluctuations grow in time, the Lagrangian bias b_L will decay at the same rate. In the region where the background overdensity grows to δ, the treshold density for the collapse will be reduced to $\delta_c - \delta$, and the number of halos of mass M will increase by a factor [173]:

$$\frac{n(M, z, \delta_c - \delta)}{n(M, z, \delta_c)} \ .$$

The Lagrangian bias is then

$$b_L = \frac{dn/d\delta}{n} = -\frac{dn/d\delta_c}{n} \ . \tag{5.8}$$

As a result, we see that within the above formalism (assumptions), asymptotic bias is directly related to the derivative of the mass function. As mass functions are usually defined in the universal form $f(\sigma)$ (§3.1.2) – and are commonly just fits to some functional form of σ, the number density of halos in a mass bin M, $M + dM$ at a redshift z is given by:

$$n(M, z)dM = \frac{\rho_b}{M}\left(-\frac{1}{\sigma}\frac{d\sigma}{dM}\right) f(\sigma)dM \ . \tag{5.9}$$

Thus for a Press-Schechter mass function where

$$n \propto \delta_c \exp\left(-\frac{\delta_c^2}{2\sigma^2}\right) \ , \tag{5.10}$$

the (Eulerian) bias has a simple form:

$$b_{\rm PS} = 1 - \frac{1}{\delta_c} + \frac{\delta_c}{\sigma^2} \ . \tag{5.11}$$

This expression was given by Cole and Kaiser in 1989 [173], and rederived by Mo and White in 1996 [36]. For the case of very light halos, where $\sigma \gg \delta_c$, the second term in the above equation will dominate the third, thus such halos are expected to be anti-biased ($b < 1$). The most massive halos ($\sigma \ll \delta_c$), on the other hand, should be biased, and the more massive the halo sample – the more the bias. Similarly, for the case where the number density of objects is given by the Sheth-Tormen formula the asymptotic bias is [37]:

$$b_{\rm ST} = 1 - \frac{1}{\delta_c} + a\frac{\delta_c}{\sigma^2} + \frac{2p/\delta_c}{1 + (a\delta_c^2/\sigma^2)^p} \ , \tag{5.12}$$

where $p = 0.3$ and $a = 0.707$ are obtained by fitting simulation data (see §3.1.3).

Clearly, we can derive bias using the peak-background split for any desired mass function. For example, we have seen in Chapter 3 that the Warren et al. fit [8] is a description of the mass function, good to ~10%. As it is a pure fit to $f(\sigma)$, we will have to rewrite it to include δ_c, enabling us to apply the peak-background split. Therefore, the number density of objects reads:

$$n_{\rm W} \propto \left[\frac{1}{1.686^a}\left(\frac{\delta_c}{\sigma}\right)^a + b\right] \exp\left[-\frac{c}{1.686^2}\left(\frac{\delta_c}{\sigma}\right)^2\right] . \tag{5.13}$$

Here, $a = 1.625$, $b = 0.2538$, and $c = 1.1982$ (see table 3.1). Using this, we find the corresponding bias as:

$$b_{\rm W} = 1 + \frac{2c}{\delta_c}\frac{1}{\sigma^2} - \frac{a}{\delta_c}\frac{\sigma^{-a}}{(\sigma^{-a}+b)} . \tag{5.14}$$

In the spherical model, the treshold for the collapse – the barrier – is independent of the mass scale considered, and is $\delta_c \approx 1.686$. The mass function (PS) can then be derived by assuming that the overdensity around a given point exhibits a Brownian motion random walk as a function of radius (i.e. the smoothing scale) [72]. Of course, the collapse occurs when the barrier is first crossed. In this formalism, the mass function, and thus halo bias, depends on the $\sigma(M)$ relation specifyed by the initial power spectrum, and the shape of the barrier. As an improvement of this, Sheth, Mo and Tormen (SMT) considered an elliptical collapse moving barrier model. While the barrier shape cannot be derived analytically, one can simulate an ensemble of random walk realizations, record the distribution of first crossings, and fit it by some convenient function of $\sigma(M)$. SMT find from their set of simulations, that a good barrier approximation is given by

$$B_{\rm SMT}(\sigma, z) = \sqrt{a}\delta_c(z)\left[1 + b\left(\frac{\sigma^2}{a\delta_c(z)}\right)^c\right] , \tag{5.15}$$

the best fit being $a = 0.707$, $b = 0.5$, and $c = 0.6$. The asymtotic bias is then related to the random walks which travel far from the origin before crossing the barrier. To ensure this, SMT consider a barrier with a high height. The bias is then:

$$b_{\rm SMT} = 1 + \frac{1}{\sqrt{a}\delta_c(z)}\left[\sqrt{a}x + \sqrt{a}bx^{1-c} - \frac{x^c}{x^c + b(1-c)(1-c/2)}\right] , \tag{5.16}$$

with $x \equiv a[\delta_c(z)/\sigma]^2$.

5.2.1 Bias Measurements from Simulations

To calculate halo bias, we turn to the same simulation set used for the mass function determination. These 60 simulations are summarized in Table 3.2. Due to the modest mass resolution of each simulation (256^3 particles in a box), the

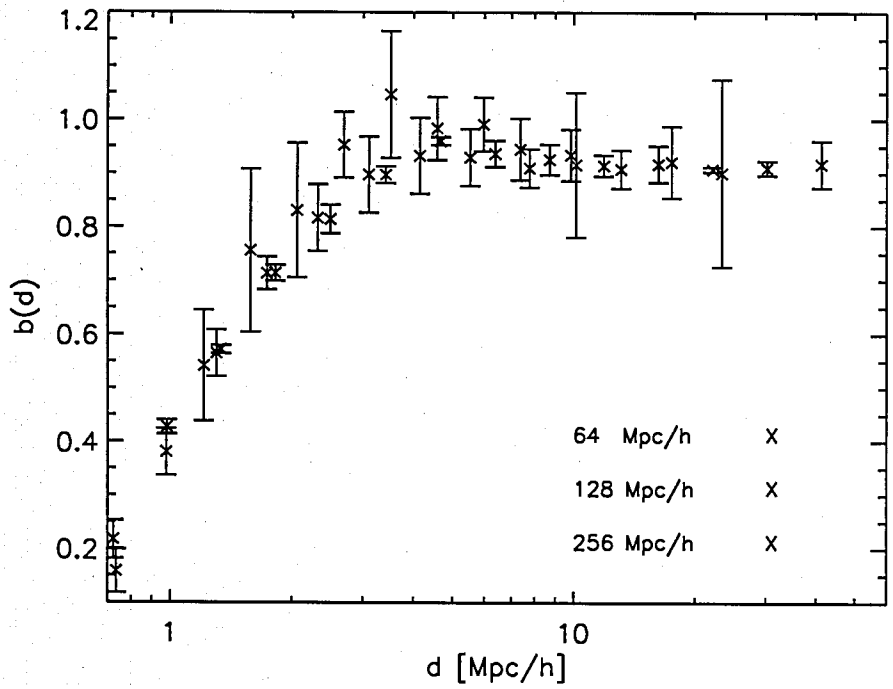

Figure 5.1: Halo bias as a function of distance for the same mass bin ($3.34 \times 10^{12} - 2.67 \times 10^{13} h^{-1} M_{\odot}$) from three simulations with different box sizes. The data are in good agreement, showing that the finite mass and force resolution do not affect the bias. 1σ errors between realizations are shown.

range of masses we can probe in any single simulation is limited. The reliable determination of the correlation function requires many more halos per mass bin than the calculation of the mass function itself. As the mass function drops steaply for high values of $1/\sigma$ (see Figs. 3.13 and 3.14), at high redshifts we commonly have only a narrow mass range for calculating the bias. On the other hand, the large ensemble of simulations with 7 different box-sizes and different mass resolutions (in physical units) will enable us to probe the bias over a wider range of masses than it was possible in any study before. Since we have many statistically independent realizations of each box-size, we can also keep the statistical errors low.

When calculating the halo bias, the goal is of course, to keep the statistical errors as small as possible, but at the same time to have mass bins which are contingent across different simulation boxes. The latter will enable better control of systematic errors, and in particular over errors arising from finite box sizes, as well as finite mass and force resolution. The separation of the mass bins by a factor of two results in a noisy correlation functions due to small number of halos in a bin. For this reason, we separated our bins by a factor of eight[2] in mass.

If the results for the same mass bin agree across different box-size simula-

[2]The mass resolution increases by a factor of 8 when going to smaller boxes – see Table 3.2.

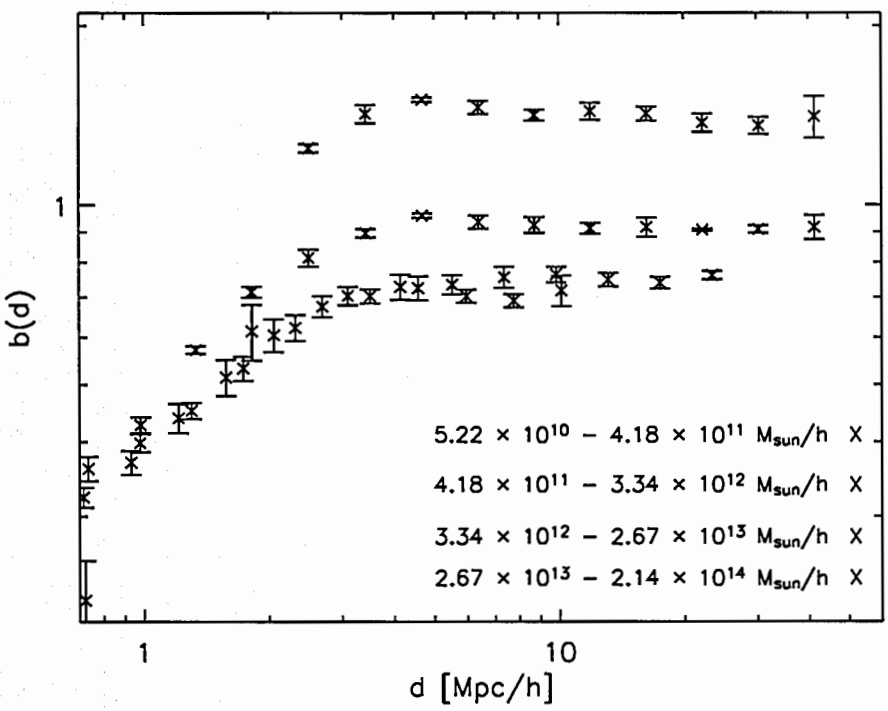

Figure 5.2: Halo bias for $z = 0$; shown are 4 mass bins, taken from the largest available boxes in order to reduce the errorbars.

tions, we can be confident that they are systematics-free, considering that the simulations themselves have being throughly tested for their accuracy (§3.3). Figure 5.1 presents one such resolution test; at redshift $z = 0$ we compare the halo bias from 64, 128, and $256h^{-1}$Mpc boxes for a $3.34 \times 10^{12} - 2.67 \times 10^{13}h^{-1}M_{\odot}$ mass bin. The errorbars are 1σ deviations between realizations. Naturally, errorbars are decreasing when increasing the box size, as there are more and more halos, but the agreement between different box-sizes is excellent, especially in the asymptotic regime ($d > 5h^{-1}$Mpc) that is of most interest here. The fact that the halo bias is not affected by the finitness of the simulation domain, was also shown in a recent paper by Reed et al. [7], on boxes as small as $1h^{-1}$Mpc. This is due to the fact that the large-scale power is missing equaly in both halo and matter correlation functions. Thus, the correlation functions themselves show a clear suppresion due to the missing power, but the bias is still accurate. Figure 5.1 also shows that the bias can be still trusted at 1/5 of the box size; the conservative requirement when calculating correlations is that the volume should be at least an order of magnitude larger than the correlation length, but this seems to be too conservative for this application.

Figure 5.2 shows the halo bias from our simulations at redshift $z = 0$. As before, we consider only the data up to one fifth of the box size. As expected, at all masses halo bias shows scale dependence: first there is an exclusion zone at small distances which comes from the fact that halos have a finite size (as opposed to the matter distribution, sampled via particles). Next, there is a

region where the bias strongly depends on separation, and clearly depends on the halo mass considered, as different bins have very different fall-offs towards smaller scales. Hamana et al. [174] express the scale dependence of the bias as:

$$b(\sigma, z, d) = b_{\rm ST}\left[1 + b_{\rm ST}\sigma(d, z)\right]^{\alpha} \ , \tag{5.17}$$

with $\alpha = 0.15$. Diaferio et al. [175] find steeper scale dependence: $\alpha = 0.35$. Reed at al. [7] find that this functional form fits their low redshift data well, but becomes increasingly inaccurate at high redshifts. The best fit to their high redshift data is:

$$b(\sigma, z, d) = b_{\rm SMT}\left[1 + 0.03 b_{\rm SMT}^{3}\sigma^{2}(d, z)\right] \ . \tag{5.18}$$

Note that Reed et al. use the SMT bias formula rather than ST, although the two are very close. Finally, Figure 5.2 shows that, as expected, at large distances the halo bias is scale independent. Of course, smaller mass halos reach this regime – which is effectively the homogeneity scale for that mass – at smaller scales than high mass halos. For this reason, it seems dangerous to define the asymptotic bias at some beforehand prescribed distance, as done by some groups. For example, Cohn and White [176] take the bias at $1.5h^{-1}$Mpc as an asymptotic value in their $z = 10$ analysis; Reed et al. show that this leads to an overestimation of the bias, as at that scale it did not reach its asymptotic value. To avoid this problem, rather than assuming certain scale, we calculate the asymptotic bias as the average value of the last 4 bias points, calculated for the largest distances.

5.3 Results and Discussion

Figure 5.3 shows the asymptotic bias from our simulations (blue points). We also plot halo bias from the Millenium simulation [6], which covers redshifts z=0–5. As the Millenium simulation has ~500 times more particles than our simulations, Gao et al. were able to bin their data into much narrower mass bins. Still, we see that our data is in very good agreement with the Millenium simulation, showing the usefullness of nested-box simulations. Moreover, the agreement gives us additional confidence in our data, as the Millenium simulation covers $500h^{-1}$Mpc on a side, thus the finite box effects are negligible.

We consider the Warren-like form of bias to fit our data:

$$b = 1 + \frac{2c}{\delta_c}\nu^2 - \frac{a}{\delta_c}\frac{\nu^a}{(\nu^a + b)} \ , \tag{5.19}$$

with $\nu = \delta_c/sigma(M)$, describing how rare a mass bin is ($\nu = 1$ are characteristic halos forming at a given epoch). While the bias we calculate does not suffer from finite box effects, we apply extended Press-Schecheter box correction

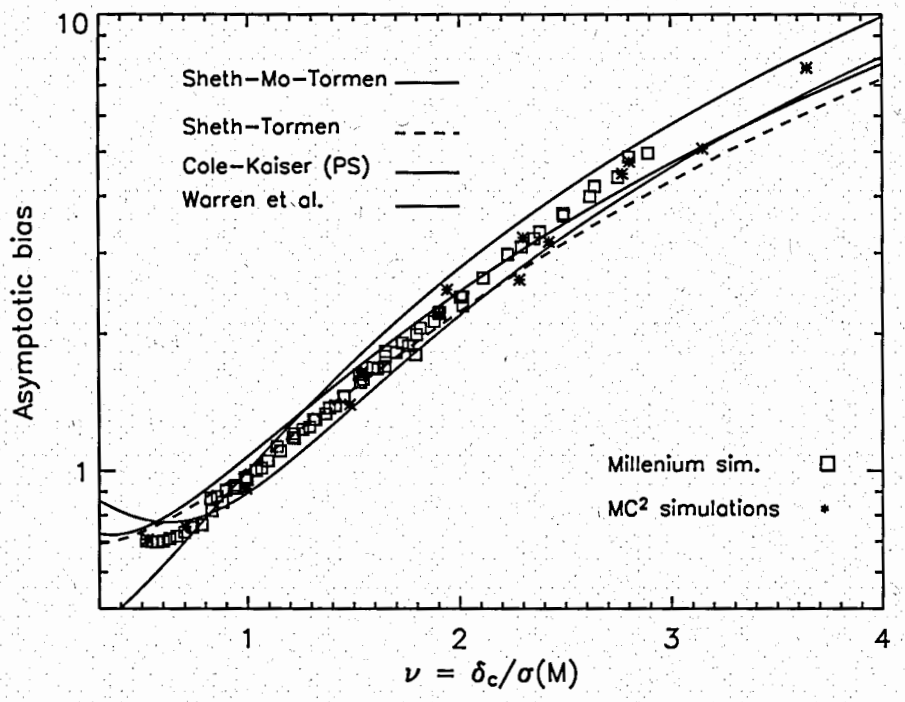

Figure 5.3: Asymptotic bias as a function of 'halo rarity'. Lines are different analytical predictions, discussed in section 5.2; red squares are data from the Millenium simulation [6], while blue points are data from our simulations.

(as explained in details in §3.4.6) when determining ν. The least–square fit to our data yields $a = 1.09$, $b = 0.21$, and $c = 0.42$. Figure 5.4 shows the ratios of several simulation data-sets to this fit. In addition to our and the Millenium data, we show here the data from Reed et al. [7], covering redshifts $z = 15 - 30$. We have excluded their $z = 10$ points (which comprised most of their data). Those are most likely to be inaccurate, as for the cosmology they consider a $1h^{-1}$Mpc box is completely non-linear at $z = 10$. All data sets are in good agreement with our bias fit. However, if we consider the mass function which would correspond to the above bias:

$$f(\nu) = A\,(\nu^a + b)\exp\left(-c\nu^2\right)\,, \tag{5.20}$$

we see large deviations from the Warren et al. [8] mass function fit, as shown in Figure 5.5. Note that fitting bias leaves the normalization of the mass function, A unconstrained; still deviations of $\pm 40\%$ from the Warren et al. fit are ruled out by the accurate mass function data (see Chapter 3). This indicates that, while peak-background split formalism is qualitatively correct, it fails at the level of accuracy needed today. It has to be emphasized that the nature of approximations in that formalism are not rigorous, as they are based on heuristic arguments. Thus, the failure in using mass function formula for describing the bias should not surprise – it simply indicates halo bias should be determined

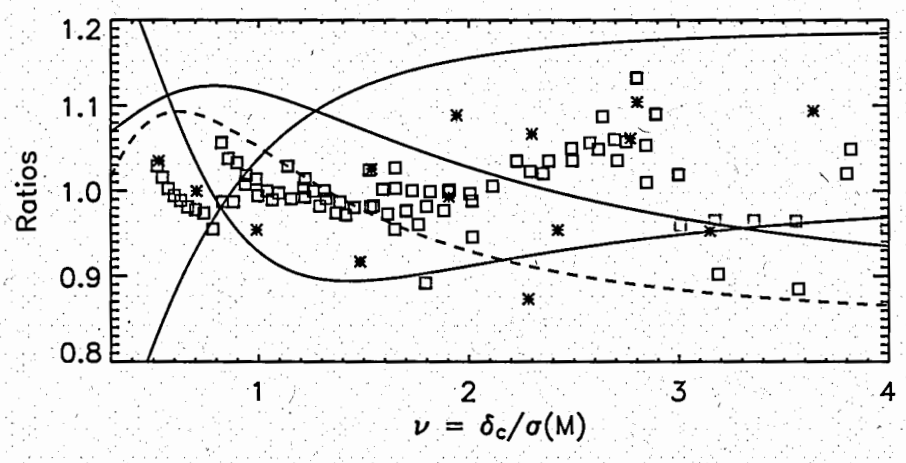

Figure 5.4: Ratio of several data sets and analytical predictions to the best fit to MC^2 data alone. In addition, we show the data from the Millenium simulation (red squares), and the data from Reed et al. [7] (yellow squares). The lines are different analytical predictions, following Fig. 5.3.

independently from simulations.

