Challenges of theoretical and numerical structure formation in a $\Lambda\mathrm{CDM}$ universe

Dissertation

zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Nina Roth

aus
Kronberg i. Ts.

Bonn, 2014

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen
Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Cristiano Porciani
2. Gutachter: Prof. Dr. Peter Schneider

Tag der Promotion: 17. 10. 2014
Erscheinungsjahr: 2015

Declaration

I hereby declare that I have produced this thesis without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such.
The thesis work was conducted under the supervision of Prof. Dr. Cristiano Porciani at the Argelander-Institut für Astronomie. It has not previously been presented in identical or similar form to any other German or foreign examination board.
A significant part of this work has been published in, or is being prepared for, peer-reviewed journals:

[1] N. Roth and C. Porciani. *Testing standard perturbation theory and the Eulerian local biasing scheme against N-body simulations.* Monthly Notices of the Royal Astronomical Society, 415:829-844, July 2011

[2] N. Roth and C. Porciani. *Can we really measure* f_{NL} *from the galaxy power spectrum?* Monthly Notices of the Royal Astronomical Society: Letters, 425:L81-L85, September 2012

[3] N. Roth, A. D. Ludlow, J. F. Navarro, et al. *Can mergers explain the origin of the Monoceros stream?* (in preparation)

Abstract

Systematic surveys of the extra-galactic sky have revealed the existence of large-scale structures in the Universe: the galaxy distribution is organized in a complex network of filaments surrounding underdense regions and crossing at density peaks which host galaxy clusters. These structures are believed to form through gravitational instability starting from quantum density fluctuations in the primordial Universe. The seeds for galaxy formation are imprinted as tiny temperature perturbations on the Cosmic Microwave Background (CMB), which provides us with a snapshot of the Universe when it was just 380,000 years old. At late times, observers can measure the statistical properties of galaxies (e. g. spatial clustering), as well as the light distribution of nearby objects in more detail.

Perhaps surprisingly – considering the huge dynamical range and complexity of the Universe – recent years have witnessed the emergence of a cosmological concordance model, dubbed ΛCDM. This six-parameter model is highly successful in describing cosmological structure formation from the very early universe to the current galaxy population. The ΛCDM model derives its name from the two components that make up the majority of the cosmic energy density: Cold Dark Matter (CDM), a species of elementary particles that dominates gravitational structure formation but has so far eluded direct detection, and the cosmological constant Λ, responsible for the observed accelerated expansion. The ΛCDM model is the basis on which we try to address the following fundamental questions:

1. How can we accurately describe the growth of structures over the cosmological time-scale of almost 14 billion years? How do we deal with the assumption that structure formation is dominated by dark matter (which only interacts gravitationally), while we generally observe galaxies via their electromagnetic emission?

2. What is the origin of the initial perturbations? Which theoretical tools allow us to gain insight into the physics of the primordial Universe?

3. Can we apply what we learned from the large-scale evolution of the density field to our Galactic neighborhood? Does the formation history of the Milky Way have implications for our cosmological model?

The theoretical foundation of cosmological structure formation is outlined in the first two chapters of this thesis. After a description of the expanding, homogeneous background space-time, we summarize in Chapter 2 how initial density perturbations grow until they collapse to form bound structures, called dark matter halos. This chapter contains an introduction to the concept of *bias* between the matter

density field and its observable tracers (e. g. galaxies). Finally, we discuss theoretical predictions for the primordial Universe and how specific initial conditions can manifest themselves in large-scale clustering statistics like the power spectrum.

During the last two decades, numerical simulations have become a powerful tool to study astronomical objects on a large range of scales. They are used extensively in all areas of astronomy and astrophysics, like the formation of single stars and planets, analyzing the dynamics of star clusters, modeling hydrodynamical processes in galaxies like our Milky Way, and last but not least studying the large-scale structure of cosmological volumes. In Chapter 3 we describe the general methodology of cosmological N-body simulations, its strengths and limitations, and the specific implementation used in this work, the GADGET-2 code. The following chapters rely heavily on results of these simulations, as they can be used to test theoretical predictions or compare the ΛCDM model with observational data.

After these rather extensive introductory chapters – necessary due to the broad scope of this thesis – Chapters 4 to 6 contain the results of three separate projects, each related to one of the questions outlined above. In Chapter 4 we describe a novel numerical method to test perturbative methods for computing the density contrast of dark matter, and compare the result against full N-body simulations. In addition, we test the validity of a popular bias model and find that it lacks the accuracy required to fully exploit the statistical power of upcoming galaxy surveys.

Chapter 5 deals with local primordial non-Gaussianity (PNG), a specific type of initial conditions for structure formation. While the most stringent constraints on primordial non-Gaussianity currently come from the CMB, galaxy clustering provides an independent validation of these results, and future large-scale structure surveys are even predicted to surpass the CMB constraints. However, we show that galaxy clustering suffers from specific parameter degeneracies which are not present in the CMB. We caution against the commonly used simple model to measure PNG parameters, and instead promote Bayesian model selection to asses the influence of these degeneracies in the data.