Finally, as we show here one example of bias fit and the corresponding mass function, we have to stress it is not the only functional form of $b(\nu)$ (or $f(\nu)$) we have tested. We experimented with several more functional forms, all ending up with the same result: the best fits to the mass function systematically depart from our measured bias, and the best fits to the bias do not fit our mass function data.

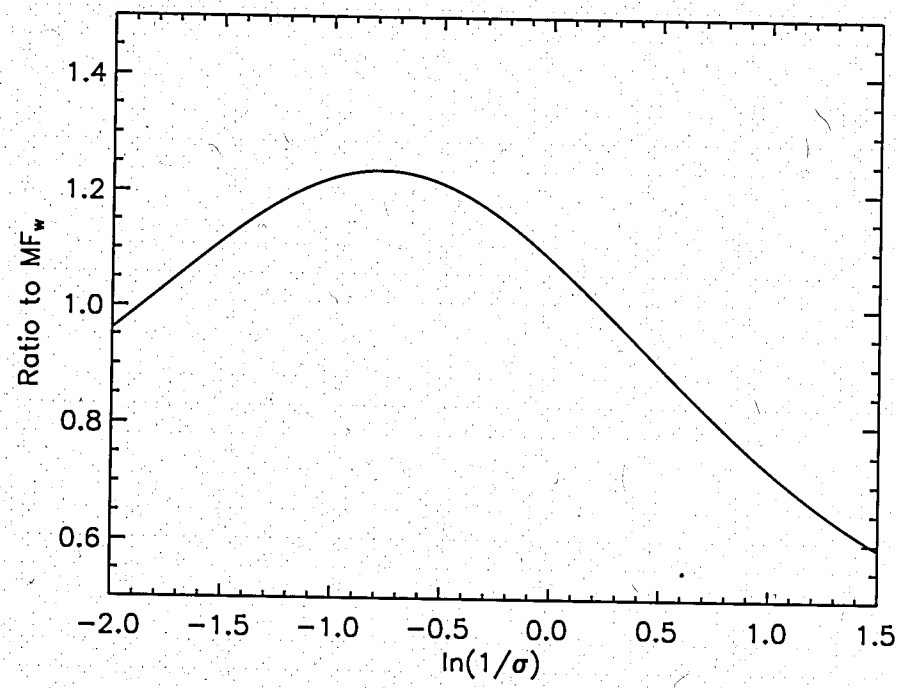

Figure 5.5: Ratio the mass function corresponding to the best fit for the halo bias, to the Warren et al. [8] mass function. While the normalization is arbitrary, the variation with respect to the Warren et al. fit is larger than allowed by the current state of the art halo mass function data sets.

6 Robustness of Cosmological Simulations

The last three decades have seen the emergence of cosmology as 'precision science', moving from order of magnitude estimates, to predictions and measurements at accuracy levels better than 10%. Cosmic microwave background observations and large galaxy surveys have led this advance in the understanding of the origin and evolution of the Universe. Future surveys promise even higher accuracy, at the one percent level, over a considerably wider dynamic range than probed earlier. In order to fully utilize the wealth of upcoming data and to address burning questions such as the dynamical nature of dark energy (most often parametrized by the equation of state parameter w), theoretical predictions must attain at least the same level of accuracy as the observations, even higher accuracy being certainly preferable. The highly nonlinear physics at the length scales probed, combined with complicated gas physics and astrophysical feedback processes at these scales, make this endeavor a daunting task.

Nature was kind that many of statistical measures, some of them analyzed in previous chapters, can be successfully modelled with gravity alone. Thus a very fist step is to test if numerical codes can reach the desired accuracy for gravitational interactions, down to the relevant nonlinear scales. Tests with exact solutions such as pancake collapse [56] are valuable for this task, but as shown in Ref. [90] the results do not easily translate into statements about the accuracy of different simulation algorithms in realistic cosmological simulations. Exactly solvable problems are typically highly symmetric and hence somewhat artificial. Codes optimized for realistic situations can break down in certain tests even if their results appear to converge in physically relevant settings. Therefore, in order to evaluate the accuracy of simulation codes, a broad suite of convergence and direct code comparison tests must be carried out.

The codes used in this comparison project are all well-established, and have been key drivers in obtaining numerous scientific results. They are based on different algorithms and are employing different methods for error control. The code developers have already carried out careful convergence tests themselves and verified to their satisfaction that the codes yield reliable results. But because of the multi-scale complexity of the dynamical problem itself, as well as the incompleteness of most convergence tests, it is necessary to do much more. Therefore, the aim here is to focus on comparing results from a suite of different codes for realistic cosmological simulations. In order to avoid uncertainties from statistical sampling, all codes are run with exactly the same initial conditions,

and all results are analyzed using the same diagnostic tools.

The organization is the following: in Section 6.1 are described ten simulation codes used for the comparison study. In Section 6.2 simulations carried out for this project are presented. Finally, comparison results are shown in Section 6.3 and conclusion in Section 6.4.

6.1 The Codes

The ten codes used in this work cover a variety of methods and application arenas. The simulation methods employed include parallel particle-in-cell (PIC) techniques (the PM codes MC^2 and PMM, the PM/AMR codes Enzo and FLASH), a hybrid of PIC and direct N-body (the AP^3M code Hydra), tree algorithms (the treecodes PKDGRAV and HOT), and hybrid tree-PM algorithms (GADGET-2, TPM, and TreePM).

The PIC method models many-body evolution problems by solving the equations of motion of a set of tracer particles which represent a sampling of the system phase space distribution function. A computational grid is used to increase the efficiency of the self-consistent inter-particle force calculation. To increase dynamic range, local force computations (e.g., P^3M, tree-PM) and AMR are often used. The grid also provides a natural basis for coupling to hydro-solvers.

Treecodes are based on the idea that the gravitational potential of a far-away group of particles is accurately given by a low-order multipole expansion. Particles are first arranged in a hierarchical system of groups in a tree structure. Computing the potential at a point turns into a descent through the tree. Treecodes naturally embody an adaptive force resolution scheme without the overhead of a computational grid. Tree-PM is a hybrid algorithm that combines a long-range force computation using a grid-based technique, with shorter-range force computation handled by a tree algorithm. In the following we give a brief description of each code used in this comparison study.

6.1.1 The Grid Codes

MC^2

The multi-species **M**esh-based **C**osmology **C**ode MC^2 code suite includes a parallel PM solver for application to large scale structure formation problems in cosmology. In part, the code descended from parallel space-charge solvers for studying high-current charged-particle beams developed at Los Alamos National Laboratory under a DOE Grand Challenge [177, 178]. MC^2 solves the Vlasov-Poisson system of equations for an expanding universe using standard mass deposition and force interpolation methods allowing for periodic or open boundary conditions with second and fourth-order (global) symplectic time-stepping and a Fast Fourier Transform (FFT) based Poisson solver. The results reported here were obtained using Cloud-In-Cell (CIC) deposition/interpolation. The overall

computational scheme has proven to be remarkably accurate and efficient: relatively large time-steps are possible with exceptional energy conservation being achieved.

PMM

Particle-Multi-Mesh (PMM) [179] is an improved PM algorithm that combines high mass resolution with moderate spatial resolution while being computationally fast and memory friendly. The current version utilizes a two-level mesh FFT-based gravity solver where the gravitational forces are separated into long-range and short-range components. The long-range force is computed on the root-level, global mesh, much like in a PM code. To obtain higher spatial resolution, the domain is decomposed into cubical regions and the short-range force is computed on a refinement-level, local mesh. This algorithm achieves a spatial resolution of 4 times better than a standard one-level mesh PM code at the same cost in memory. In [179], PMM is shown to achieve very similar accuracy to that of MC^2 when run with the same minimum grid spacing.

Enzo

Enzo[1] is a publicly available, extensively tested adaptive mesh refinement (AMR), grid-based hybrid code (hydro + N-Body) which was originally written by Greg Bryan, and is now maintained by the Laboratory for Computational Astrophysics at UC San Diego [180, 181, 89]. The code was originally designed to do simulations of cosmological structure formation, but has been modified to examine turbulence, galactic star formation, and other topics of interest. Enzo uses the Berger & Colella method of block-structured adaptive mesh refinement [182]. It couples an adaptive particle-mesh method for solving the equations of dark matter dynamics [183, 184] with a hydro solver using the piecewise parabolic method (PPM), which has been modified for cold, hypersonic astrophysical flows by the addition of a dual-energy formalism [185, 186]. In addition, the code has physics packages for radiative cooling, a metagalactic ultraviolet background, star formation and feedback, primordial gas chemistry, and turbulent driving.

FLASH

FLASH [187] originated as an AMR hydrodynamics code designed to study X-ray bursts, novae, and Type Ia supernovae as part of the DOE ASCI Alliances Program. Block-structured adaptive mesh refinement is provided via the PARAMESH library [188]. FLASH uses an oct-tree refinement scheme similar to [189] and [190]. Each mesh block contains the same number of zones (16^3 for the runs here), and its neighbors must be at the same level of refinement or one level higher or lower (mesh consistency criterion). Adjacent refinement

[1] http://lca.ucsd.edu/codes/currentcodes/enzo

levels are separated by a factor of two in spatial resolution. The refinement criterion used is based upon logarithmic density thresholds. Numerous extensions to FLASH have been developed, including solvers for thermal conduction, magnetohydrodynamics, radiative cooling, self-gravity, and particle dynamics. In particular, FLASH now includes a multigrid solver for self-gravity and an adaptive particle-mesh solver for particle dynamics. Together with the PPM hydrodynamics module, these provide the core of FLASH's cosmological simulation capabilities. FLASH uses a variable time step leapfrog integrator. In addition to other time step limiters, the FLASH particle module requires that particles travel no more than a fraction of a zone during a time step.

6.1.2 The Tree Codes

HOT

This parallel tree code [191] has been evolving for over a decade on many platforms. The basic algorithm may be divided into several stages (the method of error tolerance is described in Ref. [192]). First, particles are domain decomposed into spatial groups. Second, a distributed tree data structure is constructed. In the main stage of the algorithm, this tree is traversed independently in each processor, with requests for nonlocal data being generated as needed. A `Key` is assigned to each particle, which is based on Morton ordering. This maps the points in 3-dimensional space to a 1-dimensional list, maintaining as much spatial locality as possible. The domain decomposition is obtained by splitting this list into N_p (number of processors) pieces. An efficient mechanism for latency-hiding in the tree traversal phase of the algorithm is critical. To avoid stalls during nonlocal data access, effectively explicit 'context switching' is done using a software queue to keep track of which computations have been put aside waiting for messages to arrive. This code architecture allows HOT to perform efficiently on parallel machines with fairly high communication latencies [193]. HOT has a global time stepping scheme. The code was among the ones used for the original Santa Barbara Cluster Comparison Project [194] and also supports gas dynamics simulations via a smoothed particle hydrodynamics (SPH) module [195].

PKDGRAV

The central data structure in PKDGRAV [196, 197] is a tree structure which forms the hierarchical representation of the mass distribution. Unlike the more traditional oct-tree which is used in the Barnes-Hut algorithm [198] and is implemented in HOT, PKDGRAV uses a k-D tree, which is a binary tree. The root-cell of this tree represents the entire simulation volume. Other cells represent rectangular sub-volumes that contain the mass, center-of-mass, and moments up to hexadecapole order of their enclosed regions. PKDGRAV calculates the

gravitational accelerations using the well known tree-walking procedure of the Barnes-Hut algorithm. Periodic boundary conditions are implemented via the Ewald summation technique [199]. PKDGRAV uses adaptive time stepping. It runs efficiently on very large parallel computers and has produced some of the world's highest resolution simulations of cosmic structures. A hydrodynamics extension called GASOLINE exists.

6.1.3 The Hybrid Codes

Hydra

HYDRA [200] is an adaptive P^3M (AP^3M) code with additional SPH capability. In this work we use HYDRA only in the collisionless mode by switching off gas dynamics. The P^3M method combines mesh force calculations with direct summation of inter-particle forces on scales of two to three grid spacings. In regions of strong clustering, the direct force calculations can become significantly expensive. In AP^3M, this problem is tackled by utilizing multiple levels of subgrids in these high density regions, with direct force computations carried out on two to three spacings of the higher-resolution meshes. Two different boundary conditions are implemented in HYDRA, periodic and isolated. The time step algorithm in the dark matter-only mode is equivalent to a leapfrog algorithm.

GADGET-2

The N-body/SPH code GADGET-2 [201, 96] employs a tree method [198], to calculate gravitational forces. Optionally, the code uses a tree-PM algorithm based on an explicit split in Fourier space between long-range and short-range forces [202]. This combination provides high performance while still retaining the full spatial adaptivity of the tree algorithm, allowing the code to reach high spatial resolution throughout a large volume. By default, GADGET-2 expands the tree multipoles only to monopole order, in favor of a compact tree storage, a cache-optimized tree-walk, and consistent and efficient dynamic tree updates. The cell-opening criterion used in the tree walk is based on an estimator for the relative force error introduced by a given particle-cell interaction, such that the tree force is accurate up to a prescribed maximum relative force error. The latter can be lowered arbitrarily, if desired, at the expense of higher calculation times. The PM part of GADGET-2 solves Poisson's equation on a mesh with standard fast Fourier transforms, based on a CIC mass assignment and a four-point finite differencing scheme to compute the gravitational forces from the potential. The smoothing effects of grid assignment and interpolation are corrected by an appropriate deconvolution in Fourier space. The time-stepping of GADGET-2 uses a leap-frog integrator which is symplectic in case constant timesteps (in the log of the expansion factor) are employed for all particles. However, the code is normally run in a mode where individual and adaptive timesteps are used to

speed up the calculation time. To this end, the timesteps for the short-range dynamics are allowed to freely adapt to any power of two subdivision of the long-range timestep. GADGET-2 is fully parallelized for massively parallel computers with distributed memory, based on the MPI standard. The code can also be used to simulate hydrodynamical processes using the particle-based smoothed particles hydrodynamics (SPH) method (e.g. [203]), in an entropy conserving formulation [204], a feature which is however not exercised in the simulations considered here.

TPM

TPM [104, 105] is a publicly-available hybrid code combining a PM and a tree algorithm. The density field is broken down into many isolated high-density regions using a density threshold criterion. These contain most of the mass in the simulation but only a small fraction of the volume. In these regions, the gravitational forces are computed with the tree algorithm while for the bulk of the volume the forces are calculated via a PM algorithm, the PM time steps being large compared to the time-steps for the tree-algorithm. The PM algorithm uses the CIC deposition/interpolation scheme and solves the Poisson equation using FFTs.The time integrator in TPM is a standard leap-frog scheme: the PM time steps are fixed whereas tree particles have individual time steps, half of the PM step or smaller.

TreePM

The algorithmic structure of the TreePM code [81] is very similar to GADGET-2. The particles are integrated using a second-order leap-frog method, with position and canonical momentum as the variables. The time step is dynamically chosen as a small fraction (depending on the smoothing length) of the local free-fall time and particles have individual time steps. The force on any given particle is computed in two stages. The long-range component of the force is computed using the PM method, while the short range component is computed from a global tree. A spline softened force law is used. The tree expands forces to monopole order only, and cells are opened based upon the more conservative of a geometric and relative force error criterion. The PM force is computed by direct FFT of the density grid obtained from CIC mass assignment.

6.2 The Simulations

A previous code comparison suite [90] considered three cosmological test problems: the Santa Barbara Cluster [194], and two large-scale structure simulations of ΛCDM models in a $64h^{-1}$Mpc box and a $256h^{-1}$Mpc box. In the latter two cases, the primary target of this previous work was to investigate results in

a medium resolution regime, addressing statistical quantities such as the two-point correlation function, the density fluctuation power spectrum, and the dark matter halo mass function.

In this work we focus further attention on one of these tests, the smaller of the ΛCDM boxes. Due to the small box size, the force resolution of all codes – including the pure mesh codes – is in principle sufficient to analyze properties of individual halos themselves. This allows us to extend the dynamic range of the code comparison to higher resolution than studied earlier. In this new regime, we expect to see a much broader divergence of results because of the more demanding nature of the test. (Even in the previous analysis [90], the power spectrum was unexpectedly deviant at the larger wavenumbers considered.) Our aim is to characterize the discrepancies and attempt to understand the underlying causes.

All codes were given exactly the same particle initial conditions at a redshift $z_{\rm in} = 50$. The initial linear power spectrum was generated using a fit to the transfer function [205], a modification of the BBKS fit [145]. This fit does not capture baryon oscillations but takes baryonic suppression into account (these details are of only limited relevance for the test). The cosmology underlying the simulations is given by $\Omega_{\rm CDM} = 0.27$, $\Omega_{\rm b} = 0.044$, $\Omega_\Lambda = 0.686$, $h = 0.71$, $\sigma_8 = 0.84$, and $n = 0.99$. The simulation was run with 256^3 particles, which leads to an individual particle mass of m_p=1.362·10$^9 h^{-1}$M$_\odot$.

While performing a comprehensive code comparison study which involves very different algorithms – such as grid and particle-based methods in the present case – a central and difficult question immediately arises: what is the most informative way to compare the codes and learn from the results? The difficulty is compounded by the fact that codes are often optimized under different criteria and controlling numerical error is a complex multi-parameter problem in any case, even for codes that share the same general underlying algorithm.

As a case in point, let us consider the choice of force resolution for each code. (Since the volume and number of particles are fixed, the mass resolution is the same for each run.) One option would be to run all codes with the same formal force resolution but this, aside from wasting resolution for the high-resolution codes, also suffers from the problem that it is not easy to compare resolutions across different algorithms; moreover, time-stepping errors also must be folded into these sorts of estimates. Finally, such a comparison would be rather uninteresting, because realistic cosmological simulations are run with higher resolutions than would be possible in a conservative test of this type: Interesting effects on small scales would be missed. A more uncontrolled, but nevertheless useful option is to allow every simulator to run her or his code with close to the optimal settings they would also use for a scientific run (given the other restrictions imposed by the test problem). In this case, a more realistic comparison can be performed in which we can access the robustness of conclusions from cosmological simulations. Here, while our approach adheres more

Table 6.1: Softening lengths measured in h^{-1}kpc. The different smoothing kernels have been converted into Plummer softening equivalents by matching the potential at the origin. While this procedure is only approximate, it makes a comparison of the different force resolutions more meaningful. Mesh codes not listed here (PMM, Enzo, Flash) have the same force resolution as MC^2.

MC^2	HOT	PKDGRAV	Hydra	GADGET-2	TPM	TreePM
62.5	7.1	1.6	28.4	7.1	5.1	5.7

closely to the second strategy, we do try to assess at what length scales one should expect a specific code to break down assuming that the resolution of the code is accurately estimated by the simulator.

The nominal resolutions for the different codes for the performed runs are as given in Table 6.1. We have converted the different softening kernels into Plummer equivalents following the normalization conventions of Ref. [206]. We have matched the different softening kernels ϕ at zero and compared them at this point. With the normalization conventions in Ref. [206], we find:

$$\phi_{\rm Plummer}(0) \propto \frac{1}{\epsilon}, \tag{6.1}$$

$$\phi_{\rm Spline}(0) \propto \frac{7}{5}\frac{1}{\epsilon}, \tag{6.2}$$

$$\phi_{\rm K_3}(0) \propto \frac{2079}{512}\frac{1}{\epsilon}, \tag{6.3}$$

where ϵ is the softening length. The grid resolution of the PM and AMR codes is roughly equivalent to the Plummer softening. HOT and Hydra have Plummer force kernels implemented, PKDGRAV uses Dehnen's K_3 kernel [206] and the three tree-pm codes use spline kernels. With the above definitions, it is easy to convert the spline and K_3 kernels into Plummer via

$$\epsilon_{\rm Spline} = 1.4\epsilon_{\rm Plummer}, \tag{6.4}$$

$$\epsilon_{\rm K_3} = 4.06\epsilon_{\rm Plummer}, \tag{6.5}$$

which we used to standardize the force resolution quotes in Table 6.1. We note that some of the codes below could have been run at higher resolution, and the values below should not be thought of as resolution limits. In fact, the choices of these values represent compromises due to run time considerations as well as a (loosely) pre-planned scatter to try and determine the effects of force resolution on the simulation results.

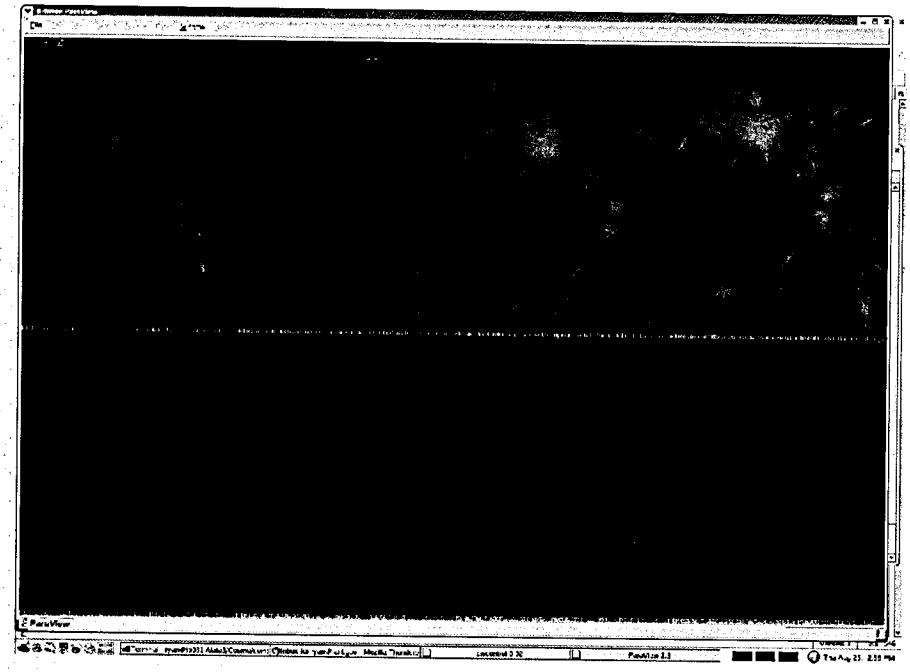

Figure 6.1: Screenshot of the comparative visualization manager in ParaView. Upper row: results from four different codes, zoomed into a dense region of the simulations. Particles are displayed as arrow glyphs, colored with respect to their velocity magnitude. Lower row: same region, the particles now displayed simply as dots.

6.3 Results

6.3.1 Results for the Full Simulation Box

As an initial test, a simple view of the simulation output at $z = 0$ proves to be very useful. ParaView [207] offers a comparative visualization option in which the results from different simulations can be shown simultaneously. Manipulation on any one output in this mode results in the same manipulation for all the others. A screenshot of the comparative visualization manager is displayed in Figure 6.1 – a zoom into an arbitrary region of the simulation box showing simultaneous results from four different codes. In the upper row a subset of the particles is shown as arrow glyphs, colored by velocity magnitude, the lower row shows the particles as dots with the same coloring scheme. A quick inspection of these snapshots reveals that the code 2 run had a problem with the velocities and code 4 had slightly incorrect boundary conditions (the whole picture being shifted upward)[2].

Figure 6.2 shows a comparison of the final GADGET-2 and Enzo outputs. We show a subsample of 20,000 particles, each displayed with vector arrow glyphs, sized and colored by their velocity magnitude. The arrow glyphs nicely represent the flows in the box to the major mass concentrations. As to be

[2]Of course these initial bugs were fixed before going on to the final results discussed below!

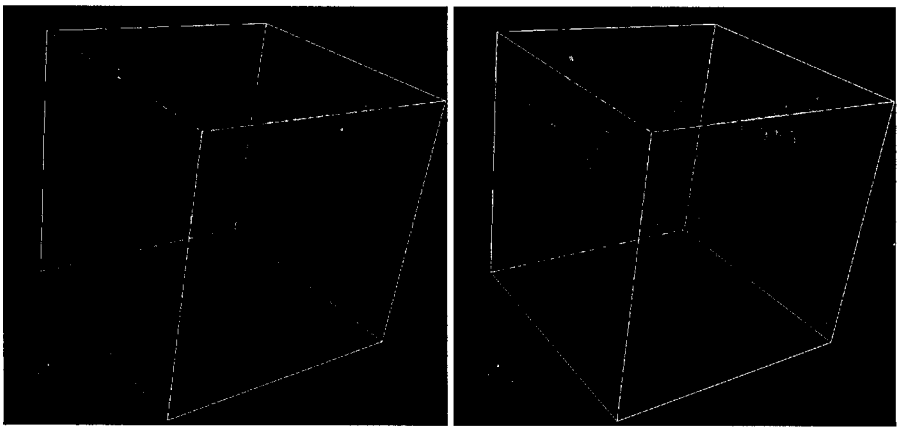

Figure 6.2: A subset of the 20,000 particles at $z = 0$ from the GADGET-2 simulation (left) and the Enzo simulation (right). The particles are shown with vector arrow glyphs which are sized and colored by their velocity magnitude (blue: slowest, red: fastest).

expected, particles in the field are slow (blue), while the particles in the halos have the largest velocities (yellow to red). While the overall appearance of both simulations shown is very similar, subtle differences can be seen (e.g., there are no small structures in the flow regions in the Enzo simulations), indicating the higher resolution employed in the GADGET-2 run. (Five of the biggest halos in the simulation will be examined in more detail below, the resolution differences becoming significantly more apparent.)

6.3.2 Dark Matter Halos

The halo paradigm is central to any large-scale structure analysis; dark matter in simulations, discretized in the form of heavy collisionless particles, forms clearly visible filaments (stripes) and halos (clumps of dark matter) through the process of gravitational instability. Figure 6.2 shows these structures clearly for the simulations studied in this paper. This picture agrees well with observations of galaxy rotation curves, and velocity dispersions of galaxies in clusters which favor scenarios where luminous, baryonic matter is embedded in massive, extended, and close to spherical conglomerates of dark matter.

As discussed in Chapter 4, the density profiles of dark matter halos are well fitted by the Navarro, Frenk, and White [130, 74] profile, Eq. 4.3. Here we are interested in the variation of the profiles produced by the different codes, tending towards the outer region of the halo. This variation may be significant for determining halo masses via the often used FOF algorithm. The mass that the halo finder will 'see', strongly depends on the density and density gradient close to the virial radius (R_{200}) of a halo. On the other hand, accurately reproducing the inner slope of a halo profile is the prime test of the code's force resolution.

Table 6.2: Halo 3 data: distance of the center from the mean value for all codes, and the mass of the halo from different simulations.

Code	ΔX_c [kpc/h]	ΔY_c [kpc/h]	ΔZ_c [kpc/h]	M [10^{14} $M_\odot$/h]
MC^2	-86.23	158.81	-14.68	2.749
PMM	201.68	33.90	10.24	2.757
Enzo	-21.36	45.16	11.36	2.745
FLASH	-41.66	-22.56	-23.10	2.726
HOT	-30.02	-120.54	43.99	2.720
PKDGRAV	38.58	52.19	-43.98	2.679
Hydra	19.91	-28.29	0.77	2.721
GADGET-2	-27.08	-59.00	-0.70	2.705
TPM	-36.37	-35.09	1.04	2.697
TreePM	-17.45	-24.62	13.63	2.727

On scales below this resolution limit, particle positions get randomized, resulting in a flattened density profile (numerical errors can also lead to a sharpening of the profile due to an associated unphysical damping).

We first compare the five heaviest halos from the simulations; their masses range between approximately 2 to $5 \cdot 10^{14} h^{-1} M_\odot$, thus each halo is sampled with 150,000 or more particles. The individual halo masses (as found by the FOF algorithm) are in agreement within 3% for all ten codes. Note that the FOF masses found for the grid codes are slightly higher. This is presumably due to their lower resolution in this comparison, resulting in less tight halos. The FOF halo finder can identify more particles in the fuzzier outskirts of lower resolution simulations as belonging to the halo than in the high resolution runs. The centers of the halos are defined by the minimum of the local potential of the halo. Here the agreement among the codes is even better than for the masses – the difference is less than 0.5% of the box size. In Table 6.2 we show the center and mass of one of the halos, Halo 3. This halo (also shown in Figure 6.7) has the size and mass of a group of galaxies. The dispersion in the mass and position of the center is similar for the other halos, whose profiles we investigate next.

In Figure 6.3 we present the spherically averaged density profiles for the five heaviest halos in the simulation. As an arbitrary reference, the black line represents the best NFW fit (Equation 4.3) for the TPM data. The fit is shown up to the inner 10 h^{-1}kpc of each halo. In addition, we show two residual panels for each halo profile. The upper panel shows the ratio of all codes with respect to GADGET-2, while the lower panel shows only the four grid codes and ratios with respect to MC^2.

The agreement in the outer part of the halos is excellent. As expected, the codes exhibit different behaviors on small scales (depending on their force resolution and time-stepping), thus the inner parts of halos are not always the same. While the high resolution codes successfully track the profile all the way in to the plotting limits of Figure 6.3, the profiles from the mesh codes depart

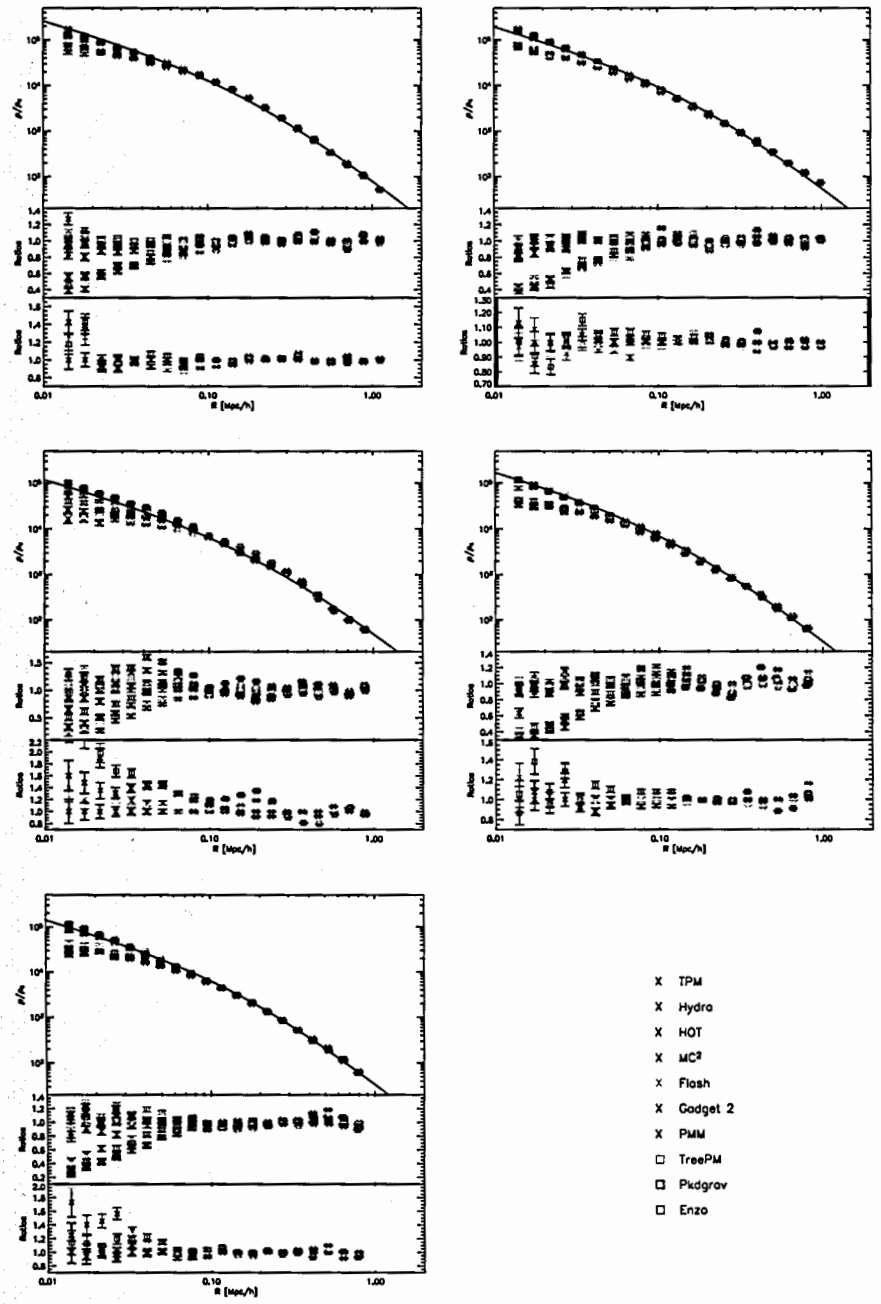

Figure 6.3: Halo profiles for the five heaviest halos in the simulation. The black line shows the best-fit NFW profile to the TPM simulation, mainly to guide the eye. In the outer regions all codes agree very well. In the inner regions the fall-off of the grid codes is as expected due to resolution limitations. The fall-off point can be predicted from the finite force resolution and agrees well with the results. The middle panel in each plot shows the ratio of the different codes with respect to GADGET-2. The lower panels show only the four grid codes and the ratio with respect to MC^2.

much earlier (60-100 h^{-1}kpc), with approximately constant density in the core. The onset of the flattening is consistent with the nominal resolution of the grid codes, which is given in Table. 6.1. Note that among the mesh codes there is no significant difference between the fixed mesh codes which ran at the highest resolution throughout the whole simulation volume, and the AMR codes whose base mesh spacing is a factor of 4 times lower.

We now study three of the five halos in more detail, restricting attention to particles within a sphere of radius $2 \cdot R_{200}$. The profiles of the largest halo, Halo 1, shown in Figure 6.3, agree well down to $R = 0.06h^{-1}$Mpc; at smaller scales the finite resolution of the grid codes becomes apparent. Nevertheless, the grid codes and the high-resolution codes among themselves yield very consistent results. Figure 6.4 shows the density of Halo 1 for the lower resolution code PMM and the higher resolution code TreePM in two-dimensional projection. The two-dimensional density field is computed on a 100×100 grid within the $2 \cdot R_{200}$ region, projected onto the z-direction (another projection along the x-direction is also shown). The projected density field has been normalized by dividing out the mean density in this area. The mean density is very close across the different codes, hence the normalization allows for direct comparisons of the projected density fields. As mentioned earlier, the positions of the halo centers (density peaks) are in remarkably good agreement. Due to its higher resolution, the density in the center of the halo from the TreePM run is slightly higher (as to be expected from the profiles). In addition, TreePM shows slightly more substructure on the outskirts of the halo, displayed by the small "hills". Overall, the halo is very smooth and well defined, which is reflected in the good agreement of the profiles. The density plots for the four grid codes are very similar. The small structures around the halo in the other codes also show only very minor variations, thus the PMM and TreePM results can be considered to be representative.

The profiles of Halo 3 show substantially more variation among the different codes in the inner region, relative to the other four halos. Studying it in more detail, we first investigate a subset of four codes: MC^2, FLASH, GADGET-2, and HOT, covering a wide range of force resolutions. In Figure 6.5 we show a zoom into the center of the halo. The particles are shown in white. Superimposed on the particle distribution is a 2-dimensional density contour evaluated on a 100×100 grid and smoothed with a Gaussian filter, projected along the z-direction. (The contouring and filtering are intrinsic functions in ParaView.)

The overall appearance of the halo is remarkably similar between the codes, a major feature of the halo being its irregular shape. The left side of the halo is elongated and a second major peak has developed on the right, leading o a triangular shape in this projection. This irregularity (seen also very clearly in Figure 6.6) is most likely the reason for the disagreement in the inner part of the profiles. The halo has probably undergone a recent merger or is in the process of merging. Comparing the lower resolution runs from MC^2 and FLASH with

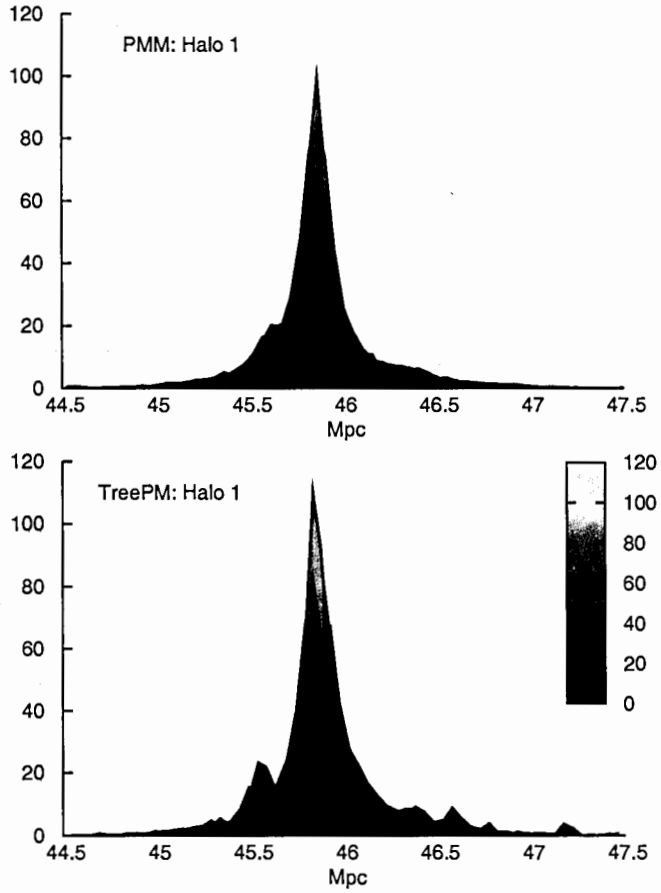

Figure 6.4: Projected and normalized two-dimensional density for Halo 1 from PMM (left) and TreePM (right). TreePM has a slightly higher density in the inner region of the halo than PMM, as to be expected from the different force resolutions. Overall the agreement is very good.

GADGET-2 and HOT, the effect of force resolution is very apparent, the high resolution runs producing significantly more substructure. GADGET-2 shows slightly more substructure than HOT, which could be due to the adaptive time stepping used in the GADGET-2 run relative to HOT's global time-step.

Figure 6.7 shows Halo 3 from another four runs. As in Figure 6.4, the two-dimensional density is shown on a 100×100 grid. The three-dimensional view underlines the rather complicated structure of the halo. PMM and Enzo show the elongated structure with two maxima, whereas the Hydra and PKDGRAV results differ somewhat from the other codes. They have a more well defined peak and do not exhibit much of the second structure. TreePM and TPM (not shown in the figure) are very similar to GADGET-2 and HOT. Overall, Halo 3 has much more interesting features than Halo 1, which leads to slight discrepancies in the halo profiles among the codes.

Last, we study Halo 4 from a subset of the codes: MC^2, GADGET-2, PKDGRAV, and HOT, covering the grid, tree-PM, and tree codes. The results are

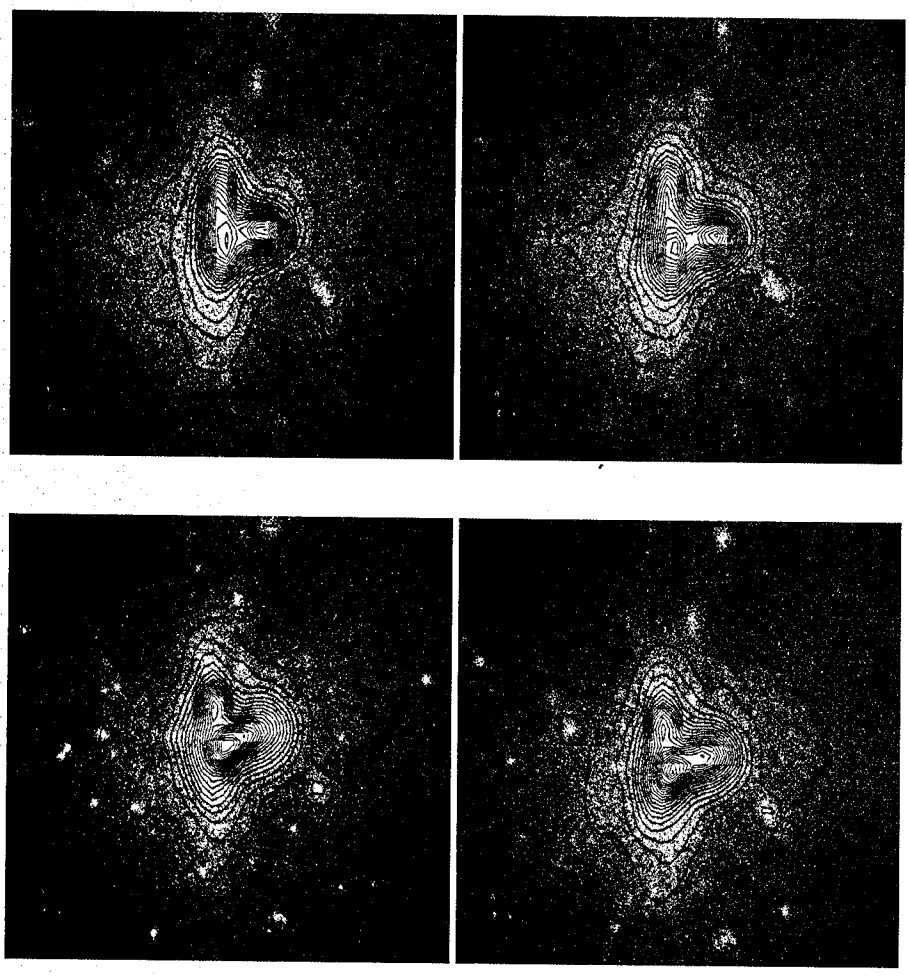

Figure 6.5: Two-dimensional contour plot of the projected density for Halo 3 from MC^2, FLASH, GADGET-2, and HOT (left upper to right lower plot). White: particles, black: contour smoothed with a Gaussian Filter.

shown in Figure 6.8. As before, the lower density of the PM code is due to its restricted resolution. Overall, the agreement is again very satisfying. The centers of the halos are in excellent agreement, and all four runs show a smaller structure on the left of the main halo. The exact details of the smallest structures are different which could be due to inaccurate time-stepping and discrepancies in the codes' output redshifts.

Overall, the comparison of the largest halos in the box is very satisfactory. The halo profiles agree on the scales expected from the code resolutions. Differences of the inner parts can be explained due to very irregular shapes as in Halo 3. The reader should keep in mind that we did not resimulate the halos with higher resolution, and that these halos ere extracted straight out of a cosmological volume simulation. Therefore, the level of agreement is in accord with theoretical expectations.

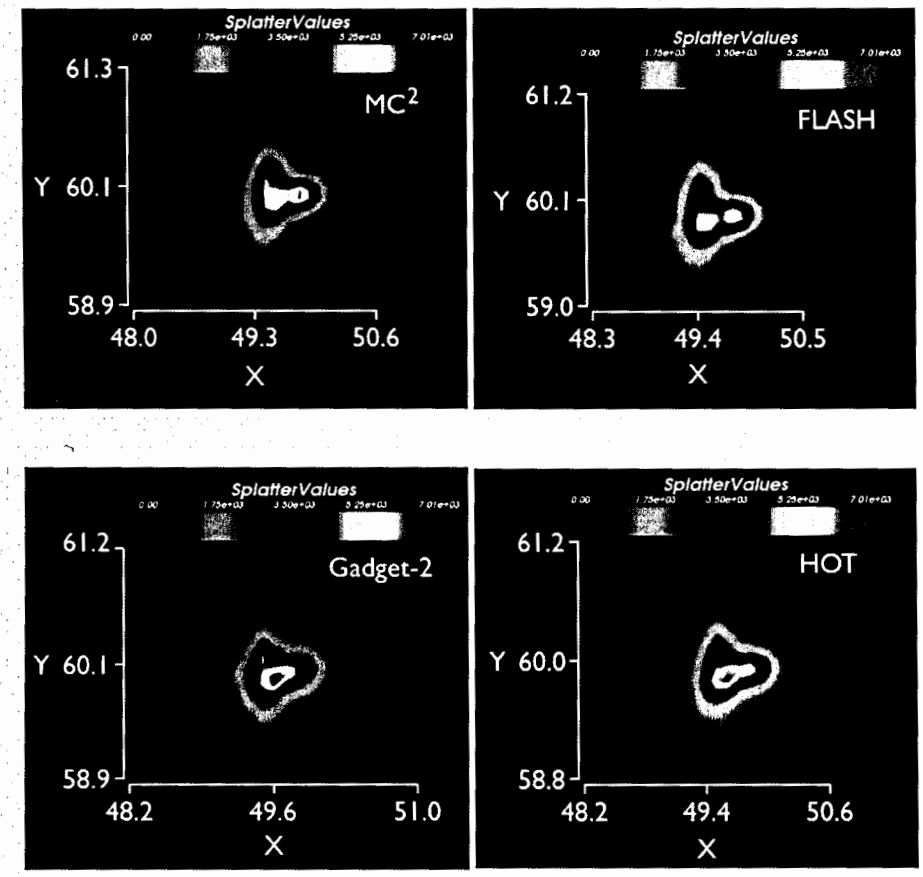

Figure 6.6: Same as in Figure 6.5: MC^2, FLASH, GADGET-2, and HOT.

6.3.3 Number Density of Halos

The Mass Function

As discussed in detail in Chapter 3, the mass function is an important statistic in cosmology. We have seen that the numerical study of the mass function poses several challenges to the simulation code, especially if one wants to obtain reliable results at the few percent accuracy level: the number of particles in a halo has to be sufficient in order to prevent systematic biases in determinations of the halo mass, the force resolution has to be adequate to capture the halos of interest, the simulation has to be started at sufficiently high redshift, and finite box corrections might have to be considered if the simulation box is small (see also Refs. [76, 124, 73, 8, 131]).

In this work we study the mass function at $z = 0$. We identify halos with a friends-of-friends algorithm [80] with linking length of $b = 0.2$. The smallest halo we consider has 10 particles, not because this is physically reasonable (usually the minimum number of particles is several times bigger), but because we are

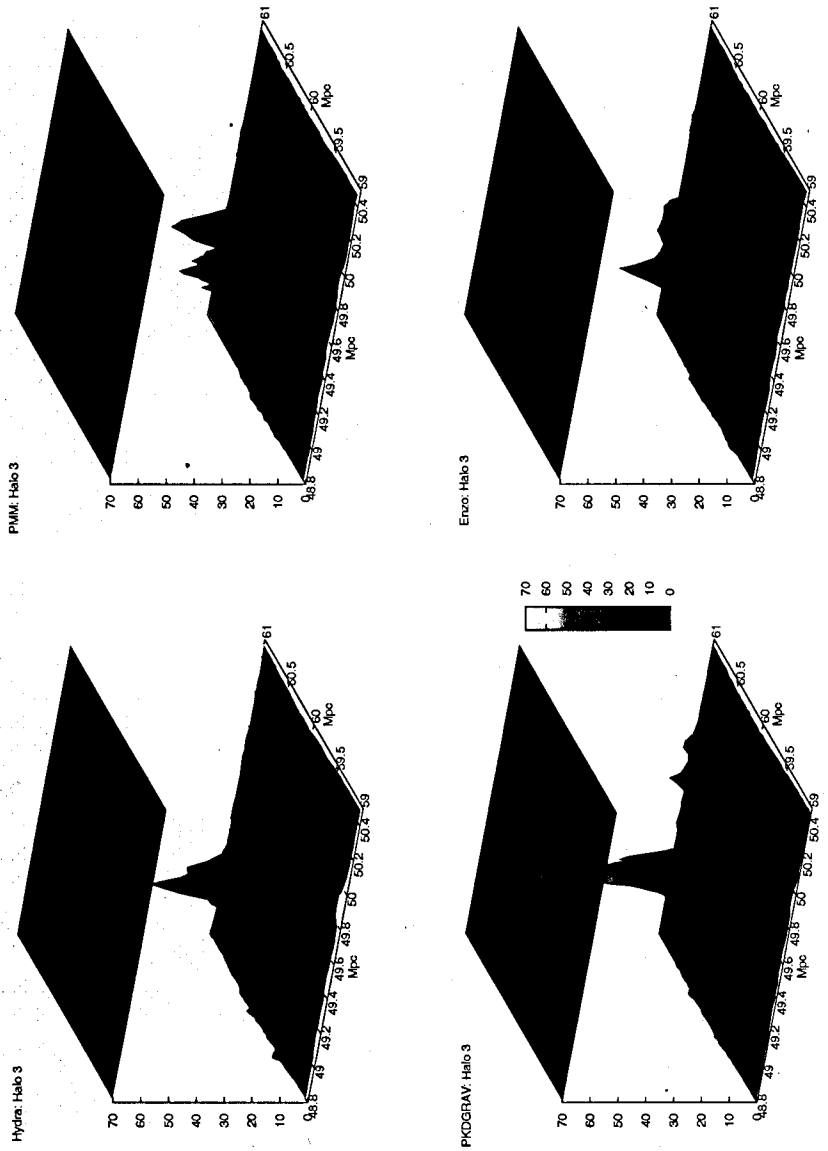

Figure 6.7: Two-dimensional densities from Hydra, PKDGRAV, PMM, and Enzo for Halo 3. The panel on the top of each graph shows the projected density. The color coding is the same for each plot, shown in the result for PKDGRAV.

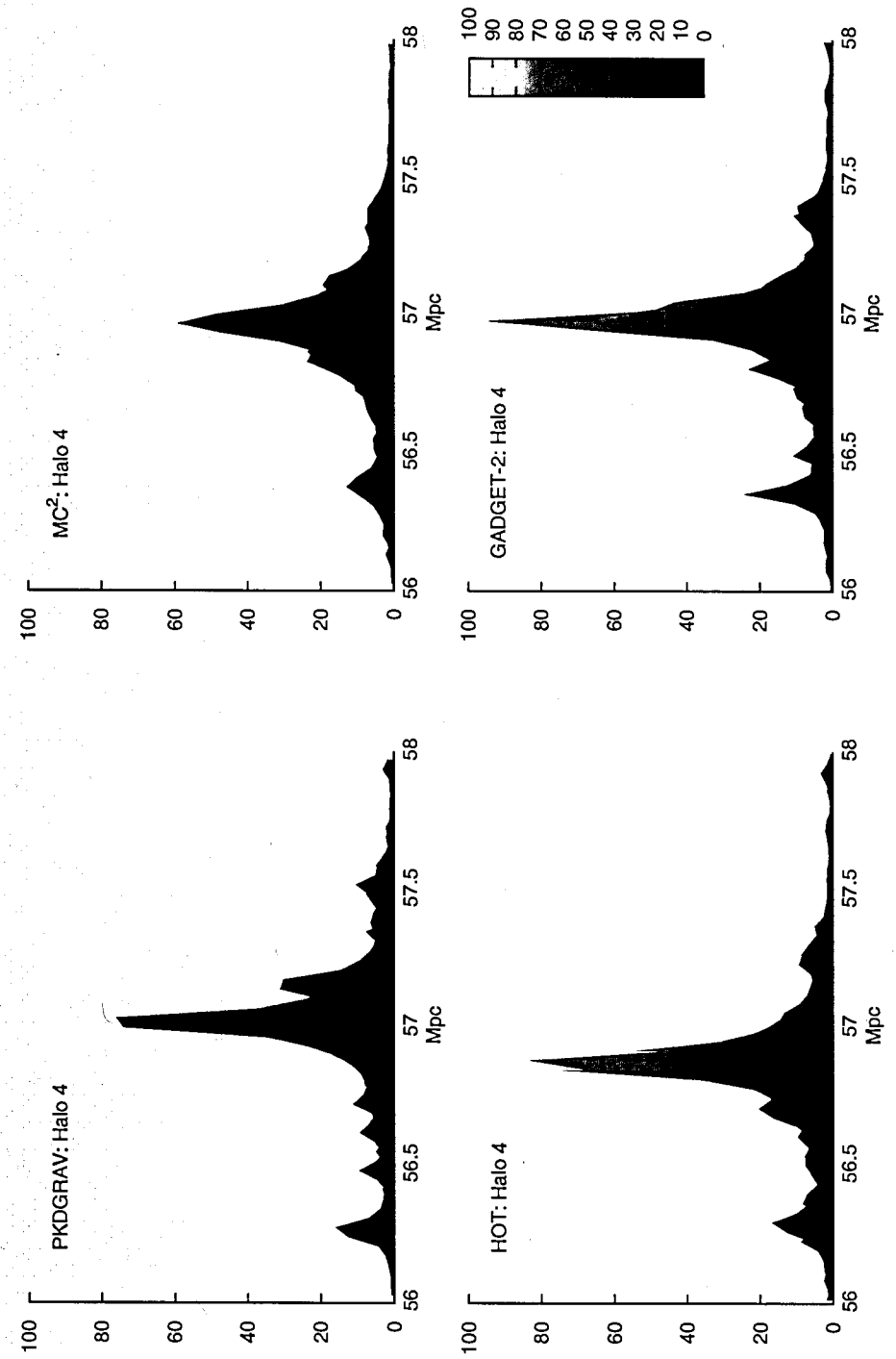

Figure 6.8: Two-dimensional density profile of Halo 4 for MC^2, GADGET-2, PKDGRAV, and HOT. MC^2 shows less substructure and is less dense in the inner region.

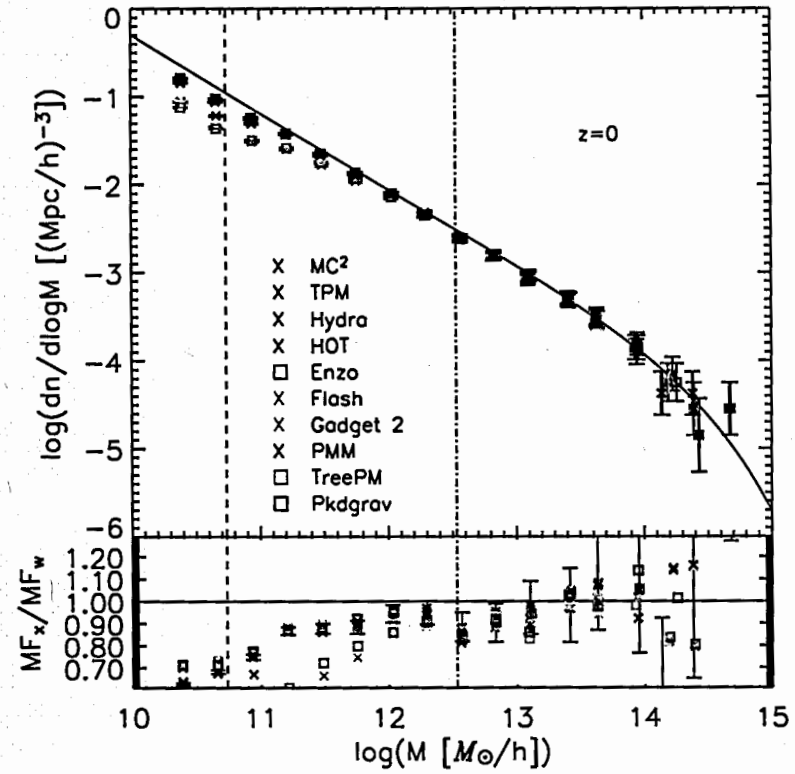

Figure 6.9: Mass function at $z = 0$, simulation results and the Warren fit (red line). Lower panel: residuals with respect to the Warren fit. For clarity we only show the error bars for one code. The dashed line indicates the threshold for 40 particles (force resolution limit for the PM codes, according to Equation (3.29)), the dotted-dashed line for 2500 particles (force resolution limit for the base grid of the AMR codes).

interested in cross-code comparison. We follow the suggestions by Warren et al. [8] and correct the halo mass for possible undersampling via Eq. 3.45 This correction lowers the masses of small mass halos considerably.

In order for small halos to be resolved, both mass and force resolution must be adequate. In Chapter 3, we derived the criterion for the force resolution:

$$\frac{\delta_f}{\Delta_p} < 0.62 \left[\frac{n_h \Omega_m(z)}{\Delta}\right]^{1/3} . \tag{6.6}$$

This equation predicts that all the non-grid codes have enough force resolution to resolve the smallest halos considered, while the two PM codes, MC^2 and PMM, have sufficient force resolution to resolve halos with more than 40 particles, and that the base grid of the two AMR codes restricts them to capturing halos with more than 2500 particles. Of course this is only a rough estimate in principle since the AMR codes increase their local resolution as a function of density threshold, the question is whether the criteria used for this is sufficient to resolve halos starting at 40 particles/halo.

We have indicated the resolution restrictions in Figure 6.9 by vertical lines

(dashed: 40 particles, dashed-dotted: 2500 particles). The predictions are good indicators of actual code results. The AMR codes fall off at slightly lower masses than given by 2500 particles. This shows that the resolution which determines the smallest halos being captured is dictated by the base grid of the AMR codes and not by the highest resolution achieved after refinement. Thus, for the AMR codes to achieve good results, significantly more aggressive density thresholding appears to be indicated. (Similar results were found in Refs. [90, 89].) As predicted, the mass functions of the PM codes start to deviate at around 40 particles from the other codes.

Overall the agreement among the codes is very good. For comparison, we show the Warren fit [8] in red. Due to limited statistics imposed by the small box-size, the purpose here is not to check the accuracy of the fitting function. At the high mass end, the scatter is as expected due to the rareness of high-mass halos. In the medium mass range between $10^{12.3}$ and $10^{13.4}h^{-1}\mathrm{M}_{\odot}$ all codes agree remarkably well, down to the percent level. In the small halo regime with as low as 40 particles, the agreement of the codes – besides the AMR codes as explained above – stays at this level. This indicates that the halo mass function is a very robust statistic and the simple resolution arguments given above can reliably predict the halo mass limits of the individual simulations.

The comparison yields one surprising result, however: the TPM code simulation has far fewer halos in the regime below 40 particles per halo than the other high resolution codes. This finding was already pointed out in Ref. [90]. In order to understand this deficit of halos in more detail we investigate the halo count as a function of environment in the following.

Halo Count and Density

In the last section we investigated the mass function and discovered a discrepancy of small halos in the two AMR codes and TPM. The hypothesis for the halo deficit in the AMR codes is, as discussed above, that the base grid resolution is too low and allows us only to catch halos with more than 2500 particles accurately. The coarse base grid in the initial state of the simulation does not allow for small halos to form and these halos cannot be recovered in the end. This would imply that the AMR simulations should have a deficit of small halos more or less independent of density: small halos should be missing everywhere, even in the highest refinement regions. A possible explanation for the missing halos in the TPM simulation could be a hand-over problem between the PM and the tree code. In this case, the number of small halos in high density regions should be correct. A qualitative comparison of three codes (HOT, MC^2, and TPM) is shown in Figure 6.10. The red points show halos with 10 particles, the white dots are a subset of the simulation particles. It is immediately obvious, that the halo counts in different environments, close to the large halo on the right, or on in the lower density regions on the left, are different. After this

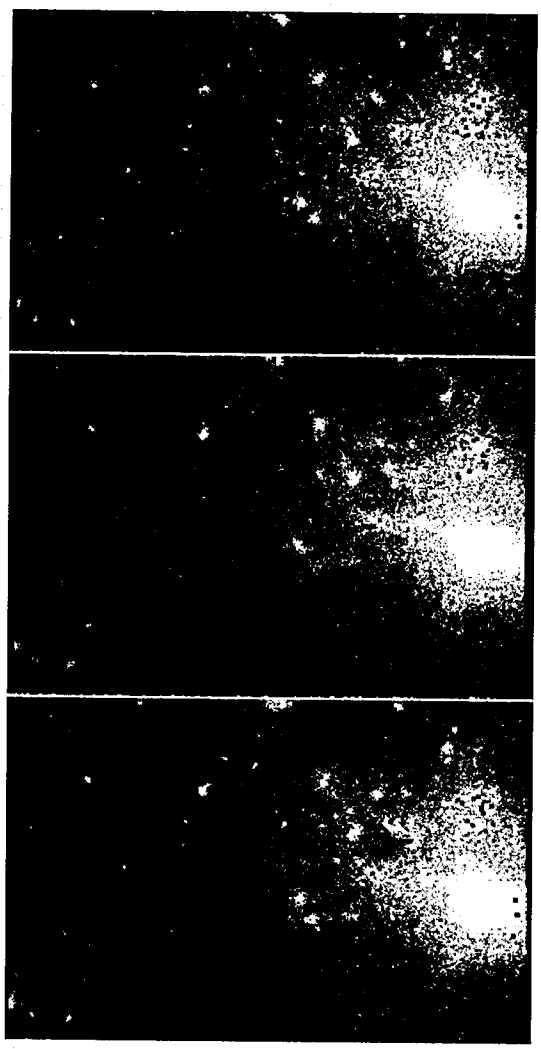

Figure 6.10: Small halos (10 particles) in the HOT, MC^2, and TPM simulation. Red points: halos, white dots: subset of the simulation particles. The distribution and number count of the small halos is different in all three codes.

qualitative result, we have to quantify this finding in order to come to a reliable conclusion about the cause for the halo deficits.

We use the ParaView visualization toolkit to implement a routine that calculates the density field on a (variable) grid from the particle distribution via a nearest grid point algorithm. The grid size for the density field is usually set by the requirement that the density field be not too noisy. As a first check we compare the density probability distribution function (PDF) for the different codes. It is clear that, if the grid for calculating the density is chosen coarse enough, details should be smoothed out and the PDFs for the different codes should be in good agreement. In Figure 6.11 we show the PDFs for all codes calculated on a 32^3 grid (upper panel) corresponding to a smoothing scale of $2h^{-1}$Mpc and a 64^3 grid (lower panel) corresponding to a smoothing scale of $1h^{-1}$Mpc. In

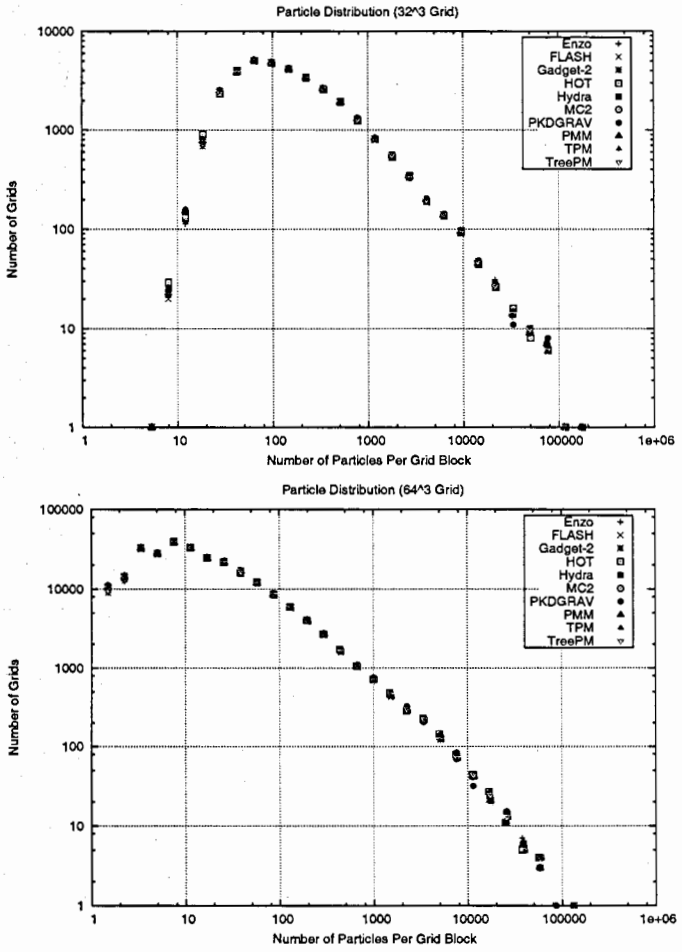

Figure 6.11: Probablity distribution function of the densities. Upper panel: calculation of the density on a 32^3 grid, lower panel: calculation of the density on a 64^3 grid.

both cases all codes agree extremely well, as to be expected since the smoothing scales are well beyond the code resolutions. We confirmed that this result holds also for finer grids, up to 256^3, which corresponds to the lowest resolution in the AMR codes Enzo and FLASH. The average number of particles in a grid cell ρ_b on the left panel is 512 particles per cell, in the right panel 64 particles per cell. If we look at the density contrast and define voids as regions with a density contrast $\delta^{\rm Void} = -0.8$, we find $\rho^{\rm Void} \simeq 100$ for the left panel and $\rho^{\rm Void} \simeq 13$ for the right panel. In both cases, this threshold is on the right of the maximum of the curves – a large fraction of the simulation volume is underdense.

To cast the results in a more quantitative light, Figure 6.12 displays the distribution of halos with respect to density for the two lower mass bins. We restrict our investigations to a density threshold of up to 100,000 particles per cell. Figure 6.12 shows the results for 10-40 particle (left panel) and 41-2500 particle halos (right panel). These thresholds were chosen because, as discussed

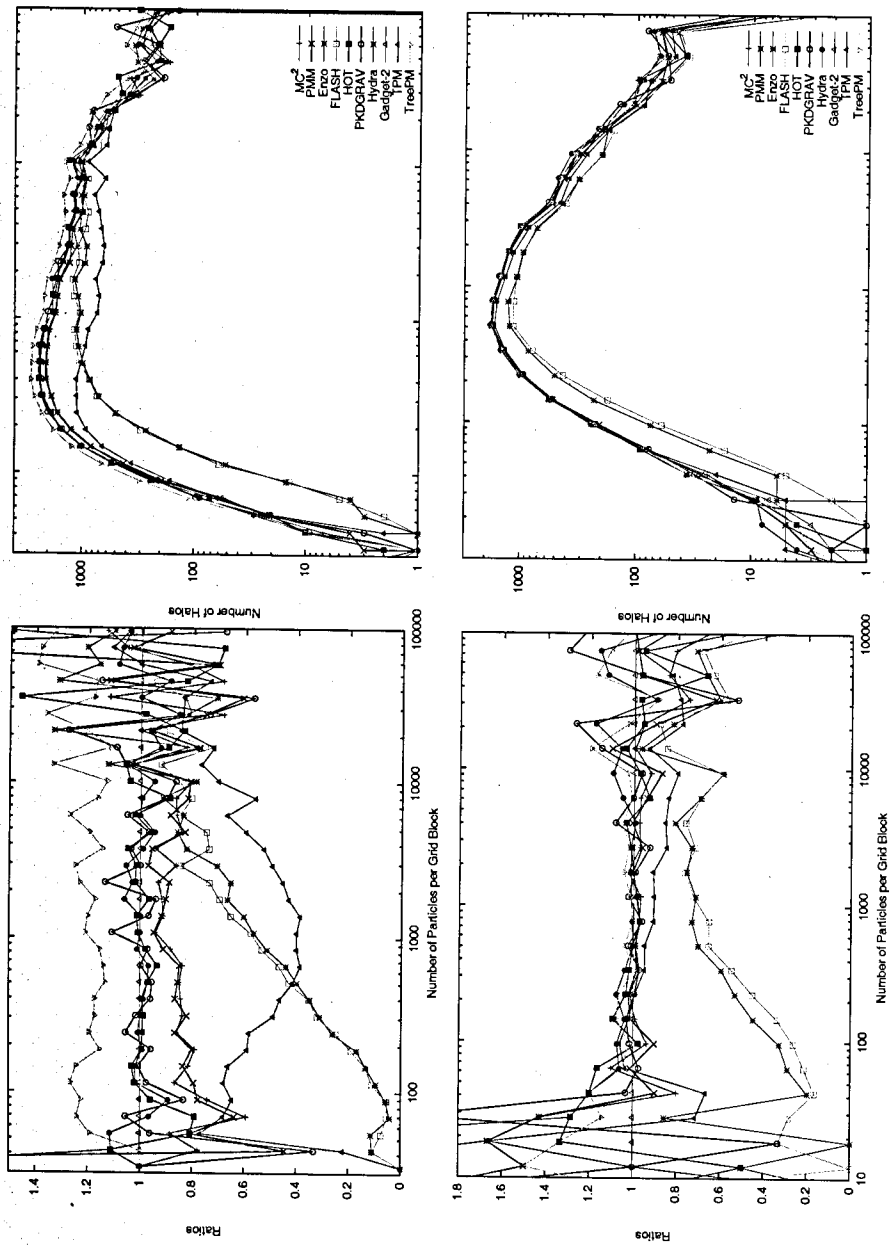

Figure 6.12: Number of halos as a function of density. Left panel: halos with 10 - 40 particles, right panel: halos with 41 - 2500 particles. The lower panels show the residuals with respect to GADGET-2. Both panels show the deficit of small halos in Enzo and FLASH over most of the density region – only at very high densities do the results catch up. The behavior of the TPM simulation is interesting: not only does this simulation have a deficit of small halos but the deficit is very significant in medium density regions, in fact falling below the two AMR codes. The slight excess of small halos shown in the TreePM run vanishes completely if the halo cut is raised to 20 particles per halo and the TreePM results are in that case in excellent agreement with GADGET-2.

earlier, the force resolution of MC^2 and PMM should be sufficient to resolve halos with more than 40 particles, while Enzo's and FLASH's base grid set this limit to more than 2500 particles. We have verified that the codes indeed agree for halos with more than 2500 particles. The lower panels show the residuals with respect to GADGET-2. The two AMR codes Enzo and FLASH have a deficit for both halo sizes over most of the density region. The small halos are mainly missing in the low density regions, and below $\delta^{\rm Void}$ there are almost no halos. Both codes only catch up with the other codes at around 10,000 particles per cell, in agreement with the our previous argument that whether halos are resolvable by the AMR codes or not is dictated by the size of the base grid. In terms of capturing smaller halos, the refinement only helps in very high density regions.

The result for the TPM simulation is somewhat paradoxical: in the low density region the result for the small halos agrees well with the other high-resolution codes, however, TPM misses a very large number of small halos in the region between 200 and 10,000 particles per cell, corresponding to a density contrast δ between 1 and 20. This suggests that the problem of the TPM code is not due to the threshold criterion for the tree but perhaps due to a hand-over problem between the grid and the tree. The two PM codes have slightly lower numbers of very small halos, in good agreement with the prediction that they only resolve halos with more than 40 particles. The agreement between MC^2 and PMM itself is excellent. The TreePM code shows a slight excess of small halos compared to the other high-resolution codes. This excess vanishes completely if the cut for the small halos is chosen to be 20 particles instead of 10 particles for the smallest allowed halo. This indicates a slightly higher force resolution in the TreePM run compared to the other runs. The agreement for the medium size halos (left panel) is very good, except for the AMR codes. For the medium size halos, the TPM code again shows a slight deficit of halos in the medium density regime, but far less pronounced than for the small halos. The overall agreement of the high-resolution codes is very good, as is to be expected from the mass function results.

6.3.4 The Power Spectrum

The matter power spectrum is one of the most important statistics for precision cosmology. Upcoming weak lensing surveys promise measurements of the power spectrum at the one percent accuracy level out to length scales of $k \sim 10h\mathrm{Mpc}^{-1}$ [208]. This poses a severe theoretical challenge: predicting the matter power spectrum at the same level of accuracy. A first step for showing that this is possible is to investigate how well the matter power spectrum can be predicted from pure dark matter simulations, baryonic physics being included as a second step. It has already been shown that at the length scales of interest, hydrodynamic effects can alter the matter power spectrum at up to 10 per-

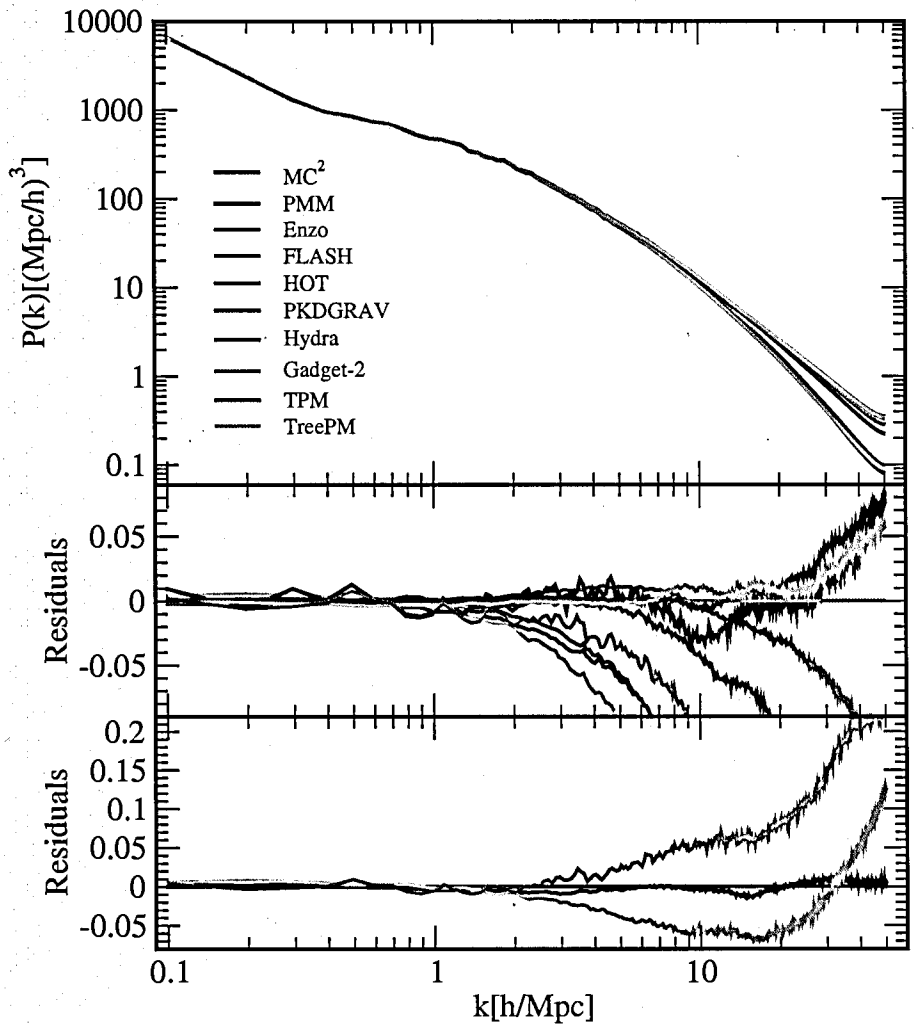

Figure 6.13: Power spectrum results and the residuals for the different codes. Upper panel: comparison of the different power spectra. Middle panel: residuals of all codes with respect to GADGET-2. Lower panel: Residuals of the mesh codes with respect to MC^2.

cent [209, 210, 211, 212]. Here, we concentrate on the first step and determine how well a diverse set of N-body codes agree with each other for the prediction of the matter power spectrum. In future work we aim to predict the dark matter power spectrum at $k \sim 1h\mathrm{Mpc}^{-1}$ at the level of one percent accuracy or better. This will include a detailed analysis of the accuracy of the initial conditions as well as of the nonlinear evolution, a task beyond the scope of the current paper.

We determine the matter power spectrum by generating the density field from the particles via a Cloud-in-Cell (CIC) routine on a 1024^3 spatial grid and then obtain the density in k-space by applying a 1024^3 FFT. The square of the k-space density yields the power spectrum: $P(k) = \langle|\delta(k)^2|\rangle$. The CIC routine introduces a filter at small length scale. We compensate for this filtering artifact by deconvolving the k-space density with a CIC window function.

The results for the different codes are shown in Figure 6.13. Note that the box size of $64h^{-1}$Mpc is too small for a realistic cosmological power spectrum calculation, as the largest modes in the box no longer evolve linearly. This leads to an incorrect onset of the nonlinear turn-over in the power spectrum. Nevertheless, the comparison of the different codes is very informative. The upper panel in Figure 6.13 shows the results for the power spectra themselves. The lower resolution of the grid codes is very apparent, their results falling away at $k \sim 2h\text{Mpc}^{-1}$. The early deviation of the grid codes is still surprising, as the nominal resolution of all codes should have been sufficient to generate agreement over a wider k-range. The middle panel shows the residuals of all codes with respect to GADGET-2. All codes agree at roughly 1% out to $k \sim 1h\text{Mpc}^{-1}$. PKDGRAV shows small scatter in the linear regime. This might be caused by imprecise periodic boundary conditions, which are not as easy to implement in tree codes as they are for grid codes. The high-resolution codes agree to better than 5% out to $k \sim 10h\text{Mpc}^{-1}$. At that point HOT and Hydra lose power, while PKDGRAV, TPM, and TreePM show slightly enhanced power compared to the GADGET-2 run. The formal force resolutions of the codes would suggest that the different runs (including the grid runs) should agree much better at the wavenumbers shown.

The 1024^3 FFT used to generate the power spectra is far below the resolution of the non-grid codes and at the resolution limit of the AMR and PM codes. The discrepancy might be due to several reasons: the number of time steps, the accuracy of the force solvers, the accuracy of reaching $z = 0$ at the end of each run, just to suggest a few. A more detailed study of the power spectrum including larger simulation boxes is certainly required to obtain the desired accuracy for upcoming surveys. In the lower panel we show a comparison of the grid codes only, with respect to MC^2. The two pure PM codes, MC^2 and PMM agree remarkably well over the whole k-range under consideration, the difference being below 1%. The two AMR codes, Flash and Enzo, deviate considerably, most likely due to different refinement criteria. It is somewhat surprising that Enzo has larger power than the two PM codes, which have the same resolution in the whole box that Enzo has only in high density regions. This could be the result of an algorithmic artifact in the AMR implementation.

6.4 Conclusions

The new era of precision cosmology requires new standards for the reach and accuracy of large cosmological simulations. While previously, qualitative answers and quantitative results at the 20% accuracy level were sufficient, we now need to robustly predict nonlinear physics at the 1% accuracy level. This demanding task can only be achieved by rigorous code verification and error control.

The results from the code comparisons are satisfactory and not unexpected, but also show that much more work is needed in order to attain the required

accuracy for upcoming surveys. The halo mass function is a very stable statistic, the agreement over wide ranges of mass being better than 5%. Additionally, the low mass cutoff for individual codes can be reliably predicted by a simple criterion.

The internal structure of halos in the outer regions of $\sim R_{200}$ also appears to be very similar between different simulation codes. Larger differences between the codes in the inner region of the halos occur if the halo is not in a relaxed state: in this case, time stepping issues might also play an important role (e.g. particle orbit phase errors, global time mismatches). For halos with a clear single center, the agreement is very good and predictions for the fall-off of the profiles from resolution criteria hold as expected. The investigation of the halo counts as a function of density revealed an interesting problem with the TPM code, the simulation suffering from a large deficit in medium density regimes. The AMR codes' showed a large deficit of small halos over almost the entire density regime, as the base grid of the AMR simulation set too low a resolution limit for the halos.

The power spectrum measurements revealed definitely more scatter among the different codes than expected. The agreement in the nonlinear regime is at the 5-10% level, even on moderate spatial scales around $k = 10h\mathrm{Mpc}^{-1}$. This disagreement on small scales is connected to differences of the codes in the inner regions of the halos.

7 Concluding Remarks

The area of structure formation is already in the mature phase. Pioneering works of Zel'dovich and Peebles were followed by several other, milestone, results, like the discovery of universal halo density profiles [130, 74], or the universality of the halo mass function [75]. As a result, we have excellent qualitative understanding of the large-scale structure, and its evolution through the gravitational instability. That theoretical picture is in a remarkable agreement with observations. On the other hand, theoretical predictions on most of statistical quantities describing the large-scale structure were not very accurate: different papers would commonly disagree by 50% or more. Partially, it was more difficult to have very precise results then, than it is now due to the constant increase in computer power. Arguably, even more important reason was that great accuracy was previously of only academic interest, as precise results could not be tested on observations.

Observational campaigns in the future (some of which have already started), will achieve a percent-level accuracy on most of cosmological measurables. In order to utilize their power, theoretical predictions of cosmology dependent statistical quantities must have at least the same level of precision. For example, the distribution of masses – mass function is of particular interest, as its high-mass tail is exponentially sensitive to the amplitude of the initial density perturbations, the mean matter density, Ω_m, and to the dark energy controlled late-time evolution of the density field. While this presents a great potential, it cannot become an accurate probe of the cosmological parameters without having good theoretical description of what the mass function is, as a function of cosmological parameters. However, historically, there has been considerable variation in the most interesting, high-mass (low σ) part of the mass function as obtained by different groups. Today, there is a clear need for the very accurate theoretical models describing statistical properties of the large-scale structure.

For that reason, this thesis contains detailed analyses of the most interesting dark matter structures which form through nonlinear gravitational evolution. Specifically, we have presented the most accurate results on the halo mass function and halo bias, covering the largest range of masses and redshifts up to date. The main goal in this research was to represent the findings as much as possible in cosmology and redshift independent ('universal') forms. For that reason, halos are defined and weighted using a percolation motivated friends–of–friends algorithm, which captures an isodensity contour, with which define

halo boundaries.

First, results published in different papers often differ by much more than what is expected from the algorithmic differences in numerical codes. In Chapter 6 we have demonstrated that the agreement on the halo mass function between different codes is as good as is needed today (5-10%), provided the codes evolve the exactly same initial conditions, and that outputs are analyzed with the same tools. In practice, different groups reported results disagreeing by much more than $\sim$10%, the difference increasing with the 'rarity' of objects – equivalent to either increasing the mass of considered halos, or by analyzing them at higher redshifts. For that reason a significant effort went into the study of accuracy requirements in cosmological simulations. Chapter 6 contains comparison of 10 codes most often used nowadays, while in Chapter 3 we derive several criteria which have to be fulfilled to obtain the required accuracy.

Having the confidence in the accuracy of our results, as well as the ability to understand most of the differences in results in the published literature, we have turned to statistical quantities of interest. For the FOF mass function of dark matter halos, we find the universality at the 10% level. Warren et al. fit [8] holds the same level of accuracy. Analysis of the halo correlation function confirms that at large scales the halo correlation is simply in a systematic offset from the matter correlation function. This offset is quantified via 'bias', which increases with halo rarity. Our data, in agreement with results from other groups, show that peak-background split theory, commonly used to relate bias to the mass function fails at $\sim$ 20%. Lacking a theory which enables us to use one formula to describe both statistical measures, we provide an accurate analytical fit to the halo bias itself.

Finally, we analyze halos forming in the ΛCDM cosmogony; we confirm that the halo definition leads to important systematic effects. However, while previous works assumed that the FOF mass corresponds to a fixed spherical overdensity mass (the exact value of 'corresponding' overdensity being varied between $\Delta = 200$ and $\Delta = 100$), we find that the concentration of the density profile plays an important role in mass mapping. Using mock halos, we quantify this concentration effect, and provide a formula for moving from one mass definition to another. We confirm that the formula leads to sensible results when applied on simulated halos, although a selection has to be applied, as a non-negligible fraction of halos is in some stage of merger. The new, interesting result arose from this study: the fraction of merging, or at least non-isolated halos, bears cosmological dependence, and the fraction seems to be universal. This can open new ways to measure growth of structure, or at least can provide a valuable cross-check to other studies. As with the mass function, the most massive structures would be of the biggest interest for this measurement. Although fraction of merging clusters as a function of mass is not as sensitive to cosmological parameters as the mass function, it is a relative probe, thus it would not (directly) suffer from completeness errors. Practical value of this

finding has yet to be tested by observations.

7.1 Future Outlook

In this thesis we analyze several theoretical aspects of growth of structure. Still, it is by no means a final word on that topic. In the following, we present some directions for the future works.

7.1.1 Emulating Halo Statistics

The work presented in this thesis focuses on universal forms of different halo statistics. The reason for this is straightforward: if we have no way of presenting the data in cosmology and redshift independent form, then our statistical measure will be different for different cosmologies. In principle, that means one has to simulate enormous number of cosmologies to make predictions accurate enough for comparison with upcoming observations, as the number of cosmological parameters is relatively large (of order 10). Needless to say, that is not very practical. To circumvent this, we have turned to the FOF definition of halos, which was shown to be universal at the 20% level by Jenkins et al. [75]; in Chapter 3 using more accurate data, we show that the universality of the mass function (and thus e.g. bias) holds to even higher level.

In practice, using this mass definition is very inconvenient for observers for many reasons, the main one being that the hydrostatic equilibrium equation becomes difficult for objects of arbitrary shape. While, as shown in Chapter 4, there are ways of retaining universality with different mass definitions by reducing the cluster sample, it is yet unknown what the accuracy of this method would be. An alternative approach, a natural extension of the work presented here, is to completely abandon the idea of universality, assume the definitions which are observationally the most convenient, and simulate halo statistics employing those definitions. As one can only have a finite number of simulations, this approach would necessarily involve developing an interpolation scheme to probe the parameter space between simulated points. We can call such an interpolating 'device' an *emulator*, and it has to conform to the following three requirements: (i) it must be possible to calibrate it using relatively modest number of expensive N-body simulations (of order 100, at most), (ii) the accuracy of the emulator should be predictable and close to constant over the whole range of interest in cosmological parameter space, and finally (iii) the computational costs of interpolation have to be negligible to that of the N-body simulation itself.

Recent statistical tests [127, 128] on the matter power spectrum show it is possible to build such an emulator. The focus there was on 5 cosmological parameters (n, h, σ_8, $\Omega_{\rm CDM}$, and $\Omega_{\rm b}$), which were kept open in a generous range, much larger than allowed nowadays by many observations. The emulator was

modeled with Gaussian process, using space-filling Latin hypercube designs for 'exact' points, obtained with 128 N-body simulations. The resulting accuracy of emulation (evaluated by comparing emulation results with independent simulations outside the training set) was $\sim$5% throughout most of the k range [128]. Of course, it is possible to achieve even higher accuracy, by reducing the parameter range considered. The reason for such successful modeling is (apart from the fact that cosmological parameters have relatively narrow ranges of interest today) that the response surface (as a function of input parameters) is very smooth; presence of any prominent features would radically increase the number of required N-body simulations.

While the success in modeling the matter power spectrum does not guarantee that any halo statistics can be modelled with the same accuracy, this is most likely the case, as the responce surface seems to be well behaved for the mass function as well as the correlation functions. The possibility of knowing those statistical measures to a great accuracy for any cosmology of interest is very alluring; even though we would still lack 'analytical understanding', its great usability makes this work the natural extension of the research presented in this thesis.

7.1.2 Mass–Observable Relations

As is emphasized throughout the thesis, clusters are a valuable cosmological probe, and arguably the most accurate one which will probe the growth of structures, as opposed to those probing the cosmic expansion. However, in all analyses of galaxy clusters, knowing their masses is essential. While finding masses in simulations, for any desired mass definition, is a straightforward process, on the observational side it is a notoriously difficult problem. Fortunately, cluster masses correlate with many observable quantities, such as temperature, X-ray luminosity, and optical richness [213, 214, 215, 216]. Moreover, properties of central galaxies in clusters and the intracluster light might also be mass proxies [217].

Although the scalings exist, neither systematic nor statistical errors in their determination are at the desired level at present. For example, when hydrostatic equilibrium of clusters is assumed (a standard assumption in observations), the resulting mass as determined from X-ray temperature can be underestimated by $\sim$ 20% [218]. Further, the differences in observational definitions of the temperature bring another systematics at $\sim$ 50% level [219]. On the statistical side, any mass–observable relation exhibits a significant scatter when applied on the whole sample of clusters.

The role of simulation is thus twofold here; on one side simulations can be testbeds for testing different observational strategies, and estimating their accuracy. On the other side, simulations can provide important insight on how to reduce the statistical scatter, possibly by some appropriate selection of clusters,

or by measuring some new quantity (e.g. [220]). In both cases, understanding physical origins of scatter is crucial. Examining the dynamical state, radiative cooling, and heating processes in galaxy clusters is in progress at many research institutions, including the University of Illinois. Better understanding of these effects will open a new avenue in precision cosmology.

A MC^2 Code

MC^2 is the code used for producing most of the simulations analyzed in this thesis. It solves Vlasov-Poisson system of equations (§2.2), discretized through particles. To speed up the force calculation, the code uses the particle-mesh (PM) scheme, in which all space dependent variables (e.g. density, potential) are calculated on an array of mesh points. The differential operators: gradient (∇) and laplacian (∇^2), are expressed as finite-differences on the same mesh.

The density at mesh points is obtained by depositing each particle mass on the surrounding grid nodes using the Cloud-In-Cell (CIC) method where the fraction of the particle mass being deposit on a node depends on the distance between a particle and a nodal point. The CIC interpolation kernel is thus:

$$W = \begin{cases} 1 - \frac{|\mathbf{r}_p - \mathbf{r}_g|}{h} & ; \quad |\mathbf{r}_p - \mathbf{r}_g| \leq h \\ 0 & ; \quad |\mathbf{r}_p - \mathbf{r}_g| > h \end{cases} , \tag{A.1}$$

where $\mathbf{r}_p$ denotes the particles, and $\mathbf{r}_g$ the grid points, while h is the cell width. The density on a mesh is then

$$\rho(n_x, n_y, n_z) = \frac{1}{L^3} \sum_{i=1}^{N_p} m_i W \, . \tag{A.2}$$

The gravitational potential on a mesh can be calculated through

$$\Phi(n_x, n_y, n_z) = \sum_{n'_x, n'_y, n'_z} G(n_x - n'_x, n_y - n'_y, n_z - n'_z) \rho(n'_x, n'_y, n'_z) \, . \tag{A.3}$$

Here $\Phi(n_x, n_y, n_z)$ is the potential at a specified mesh point, and G is the Green's function, whose Fourier transform, for a lattice discretization, is given by:

$$G(\mathbf{k}) = -\frac{3\Omega_m}{8a} \left(\sin^2 \frac{\pi n_x}{L} + \sin^2 \frac{\pi n_y}{L} + \sin^2 \frac{\pi n_z}{L} \right)^{-1} , \tag{A.4}$$

L being the box-size. To find Φ, the code calculates the discrete Fourier transforms of ρ, multiplies them with $G(\mathbf{k})$, and takes the inverse Fourier transform of the product. The gravitational potential at the position of any particular particle is then obtained by interpolation from the mesh.

The main advantage of this method versus direct force summations is the existence of very efficient algorithms for calculating Fourier transforms. Also, Fourier transforms naturally incorporate periodic boundary conditions used in

cosmological simulations. As a result, PM is a lean and very fast method, but with the drawback of having significantly lower force resolution (determined by the mesh spacing) than other methods. This does not pose a problem in simulations where dark matter is considered 'hot' since density fluctuation essentially do not have power on small scales, but may open many issues in cold dark matter simulations where strucuture forms 'bottom-up'. In addition, the speed-up does not come completely for free – the memory requierments with this method are significantly higher than for N^2 solvers, as besides the particle phase-space information and mass, density and potential on the grid have to be stored [1], as well as some auxiliary FFT arrays.

Once forces are calculated, particles can be moved one step in time, and the process repeats. The time integration is done via Leapfrog scheme, where velocity updates are staggered with respect to position updates, and the offset is half of the interval. The integration is therefore:

$$\begin{aligned} \mathbf{r}_{n+1} &= \mathbf{r}_n + \mathbf{v}_{n+1/2}\delta t \\ \mathbf{v}_{n+3/2} &= \mathbf{v}_{n+1/2} - \delta t \nabla \Phi \ . \end{aligned} \tag{A.5}$$

The scheme is second order accurate like, for example, midpoint integration, but the advantage of the Leapfrog algorithm is that it is a symplectic method. This means that the Hamiltonian nature of the equations of motion is preserved, and errors (on any integral of motion) are kept bound.

A.1 Initial Conditions

To set up initial conditions, the code produces a random realization of the desired density field, characterized by its power spectrum $P(\mathbf{k})$. Initial conditions in the code are set using the Zel'dovich [56] approximation:

$$\begin{aligned} \mathbf{r} &= \mathbf{q} - D(a)\mathbf{S}(\mathbf{q}) \\ \mathbf{v} &= -\dot{D}(a - \Delta a/2)\mathbf{S}(\mathbf{q}) \ , \end{aligned} \tag{A.6}$$

where $\mathbf{q}$ is the initial (unperturbed) position of a particle, $\mathbf{x}$ is its position at some time t, D is the linear growth function, and $\mathbf{S}(\mathbf{q})$ is a displacement vector given by the discrete Fourier transform:

$$\mathbf{S}(\mathbf{q}) = A \sum_{k_x,k_y,k_z=-k_{max}}^{k_{max}} i\mathbf{k}c_k e^{i\mathbf{k}\cdot\mathbf{q}}. \tag{A.7}$$

Here, A is the power spectrum normalization coefficient, while c_k is the complex Fourier coefficient, randomly chosen from a Gaussian distribution with the mean of zero, and dispersion $\sigma^2 = P(k)/k^4$. The summation is over all modes which

[1] Actually, one array is enough, since the density can be overwritten by the potential, which is often not done as the code becomes much more difficult to read or change.

can exist in a simulation box, determined for small k's by the physical size of the box, and for large k's by the number of particles[2]. Modes have to be initialized enforcing a 'reality' condition: $c_k = c^*_{-k}$. Thus, cosmological initial conditions reproduce *one* realization of a Gaussian random density field, and the position of any single particle (at any time) is meaningful only in a statistical sense.

[2]Nyquist wavenumber.

Bibliography

[1] M. Tegmark, D. J. Eisenstein, M. A. Strauss, D. H. Weinberg, M. R. Blanton, J. A. Frieman, M. Fukugita, J. E. Gunn, A. J. S. Hamilton, G. R. Knapp, and 57 coauthors. Cosmological constraints from the SDSS luminous red galaxies. *Physical Review D*, 74(12):123507–+, December 2006.

[2] R. K. Sheth and G. Tormen. An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier. *Monthly Notices of Royal Astronomical Society*, 329:61–75, January 2002.

[3] D. Reed, J. Gardner, T. Quinn, J. Stadel, M. Fardal, G. Lake, and F. Governato. Evolution of the mass function of dark matter haloes. *Monthly Notices of Royal Astronomical Society*, 346:565–572, December 2003.

[4] D. S. Reed, R. Bower, C. S. Frenk, A. Jenkins, and T. Theuns. The halo mass function from the dark ages through the present day. *Monthly Notices of Royal Astronomical Society*, 374:2–15, January 2007.

[5] P. Valageas. Transients from Zel'dovich initial conditions. *Astronomy & Astrophysics*, 385:761–767, April 2002.

[6] L. Gao, V. Springel, and S. D. M. White. The age dependence of halo clustering. *Monthly Notices of Royal Astronomical Society*, 363:L66–L70, October 2005.

[7] D. S. Reed, R. Bower, C. S. Frenk, A. Jenkins, and T. Theuns. The clustering of the first galaxies. *ArXiv e-prints*, 804, April 2008.

[8] M. S. Warren, K. Abazajian, D. E. Holz, and L. Teodoro. Precision Determination of the Mass Function of Dark Matter Halos. *Astrophysical Journal*, 646:881–885, August 2006.

[9] A. Friedmann. Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zeitschrift fur Physik*, 21:326–332, December 1924.

[10] A. R. Sandage. Cosmology: a search for two numbers. *Physics Today*, 23:34–41, 1970.

[11] G. Jungman, M. Kamionkowski, and K. Griest. Supersymmetric dark matter. *Physics Reports*, 267:195–373, March 1996.

[12] G. Bertone, D. Hooper, and J. Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405:279–390, January 2005.

[13] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, and 10 coauthors. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astronomical Journal*, 116:1009–1038, September 1998.

[14] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, and 23 coauthors. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *Astrophysical Journal*, 517:565–586, June 1999.

[15] A. Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 142-152.*, pages 142–152, 1917.

[16] A. G. Riess, L.-G. Strolger, J. Tonry, S. Casertano, H. C. Ferguson, B. Mobasher, P. Challis, A. V. Filippenko, S. Jha, W. Li, and 9 coauthors. Type Ia Supernova Discoveries at $z \geq 1$ from the Hubble Space Telescope: Evidence for Past Deceleration and Constraints on Dark Energy Evolution. *Astrophysical Journal*, 607:665–687, June 2004.

[17] R. R. Caldwell. A phantom menace? Cosmological consequences of a dark energy component with super-negative equation of state. *Physics Letters B*, 545:23–29, October 2002.

[18] S. Weinberg. The cosmological constant problem. *Reviews of Modern Physics*, 61:1–23, January 1989.

[19] R. V. Wagoner. Determining q0 from Supernovae. *Astrophysical Journal Letter*, 214:L5+, May 1977.

[20] S. A. Colgate. Supernovae as a standard candle for cosmology. *Astrophysical Journal*, 232:404–408, September 1979.

[21] M. Kowalski, D. Rubin, G. Aldering, R. J. Agostinho, A. Amadon, R. Amanullah, C. Balland, K. Barbary, G. Blanc, P. J. Challis, and 60 coauthors. Improved Cosmological Constraints from New, Old and Combined Supernova Datasets. *ArXiv e-prints*, 804, April 2008.

[22] P. J. E. Peebles and J. T. Yu. Primeval Adiabatic Perturbation in an Expanding Universe. *Astrophysical Journal*, 162:815–+, December 1970.

[23] R. A. Sunyaev and Y. B. Zeldovich. Small-Scale Fluctuations of Relic Radiation. *Astrophysics and Space Science*, 7:3–19, April 1970.

[24] P. Coles and F. Lucchin. *Cosmology. The origin and evolution of cosmic structure.* Chichester: Wiley, —c1995, 1995.

[25] V. Mukhanov. *Physical foundations of cosmology.* Physical foundations of cosmology, by V. Mukhanov. Cambridge, UK: Cambridge University Press, 2005, 2005.

[26] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, and 9 coauthors. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ArXiv e-prints*, 803, March 2008.

[27] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol, R. Scranton, H.-J. Seo, M. Tegmark, Z. Zheng, and 39 coauthors. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophysical Journal*, 633:560–574, November 2005.

[28] M. Bartelmann and P. Schneider. Weak gravitational lensing. *Physics Reports*, 340:291–472, January 2001.

[29] D. Huterer. Weak lensing and dark energy. *Physical Review D*, 65(6):063001–+, March 2002.

[30] M. Takada and B. Jain. Cosmological parameters from lensing power spectrum and bispectrum tomography. *Monthly Notices of Royal Astronomical Society*, 348:897–915, March 2004.

[31] H. Hoekstra and B. Jain. Weak Gravitational Lensing and its Cosmological Applications. *ArXiv e-prints*, 805, May 2008.

[32] G. Holder, Z. Haiman, and J. J. Mohr. Constraints on Ω_m, Ω_Λ, and σ_8 from Galaxy Cluster Redshift Distributions. *Astrophysical Journal Letter*, 560:L111–L114, October 2001.

[33] F. Villa, M. Bersanelli, C. Burigana, R. C. Butler, N. Mandolesi, A. Mennella, G. Morgante, M. Sandri, L. Terenzi, and L. Valenziano. The Planck Telescope. In M. de Petris and M. Gervasi, editors, *Experimental Cosmology at Millimetre Wavelengths*, volume 616 of *American Institute of Physics Conference Series*, pages 224–228, May 2002.

[34] A. Kosowsky. The Atacama Cosmology Telescope. *New Astronomy Review*, 47:939–943, December 2003.

[35] J. Ruhl, P. A. R. Ade, J. E. Carlstrom, H.-M. Cho, T. Crawford, M. Dobbs, C. H. Greer, N. w. Halverson, W. L. Holzapfel, T. M. Lanting, and 17 coauthors. The South Pole Telescope. In C. M. Bradford, P. A. R. Ade, J. E. Aguirre, J. J. Bock, M. Dragovan, L. Duband, L. Earle, J. Glenn, H. Matsuhara, B. J. Naylor, H. T. Nguyen, M. Yun, and J. Zmuidzinas, editors, *Millimeter and Submillimeter Detectors for Astronomy II. Edited by Jonas Zmuidzinas, Wayne S. Holland and Stafford Withington Proceedings of the SPIE, Volume 5498, pp. 11-29 (2004).*, volume 5498 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 11–29, October 2004.

[36] H. J. Mo and S. D. M. White. An analytic model for the spatial clustering of dark matter haloes. *Monthly Notices of Royal Astronomical Society*, 282:347–361, September 1996.

[37] R. K. Sheth and G. Tormen. Large-scale bias and the peak background split. *Monthly Notices of Royal Astronomical Society*, 308:119–126, September 1999.

[38] A. H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23:347–356, January 1981.

[39] J. E. Lidsey, A. R. Liddle, E. W. Kolb, E. J. Copeland, T. Barreiro, and M. Abney. Reconstructing the inflaton potential-an overview. *Reviews of Modern Physics*, 69:373–410, April 1997.

[40] I. J. Grivell and A. R. Liddle. Accurate determination of inflationary perturbations. *Physical Review D*, 54:7191–7198, December 1996.

[41] T. Padmanabhan. *Structure Formation in the Universe.* Structure Formation in the Universe, by T. Padmanabhan, pp. 499. ISBN 0521424860. Cambridge, UK: Cambridge University Press, June 1993., June 1993.

[42] S. W. Hawking. Perturbations of an Expanding Universe. *Astrophysical Journal*, 145:544–+, August 1966.

[43] L. V. Soloveva and A. A. Starobinskii. Gravitational Instability of a Two-Component Medium in an Expanding Universe. *Soviet Astronomy*, 29:367–+, August 1985.

[44] M. Bruni, P. K. S. Dunsby, and G. F. R. Ellis. Cosmological perturbations and the physical meaning of gauge-invariant variables. *Astrophysical Journal*, 395:34–53, August 1992.

[45] P. J. E. Peebles. *The large-scale structure of the universe.* Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p., 1980.

[46] J. Binney and S. Tremaine. *Galactic dynamics.* Princeton, NJ, Princeton University Press, 1987, 747 p., 1987.

[47] L. D. Landau and E. M. Lifshitz. *Fluid mechanics.* Course of theoretical physics, Oxford: Pergamon Press, 1959, 1959.

[48] S. Dodelson. *Modern cosmology.* Modern cosmology / Scott Dodelson. Amsterdam (Netherlands): Academic Press. ISBN 0-12-219141-2, 2003, XIII + 440 p., 2003.

[49] O. Lahav, P. B. Lilje, J. R. Primack, and M. J. Rees. Dynamical effects of the cosmological constant. *Monthly Notices of Royal Astronomical Society*, 251:128–136, July 1991.

[50] V. Sahni and Y. Shtanov. Braneworld models of dark energy. *Journal of Cosmology and Astro-Particle Physics*, 11:14–+, November 2003.

[51] E. V. Linder. Cosmic growth history and expansion history. *Physical Review D*, 72(4):043529–+, August 2005.

[52] E. V. Linder. Redshift distortions as a probe of gravity. *Astroparticle Physics*, 29:336–339, June 2008.

[53] L. Guzzo, M. Pierleoni, B. Meneux, E. Branchini, O. Le Fèvre, C. Marinoni, B. Garilli, J. Blaizot, G. De Lucia, A. Pollo, and 41 coauthors. A test of the nature of cosmic acceleration using galaxy redshift distortions. *Nature*, 451:541–544, January 2008.

[54] R. Juszkiewicz, D. H. Sonoda, and J. D. Barrow. Non-linear gravitational clustering. *Monthly Notices of Royal Astronomical Society*, 209:139–144, July 1984.

[55] V. Sahni and P. Coles. Approximation methods for non-linear gravitational clustering. *Physics Reports*, 262:1–135, November 1995.

[56] Y. B. Zel'Dovich. Gravitational instability: An approximate theory for large density perturbations. *Astronomy & Astrophysics*, 5:84–89, March 1970.

[57] A. G. Doroshkevich. The space structure of perturbations and the origin of rotation of galaxies in the theory of fluctuation. *Astrofizika*, 6:581–600, 1970.

[58] A. G. Doroshkevich. Spatial structure of perturbations and origin of galactic rotation in fluctuation theory. *Astrophysics*, 6:320–330, October 1970.

[59] C. C. Lin, L. Mestel, and F. H. Shu. The Gravitational Collapse of a Uniform Spheroid. *Astrophysical Journal*, 142:1431–+, November 1965.

[60] L. Kofman, E. Bertschinger, J. M. Gelb, A. Nusser, and A. Dekel. Evolution of one-point distributions from Gaussian initial fluctuations. *Astrophysical Journal*, 420:44–57, January 1994.

[61] F. Bernardeau and L. Kofman. Properties of the cosmological density distribution function. *Astrophysical Journal*, 443:479–498, April 1995.

[62] S. F. Shandarin and Y. B. Zeldovich. The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium. *Reviews of Modern Physics*, 61:185–220, April 1989.

[63] H. Bondi. Spherically symmetrical models in general relativity. *Monthly Notices of Royal Astronomical Society*, 107:410–+, 1947.

[64] J. E. Gunn and J. R. I. Gott. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *Astrophysical Journal*, 176:1–+, August 1972.

[65] T. Padmanabhan. Cosmological constant-the weight of the vacuum. *Physics Reports*, 380:235–320, July 2003.

[66] P. J. E. Peebles. The Gravitational Instability of the Universe. *Astrophysical Journal*, 147:859–+, March 1967.

[67] G. L. Bryan and M. L. Norman. Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. *Astrophysical Journal*, 495:80–+, March 1998.

[68] R. Schaeffer, J. Silk, M. Spiro, and J. Zinn-Justin, editors. *Cosmology and large scale structure*, January 1996.

[69] Z. Haiman and A. Loeb. What Is the Highest Plausible Redshift of Luminous Quasars? *Astrophysical Journal*, 552:459–463, May 2001.

[70] S. R. Furlanetto, M. McQuinn, and L. Hernquist. Characteristic scales during reionization. *Monthly Notices of Royal Astronomical Society*, 365:115–126, January 2006.

[71] W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *Astrophysical Journal*, 187:425–438, February 1974.

[72] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion set mass functions for hierarchical Gaussian fluctuations. *Astrophysical Journal*, 379:440–460, October 1991.

[73] K. Heitmann, Z. Lukić, S. Habib, and P. M. Ricker. Capturing Halos at High Redshifts. *Astrophysical Journal Letter*, 642:L85–L88, May 2006.

[74] J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. *Astrophysical Journal*, 490:493–+, December 1997.

[75] A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida. The mass function of dark matter haloes. *Monthly Notices of Royal Astronomical Society*, 321:372–384, February 2001.

[76] R. Barkana and A. Loeb. Unusually Large Fluctuations in the Statistics of Galaxy Formation at High Redshift. *Astrophysical Journal*, 609:474–481, July 2004.

[77] G. M. Voit. Tracing cosmic evolution with clusters of galaxies. *Reviews of Modern Physics*, 77:207–258, April 2005.

[78] M. White. The mass of a halo. *Astronomy & Astrophysics*, 367:27–32, February 2001.

[79] J. Einasto, A. A. Klypin, E. Saar, and S. F. Shandarin. Structure of superclusters and supercluster formation. III Quantitative study of the local supercluster. *Monthly Notices of Royal Astronomical Society*, 206:529–558, February 1984.

[80] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophysical Journal*, 292:371–394, May 1985.

[81] M. White. The Mass Function. *Astrophysical Journal Supplement Series*, 143:241–255, December 2002.

[82] W. Hu and N. Sugiyama. Small-Scale Cosmological Perturbations: an Analytic Approach. *Astrophysical Journal*, 471:542–+, November 1996.

[83] N. Yoshida, N. Sugiyama, and L. Hernquist. The evolution of baryon density fluctuations in multicomponent cosmological simulations. *Monthly Notices of Royal Astronomical Society*, 344:481–491, September 2003.

[84] S. Naoz and R. Barkana. The formation and gas content of high-redshift galaxies and minihaloes. *Monthly Notices of Royal Astronomical Society*, 377:667–676, May 2007.

[85] C. Lacey and S. Cole. Merger rates in hierarchical models of galaxy formation. *Monthly Notices of Royal Astronomical Society*, 262:627–649, June 1993.

[86] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, and 7 coauthors. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636, June 2005.

[87] R. K. Sheth, H. J. Mo, and G. Tormen. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *Monthly Notices of Royal Astronomical Society*, 323:1–12, May 2001.

[88] J. Lee and S. F. Shandarin. The Cosmological Mass Distribution Function in the Zeldovich Approximation. *Astrophysical Journal*, 500:14–+, June 1998.

[89] B. W. O'Shea, K. Nagamine, V. Springel, L. Hernquist, and M. L. Norman. Comparing AMR and SPH Cosmological Simulations. I. Dark Matter and Adiabatic Simulations. *Astrophysical Journal Supplement Series*, 160:1–27, September 2005.

[90] K. Heitmann, P. M. Ricker, M. S. Warren, and S. Habib. Robustness of Cosmological Simulations. I. Large-Scale Structure. *Astrophysical Journal Supplement Series*, 160:28–58, September 2005.

[91] K. Heitmann, Z. Lukic, P. Fasel, S. Habib, M. S. Warren, M. White, J. Ahrens, L. Ankeny, R. Armstrong, B. O'Shea, P. M. Ricker, V. Springel, J. Stadel, and H. Trac. The Cosmic Code Comparison Project. *ArXiv e-prints*, 706, June 2007.

[92] O. Le Fèvre, S. Paltani, S. Arnouts, S. Charlot, S. Foucaud, O. Ilbert, H. J. McCracken, G. Zamorani, D. Bottini, B. Garilli, and 39 coauthors. A large population of galaxies 9 to 12 billion years back in the history of the Universe. *Nature*, 437:519–521, September 2005.

[93] B. Mobasher, M. Dickinson, H. C. Ferguson, M. Giavalisco, T. Wiklind, D. Stark, R. S. Ellis, S. M. Fall, N. A. Grogin, L. A. Moustakas, N. Panagia, M. Sosey, M. Stiavelli, E. Bergeron, S. Casertano, P. Ingraham, A. Koekemoer, I. Labbé, M. Livio, B. Rodgers, C. Scarlata, J. Vernet, A. Renzini, P. Rosati, H. Kuntschner, M. Kümmel, J. R. Walsh, R. Chary, P. Eisenhardt, N. Pirzkal, and D. Stern. Evidence for a Massive Poststarburst Galaxy at z~6.5. *Astrophysical Journal*, 635:832–844, December 2005.

[94] J. D. Bowman, M. F. Morales, and J. N. Hewitt. The Sensitivity of First-Generation Epoch of Reionization Observatories and Their Potential for Differentiating Theoretical Power Spectra. *Astrophysical Journal*, 638:20–26, February 2006.

[95] H. Jang-Condell and L. Hernquist. First Structure Formation: A Simulation of Small-Scale Structure at High Redshift. *Astrophysical Journal*, 548:68–78, February 2001.

[96] V. Springel. The cosmological simulation code GADGET-2. *Monthly Notices of Royal Astronomical Society*, 364:1105–1134, December 2005.

[97] C. M. Baugh, E. Gaztanaga, and G. Efstathiou. A comparison of the evolution of density fields in perturbation theory and numerical simulations - II. Counts-in-cells analysis. *Monthly Notices of Royal Astronomical Society*, 274:1049–1070, June 1995.

[98] S. D. M. White. Formation and Evolution of Galaxies. In R. Schaeffer, J. Silk, M. Spiro, and J. Zinn-Justin, editors, *Cosmology and Large Scale Structure*, pages 349–+, January 1996.

[99] N. Yoshida, T. Abel, L. Hernquist, and N. Sugiyama. Simulations of Early Structure Formation: Primordial Gas Clouds. *Astrophysical Journal*, 592:645–663, August 2003.

[100] N. Yoshida, A. Sokasian, L. Hernquist, and V. Springel. Early Structure Formation and Reionization in a Warm Dark Matter Cosmology. *Astrophysical Journal Letter*, 591:L1–L4, July 2003.

[101] N. Yoshida, A. Sokasian, L. Hernquist, and V. Springel. Early Structure Formation and Reionization in a Cosmological Model with a Running Primordial Power Spectrum. *Astrophysical Journal*, 598:73–85, November 2003.

[102] U. Maio, K. Dolag, M. Meneghetti, L. Moscardini, N. Yoshida, C. Baccigalupi, M. Bartelmann, and F. Perrotta. Early structure formation in quintessence models and its implications for cosmic reionization from first stars. *Monthly Notices of Royal Astronomical Society*, 373:869–878, December 2006.

[103] R. Cen, F. Dong, P. Bode, and J. P. Ostriker. Properties of Cold Dark Matter Halos at z¿6. *ArXiv Astrophysics e-prints*, March 2004.

[104] G. Xu. A New Parallel N-Body Gravity Solver: TPM. *Astrophysical Journal Supplement Series*, 98:355–+, May 1995.

[105] P. Bode, J. P. Ostriker, and G. Xu. The Tree Particle-Mesh N-Body Gravity Solver. *Astrophysical Journal Supplement Series*, 128:561–569, June 2000.

[106] E. Bertschinger and J. M. Gelb. Cosmological N-body simulations. *Computers in Physics*, 5:164–175, April 1991.

[107] I. T. Iliev, G. Mellema, U.-L. Pen, H. Merz, P. R. Shapiro, and M. A. Alvarez. Simulating cosmic reionization at large scales - I. The geometry of reionization. *Monthly Notices of Royal Astronomical Society*, 369:1625–1638, July 2006.

[108] H. Merz, U.-L. Pen, and H. Trac. Towards optimal parallel PM N-body codes: PMFAST. *New Astronomy*, 10:393–407, April 2005.

[109] O. Zahn, A. Lidz, M. McQuinn, S. Dutta, L. Hernquist, M. Zaldarriaga, and S. R. Furlanetto. Simulations and Analytic Calculations of Bubble Growth during Hydrogen Reionization. *Astrophysical Journal*, 654:12–26, January 2007.

[110] H. Trac and R. Cen. Radiative Transfer Simulations of Cosmic Reionization. I. Methodology and Initial Results. *Astrophysical Journal*, 671:1–13, December 2007.

[111] C. J. MacTavish, P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, A. Boscaleri, P. Cabella, C. R. Contaldi, B. P. Crill, P. de Bernardis, G. De Gasperis, A. de Oliveira-Costa, G. De Troia, G. di Stefano, E. Hivon, A. H. Jaffe, W. C. Jones, T. S. Kisner, A. E. Lange, A. M. Lewis, S. Masi, P. D. Mauskopf, A. Melchiorri, T. E. Montroy, P. Natoli, C. B. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, G. Polenta, S. Prunet, S. Ricciardi, G. Romeo, J. E. Ruhl, P. Santini, M. Tegmark, M. Veneziani, and N. Vittorio. Cosmological Parameters from the 2003 Flight of BOOMERANG. *Astrophysical Journal*, 647:799–812, August 2006.

[112] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. *Astrophysical Journal Supplement Series*, 170:377–408, June 2007.

[113] U. Seljak and M. Zaldarriaga. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *Astrophysical Journal*, 469:437–+, October 1996.

[114] J. M. Gelb and E. Bertschinger. Cold dark matter. 1: The formation of dark halos. *Astrophysical Journal*, 436:467–490, December 1994.

[115] F. J. Summers, M. Davis, and A. E. Evrard. Galaxy Tracers and Velocity Bias. *Astrophysical Journal*, 454:1–+, November 1995.

[116] D. S. Reed, R. Bower, C. S. Frenk, L. Gao, A. Jenkins, T. Theuns, and S. D. M. White. The first generation of star-forming haloes. *Monthly Notices of Royal Astronomical Society*, 363:393–404, October 2005.

[117] E. F. Bunn and M. White. The 4 Year COBE Normalization and Large-Scale Structure. *Astrophysical Journal*, 480:6–+, May 1997.

[118] M. Crocce, S. Pueblas, and R. Scoccimarro. Transients from initial conditions in cosmological simulations. *Monthly Notices of Royal Astronomical Society*, 373:369–381, November 2006.

[119] A. E. Evrard. Galaxy Clusters as Probes of Cosmology and Astrophysics. In J. S. Mulchaey, A. Dressler, and A. Oemler, editors, *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution*, pages 1–+, 2004.

[120] A. E. Evrard, J. Bialek, M. Busha, M. White, S. Habib, K. Heitmann, M. Warren, E. Rasia, G. Tormen, L. Moscardini, C. Power, A. R. Jenkins, L. Gao, C. S. Frenk, V. Springel, S. D. M. White, and J. Diemand. Virial Scaling of Massive Dark Matter Halos: Why Clusters Prefer a High Normalization Cosmology. *Astrophysical Journal*, 672:122–137, January 2008.

[121] J. Heinrich. Coverage of Error Bars for Poisson Data. *CDF Note*, 6438, May 2003.

[122] M. Tegmark, J. Silk, M. J. Rees, A. Blanchard, T. Abel, and F. Palla. How Small Were the First Cosmological Objects? *Astrophysical Journal*, 474:1–+, January 1997.

[123] N. Katz, T. Quinn, and J. M. Gelb. Galaxy Formation and the Peaks Formalism. *Monthly Notices of Royal Astronomical Society*, 265:689–+, December 1993.

[124] J. S. Bagla and J. Prasad. Effects of the size of cosmological N-body simulations on physical quantities - I. Mass function. *Monthly Notices of Royal Astronomical Society*, 370:993–1002, August 2006.

[125] J. E. Betancort-Rijo and A. D. Montero-Dorta. The Halo Mass Function Redshift Dependence. *Astrophysical Journal Letter*, 653:L77–L80, December 2006.

[126] J. E. Betancort-Rijo and A. D. Montero-Dorta. Understanding the Cosmic Mass Function High-Mass Behavior. *Astrophysical Journal Letter*, 650:L95–L98, October 2006.

[127] K. Heitmann, D. Higdon, C. Nakhleh, and S. Habib. Cosmic Calibration. *Astrophysical Journal Letter*, 646:L1–L4, July 2006.

[128] S. Habib, K. Heitmann, D. Higdon, C. Nakhleh, and B. Williams. Cosmic calibration: Constraints from the matter power spectrum and the cosmic microwave background. *Physical Review D*, 76(8):083503–+, October 2007.

[129] A. Cooray and R. Sheth. Halo models of large scale structure. *Physics Reports*, 372:1–129, December 2002.

[130] J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. *Astrophysical Journal*, 462:563–+, May 1996.

[131] Z. Lukić, K. Heitmann, S. Habib, S. Bashinsky, and P. M. Ricker. The Halo Mass Function: High-Redshift Evolution and Universality. *Astrophysical Journal*, 671:1160–1181, December 2007.

[132] A. E. Evrard, T. J. MacFarland, H. M. P. Couchman, J. M. Colberg, N. Yoshida, S. D. M. White, A. Jenkins, C. S. Frenk, F. R. Pearce, J. A. Peacock, and P. A. Thomas. Galaxy Clusters in Hubble Volume Simulations: Cosmological Constraints from Sky Survey Populations. *Astrophysical Journal*, 573:7–36, July 2002.

[133] J. L Tinker, A. V Kravtsov, A. Klypin, K. Abazajian, M. S Warren, G. Yepes, S. Gottlober, and D. E Holz. Toward a halo mass function for precision cosmology: the limits of universality. *ArXiv e-prints*, 803, March 2008.

[134] A. E. Evrard, J. Bialek, M. Busha, M. White, S. Habib, K. Heitmann, M. Warren, E. Rasia, G. Tormen, L. Moscardini, C. Power, A. R. Jenkins, L. Gao, C. S. Frenk, V. Springel, S. D. M. White, and J. Diemand. Virial Scaling of Massive Dark Matter Halos: Why Clusters Prefer a High Normalization Cosmology. *Astrophysical Journal*, 672:122–137, January 2008.

[135] C. S. Frenk, S. D. M. White, M. Davis, and G. Efstathiou. The formation of dark halos in a universe dominated by cold dark matter. *Astrophysical Journal*, 327:507–525, April 1988.

[136] S. Ghigna, B. Moore, F. Governato, G. Lake, T. Quinn, and J. Stadel. Density Profiles and Substructure of Dark Matter Halos: Converging Results at Ultra-High Numerical Resolution. *Astrophysical Journal*, 544:616–628, December 2000.

[137] Y. P. Jing and Y. Suto. The Density Profiles of the Dark Matter Halo Are Not Universal. *Astrophysical Journal Letter*, 529:L69–L72, February 2000.

[138] A. Klypin, A. V. Kravtsov, J. S. Bullock, and J. R. Primack. Resolving the Structure of Cold Dark Matter Halos. *Astrophysical Journal*, 554:903–915, June 2001.

[139] J. F. Navarro, E. Hayashi, C. Power, A. R. Jenkins, C. S. Frenk, S. D. M. White, V. Springel, J. Stadel, and T. R. Quinn. The inner structure of ΛCDM haloes - III. Universality and asymptotic slopes. *Monthly Notices of Royal Astronomical Society*, 349:1039–1051, April 2004.

[140] D. Reed, F. Governato, L. Verde, J. Gardner, T. Quinn, J. Stadel, D. Merritt, and G. Lake. Evolution of the density profiles of dark matter haloes. *Monthly Notices of Royal Astronomical Society*, 357:82–96, February 2005.

[141] J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes: evolution, scatter and environment. *Monthly Notices of Royal Astronomical Society*, 321:559–575, March 2001.

[142] V. R. Eke, J. F. Navarro, and M. Steinmetz. The Power Spectrum Dependence of Dark Matter Halo Concentrations. *Astrophysical Journal*, 554:114–125, June 2001.

[143] S. F. Kasun and A. E. Evrard. Shapes and Alignments of Galaxy Cluster Halos. *Astrophysical Journal*, 629:781–790, August 2005.

[144] B. Allgood, R. A. Flores, J. R. Primack, A. V. Kravtsov, R. H. Wechsler, A. Faltenbacher, and J. S. Bullock. The shape of dark matter haloes: dependence on mass, redshift, radius and formation. *Monthly Notices of Royal Astronomical Society*, 367:1781–1796, April 2006.

[145] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay. The statistics of peaks of Gaussian random fields. *Astrophysical Journal*, 304:15–61, May 1986.

[146] C. Power, J. F. Navarro, A. Jenkins, C. S. Frenk, S. D. M. White, V. Springel, J. Stadel, and T. Quinn. The inner structure of ΛCDM haloes - I. A numerical convergence study. *Monthly Notices of Royal Astronomical Society*, 338:14–34, January 2003.

[147] A. F. Neto, L. Gao, P. Bett, S. Cole, J. F. Navarro, C. S. Frenk, S. D. M. White, V. Springel, and A. Jenkins. The statistics of Λ CDM halo concentrations. *Monthly Notices of Royal Astronomical Society*, 381:1450–1462, November 2007.

[148] Y. P. Jing. The Density Profile of Equilibrium and Nonequilibrium Dark Matter Halos. *Astrophysical Journal*, 535:30–36, May 2000.

[149] A. V. Macciò, A. A. Dutton, F. C. van den Bosch, B. Moore, D. Potter, and J. Stadel. Concentration, spin and shape of dark matter haloes: scatter and the dependence on mass and environment. *Monthly Notices of Royal Astronomical Society*, 378:55–71, June 2007.

[150] M. T. Busha, A. E. Evrard, F. C. Adams, and R. H. Wechsler. The ultimate halo mass in a ΛCDM universe. *Monthly Notices of Royal Astronomical Society*, 363:L11–L15, October 2005.

[151] S. P. D. Gill, A. Knebe, and B. K. Gibson. The evolution of substructure - III. The outskirts of clusters. *Monthly Notices of Royal Astronomical Society*, 356:1327–1332, February 2005.

[152] A. D. Ludlow, J. F. Navarro, V. Springel, A. Jenkins, C. S. Frenk, and A. Helmi. The Unorthodox Orbits of Substructure Halos. *ArXiv e-prints*, 801, January 2008.

[153] S. L. W. McMillan, M. P. Kowalski, and M. P. Ulmer. X-ray morphologies of Abell clusters. *Astrophysical Journal Supplement Series*, 70:723–730, August 1989.

[154] J. J. Mohr, A. E. Evrard, D. G. Fabricant, and M. J. Geller. Cosmological Constraints from Observed Cluster X-Ray Morphologies. *Astrophysical Journal*, 447:8–+, July 1995.

[155] A. I. Zabludoff and D. Zaritsky. A Collision of Subclusters in Abell 754. *Astrophysical Journal Letter*, 447:L21+, July 1995.

[156] C. Jones and W. Forman. Einstein Observatory Images of Clusters of Galaxies. *Astrophysical Journal*, 511:65–83, January 1999.

[157] V. Kolokotronis, S. Basilakos, M. Plionis, and I. Georgantopoulos. Searching for cluster substructure using APM and ROSAT data. *Monthly Notices of Royal Astronomical Society*, 320:49–60, January 2001.

[158] T. E. Jeltema, C. R. Canizares, M. W. Bautz, and D. A. Buote. The Evolution of Structure in X-Ray Clusters of Galaxies. *Astrophysical Journal*, 624:606–629, May 2005.

[159] M. Ramella, A. Biviano, A. Pisani, J. Varela, D. Bettoni, W. J. Couch, M. D'Onofrio, A. Dressler, G. Fasano, P. Kjørgaard, M. Moles, E. Pignatelli, and B. M. Poggianti. Substructures in WINGS clusters. *Astronomy & Astrophysics*, 470:39–51, July 2007.

[160] A. A. Berlind and D. H. Weinberg. The Halo Occupation Distribution: Toward an Empirical Determination of the Relation between Galaxies and Mass. *Astrophysical Journal*, 575:587–616, August 2002.

[161] J. L. Tinker, D. H. Weinberg, Z. Zheng, and I. Zehavi. On the Mass-to-Light Ratio of Large-Scale Structure. *Astrophysical Journal*, 631:41–58, September 2005.

[162] A. E. Schulz and M. White. Surveys of Galaxy Clusters with the Sunyaev-Zel'dovich Effect. *Astrophysical Journal*, 586:723–730, April 2003.

[163] J. R. Bond, L. Kofman, and D. Pogosyan. How filaments of galaxies are woven into the cosmic web. *Nature*, 380:603–+, April 1996.

[164] M. Tegmark, M. R. Blanton, M. A. Strauss, F. Hoyle, D. Schlegel, R. Scoccimarro, M. S. Vogeley, D. H. Weinberg, I. Zehavi, A. Berlind, and 55 coauthors. The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey. *Astrophysical Journal*, 606:702–740, May 2004.

[165] J. Estrada, E. Sefusatti, and J. A. Frieman. The Correlation Function of Optically Selected Galaxy Clusters in the Sloan Digital Sky Survey. *ArXiv e-prints*, 801, January 2008.

[166] U. Seljak and M. S. Warren. Large-scale bias and stochasticity of haloes and dark matter. *Monthly Notices of Royal Astronomical Society*, 355:129–136, November 2004.

[167] S. Majumdar and J. J. Mohr. Self-Calibration in Cluster Studies of Dark Energy: Combining the Cluster Redshift Distribution, the Power Spectrum, and Mass Measurements. *Astrophysical Journal*, 613:41–50, September 2004.

[168] M. Lima and W. Hu. Self-calibration of cluster dark energy studies: Counts in cells. *Physical Review D*, 70(4):043504–+, August 2004.

[169] M. Lima and W. Hu. Self-calibration of cluster dark energy studies: Observable-mass distribution. *Physical Review D*, 72(4):043006–+, August 2005.

[170] H.-Y. Wu, E. Rozo, and R. H. Wechsler. The Effect of Halo Assembly Bias on Self Calibration in Galaxy Cluster Surveys. *ArXiv e-prints*, 803, March 2008.

[171] M. Davis and P. J. E. Peebles. On the integration of the BBGKY equations for the development of strongly nonlinear clustering in an expanding universe. *Astrophysical Journal Supplement Series*, 34:425–450, August 1977.

[172] N. Dalal, M. White, J. R. Bond, and A. Shirokov. Halo Assembly Bias in Hierarchical Structure Formation. *ArXiv e-prints*, 803, March 2008.

[173] S. Cole and N. Kaiser. Biased clustering in the cold dark matter cosmogony. *Monthly Notices of Royal Astronomical Society*, 237:1127–1146, April 1989.

[174] T. Hamana, N. Yoshida, Y. Suto, and A. E. Evrard. Clustering of Dark Matter Halos on the Light Cone: Scale, Time, and Mass Dependence of the Halo Biasing in the Hubble Volume Simulations. *Astrophysical Journal Letter*, 561:L143–L146, November 2001.

[175] A. Diaferio, A. Nusser, N. Yoshida, and R. A. Sunyaev. Superclusters with thermal Sunyaev-Zel'dovich effect surveys. *Monthly Notices of Royal Astronomical Society*, 338:433–442, January 2003.

[176] J. D. Cohn and M. White. Dark matter halo abundances, clustering and assembly histories at high redshift. *Monthly Notices of Royal Astronomical Society*, 385:2025–2033, April 2008.

[177] R.D. Ryne, S. Habib, J. Qiang, K. Ko, Z. Li, B. McCandless, W. Mi, C.-K. Ng, and 7 coauthors. The U.S. DOE Grand Challenge in Computational Accelerator Physics. In *XIX International Linear Accelerator Conference*, pages 701–706, 1998.

[178] J. Qiang, R. D. Ryne, S. Habib, and V. Decyk. An Object-Oriented Parallel Particle-in-Cell Code for Beam Dynamics Simulation in Linear Accelerators. *Journal of Computational Physics*, 163:434–451, September 2000.

[179] H. Trac and U.-L. Pen. Out-of-core hydrodynamic simulations for cosmological applications. *New Astronomy*, 11:273–286, January 2006.

[180] G. L. Bryan and M. L. Norman. A Hybrid AMR Application for Cosmology and Astrophysics. *ArXiv Astrophysics e-prints*, October 1997.

[181] B. W. O'Shea, G. Bryan, J. Bordner, M. L. Norman, T. Abel, R. Harkness, and A. Kritsuk. Introducing Enzo, an AMR Cosmology Application. In T. Plewa, T. Linde, and G. Weirs, editors, *Adaptive Mesh Refinement - Theory and Applications*, pages 341–351, 2004.

[182] M. J. Berger and P. Colella. Local adaptive mesh refinement for shock hydrodynamics. *Journal of Computational Physics*, 82:64–84, May 1989.

[183] G. Efstathiou, M. Davis, S. D. M. White, and C. S. Frenk. Numerical techniques for large cosmological N-body simulations. *Astrophysical Journal Supplement Series*, 57:241–260, February 1985.

[184] R. W. Hockney and J. W. Eastwood. *Computer simulation using particles.* Bristol: Hilger, 1988, 1988.

[185] P. Colella and P. R. Woodward. The Piecewise Parabolic Method (PPM) for Gas-Dynamical Simulations. *Journal of Computational Physics*, 54:174–201, September 1984.

[186] G. L. Bryan, M. L. Norman, J. M. Stone, R. Cen, and J. P. Ostriker. A piecewise parabolic method for cosmological hydrodynamics. *Computer Physics Communications*, 89:149–168, August 1995.

[187] B. Fryxell, K. Olson, P. Ricker, F. X. Timmes, M. Zingale, D. Q. Lamb, P. MacNeice, R. Rosner, J. W. Truran, and H. Tufo. FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes. *Astrophysical Journal Supplement Series*, 131:273–334, November 2000.

[188] P. MacNeice, K. M. Olson, C. Mobarry, R. de Fainchtein, and C. Packer. PARAMESH: A parallel adaptive mesh refinement community toolkit. *Computer Physics Communications*, 126:330–354, April 2000.

[189] J. J. Quirk. PhD thesis, , Cranfield Institute of Technology, (1991), 1991.

[190] D. Dezeeuw and K. G. Powell. An Adaptively Refined Cartesian Mesh Solver for the Euler Equations. *Journal of Computational Physics*, 104:56–68, January 1993.

[191] M. S. Warren and J. K. Salmon. A Parallel Hashed Oct-Tree N-Body Algorithm. In *Supercomputing 1993*, pages 1–+, 1993.

[192] J. K. Salmon and M. S. Warren. Skeletons from the treecode closet. *Journal of Computational Physics*, 111:136–155, March 1994.

[193] M. S. Warren, C. L. Fryer, and M. P. Goda. The Space Simulator: Modeling the Universe from Supernovae to Cosmology. In *Proceedings of the ACM/IEEE SC2003 Conference*, pages 30–, 2003.

[194] C. S. Frenk, S. D. M. White, P. Bode, J. R. Bond, G. L. Bryan, R. Cen, H. M. P. Couchman, A. E. Evrard, N. Gnedin, A. Jenkins, and 15 coauthors. The Santa Barbara Cluster Comparison Project: A Comparison of Cosmological Hydrodynamics Solutions. *Astrophysical Journal*, 525:554–582, November 1999.

[195] C. L. Fryer and M. S. Warren. Modeling Core-Collapse Supernovae in Three Dimensions. *Astrophysical Journal Letter*, 574:L65–L68, July 2002.

[196] M. D. Dikaiakos and J. Stadel. A performance study of cosmological simulations on message-passing and shared-memory multiprocessors. In *Proceedings of the 10th international conference on Supercomputing*, pages 94–101, 1996.

[197] J. G. Stadel. *Cosmological N-body simulations and their analysis.* PhD thesis, AA(UNIVERSITY OF WASHINGTON), 2001.

[198] J. Barnes and P. Hut. A Hierarchical O(NlogN) Force-Calculation Algorithm. *Nature*, 324:446–449, December 1986.

[199] L. Hernquist, F. R. Bouchet, and Y. Suto. Application of the Ewald method to cosmological N-body simulations. *Astrophysical Journal Supplement Series*, 75:231–240, February 1991.

[200] H. M. P. Couchman. Simulating the formation of large-scale cosmic structure with particle-grid methods. *Journal of Computational and Applied Mathematics*, 109:373–406, September 1999.

[201] V. Springel, N. Yoshida, and S. D. M. White. GADGET: a code for collisionless and gasdynamical cosmological simulations. *New Astronomy*, 6:79–117, April 2001.

[202] J. S. Bagla. TreePM: A Code for Cosmological N-Body Simulations. *Journal of Astrophysics and Astronomy*, 23:185–196, December 2002.

[203] J. J. Monaghan. Smoothed particle hydrodynamics. *Annual Review of Astronomy and Astrophysics*, 30:543–574, 1992.

[204] V. Springel and L. Hernquist. Cosmological smoothed particle hydrodynamics simulations: the entropy equation. *Monthly Notices of Royal Astronomical Society*, 333:649–664, July 2002.

[205] A. Klypin and J. Holtzman. Particle-Mesh code for cosmological simulations. *ArXiv Astrophysics e-prints*, December 1997.

[206] W. Dehnen. Towards optimal softening in three-dimensional N-body codes - I. Minimizing the force error. *Monthly Notices of Royal Astronomical Society*, 324:273–291, June 2001.

[207] J. Ahrens, B. Geveci, and Law C. ParaView: An End-User Tool for Large Data Visualization. In C. Hansen and C. Johnson, editors, *The Visualization Handbook*, pages 717–731, 2005.

[208] D. Huterer and M. Takada. Calibrating the nonlinear matter power spectrum: Requirements for future weak lensing surveys. *Astroparticle Physics*, 23:369–376, May 2005.

[209] M. White. Baryons and weak lensing power spectra. *Astroparticle Physics*, 22:211–217, November 2004.

[210] H. Zhan and L. Knox. Effect of Hot Baryons on the Weak-Lensing Shear Power Spectrum. *Astrophysical Journal Letter*, 616:L75–L78, December 2004.

[211] Y. P. Jing, P. Zhang, W. P. Lin, L. Gao, and V. Springel. The Influence of Baryons on the Clustering of Matter and Weak-Lensing Surveys. *Astrophysical Journal Letter*, 640:L119–L122, April 2006.

[212] D. H. Rudd, A. R. Zentner, and A. V. Kravtsov. Effects of Baryons and Dissipation on the Matter Power Spectrum. *Astrophysical Journal*, 672:19–32, January 2008.

[213] J. J. Mohr, B. Mathiesen, and A. E. Evrard. Properties of the Intracluster Medium in an Ensemble of Nearby Galaxy Clusters. *Astrophysical Journal*, 517:627–649, June 1999.

[214] P. Popesso, H. Böhringer, M. Romaniello, and W. Voges. RASS-SDSS galaxy cluster survey. II. A unified picture of the cluster luminosity function. *Astronomy & Astrophysics*, 433:415–429, April 2005.

[215] R. Stanek, A. E. Evrard, H. Böhringer, P. Schuecker, and B. Nord. The X-Ray Luminosity-Mass Relation for Local Clusters of Galaxies. *Astrophysical Journal*, 648:956–968, September 2006.

[216] Y.-T. Lin, J. J. Mohr, A. H. Gonzalez, and S. A. Stanford. Evolution of the K-Band Galaxy Cluster Luminosity Function and Scaling Relations. *Astrophysical Journal Letter*, 650:L99–L102, October 2006.

[217] Y.-T. Lin and J. J. Mohr. K-band Properties of Galaxy Clusters and Groups: Brightest Cluster Galaxies and Intracluster Light. *Astrophysical Journal*, 617:879–895, December 2004.

[218] D. Nagai, A. Vikhlinin, and A. V. Kravtsov. Testing X-Ray Measurements of Galaxy Clusters with Cosmological Simulations. *Astrophysical Journal*, 655:98–108, January 2007.

[219] E. Rasia, P. Mazzotta, S. Borgani, L. Moscardini, K. Dolag, G. Tormen, A. Diaferio, and G. Murante. Mismatch between X-Ray and Emission-weighted Temperatures in Galaxy Clusters: Cosmological Implications. *Astrophysical Journal Letter*, 618:L1–L4, January 2005.

[220] T. B. O'Hara, J. J. Mohr, J. J. Bialek, and A. E. Evrard. Effects of Mergers and Core Structure on the Bulk Properties of Nearby Galaxy Clusters. *Astrophysical Journal*, 639:64–80, March 2006.