Chapter 6 addresses the subject of *near-field cosmology*, i. e. the concept of supplementing the constraints on the overall cosmological model with small-scale information from the local Universe. In particular, this chapter deals with the dynamical properties of simulated subhalos around high-resolution Milky-Way sized dark matter halos, representative of the population of dwarf galaxies around our own Galaxy. We investigate the physical conditions needed to rapidly circularize the orbits of infalling systems, possibly giving rise to peculiar stellar streams such as the recently observed Monoceros overdensity.

Contents

1 Introduction

1.1 ΛCDM – A brief description

Before delving into the theoretical aspects of cosmology we want to state a few key concepts and observations that established the Lambda Cold Dark Matter (ΛCDM) model as the standard model of cosmology over the past ~ 20 years.

Notation and Units We denote 3-vectors in boldface ($\boldsymbol{x}$), while 4-vectors and other tensors are generally distinguished by having Greek indices. When using the Einstein sum convention, Roman indices run from 1 to 3 and Greek indices run from 0 to 3. In all derivations we will assume natural units, i. e. $c = \hbar = 1$. Typical cosmological units are $1\,\mathrm{M}_\odot \approx 2 \cdot 10^{30}$ kg (solar mass), $1\,\mathrm{Mpc} = 10^6\mathrm{pc} \approx 3.1 \cdot 10^{22}\mathrm{m}$ (Megaparsec) and $1\,\mathrm{Gyr} = 10^9$ years (Gigayear). Mass and length units are also often expressed in $h^{-1}\,\mathrm{M}_\odot$ or $h^{-1}\,\mathrm{Mpc}$ where $h \sim 0.7$ parameterises our ignorance of the local expansion rate (see Equation (1.7)). Our Fourier convention is

$$\tilde{f}(\boldsymbol{k}) = \int \mathrm{d}^3x \; \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{x}} \, f(\boldsymbol{x}) \Longleftrightarrow f(\boldsymbol{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3} \; \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\boldsymbol{x}} \, \tilde{f}(\boldsymbol{k}). \tag{1.1}$$

A list of commonly used acronyms is provided at the end of the thesis.

Cosmological redshift Observations of nearby galaxies show that on average, galaxies move away from us. This recession velocity is caused by the fact that the Universe is expanding, and leads to a time dilation which changes the frequency of the photons between emitter and observer. If the peculiar velocities of the objects are negligible, the redshift is a distance indicator, and, because of the finite speed of light, also a measure of the look-back time.

The Cosmological Principle On the largest scales (or at early times), the mass distribution in the Universe is considered to be homogeneous and isotropic. Redshift surveys that measure the angular coordinates and distance of large numbers of galaxies show that their distribution is isotropic on scales $\gtrsim 200$ Mpc (Maddox et al., 1996; Percival et al., 2001), and homogeneity follows from the assumption that our position in the Universe is not special (the Copernican Principle). In addition, the temperature of the CMB radiation, emitted at a time when the Universe was only 380.000 years old (Penzias & Wilson, 1965; Dicke et al., 1965), is isotropic to within 1 part in 10^5 over the whole sky, reflecting the accuracy of the Cosmological Principle at very early times.

Dark Matter (DM) The matter content of the Universe is dominated by dark matter, a species of yet unknown elementary particles, which interact through gravity but do not emit or absorb light (although some species may self-annihilate into standard model particles, and eventually photons). *Cold* dark matter is non-relativistic at the time when the CMB is released, while the terms *warm* and *hot* dark matter refer to particles which are relativistic/ultra-relativistic at early times. Therefore the CDM in ΛCDM stands for cold dark matter.

The existence of dark matter has been deduced indirectly from the rotation curves of spiral galaxies (Sofue & Rubin, 2001), gravitational lensing (Schneider, 2006), the dynamics and x-ray temperature of galaxy clusters (Allen et al., 2011), the galaxy correlation function (Eisenstein et al., 2005), and the amplitude of the Cosmic Microwave Background fluctuations. These observations consistently show that there is about 5 times as much dark matter as baryons[1], and DM contributes about 25% to the total energy budget of the Universe.

DM particles have so far eluded direct detection, apart from the hotly disputed claim from the DAMA/LIBRA[2] experiment (e. g. Bernabei et al., 2008), which seems to be incompatible with several other experiments (Kopp et al., 2010). As the allowed parameter range is rapidly shrinking (Baudis, 2012), alternative models that predict deviations of gravity from General Relativity (GR) are becoming more popular. However, so far none of these *Modified Gravity* models can collectively explain the astronomical observations (Clifton et al., 2012).

Accelerated expansion and Dark Energy (DE) Not only is the Universe expanding, but the expansion is currently *accelerating.* This is attributed to another rather mysterious component called *Dark Energy* (DE). In ΛCDM, DE is expressed by the cosmological constant Λ, while the term Dark Energy often refers to models with time-dependent energy density. So far, no indication of a deviation of Λ from a constant has been found, but current constraints are still nowhere near the percent (or even sub-percent) accuracy of other cosmological parameters (e. g. Suzuki et al., 2012, and Table 2.1). While there were already hints for a non-zero cosmological constant in the early 1990's (Efstathiou et al., 1990), the definitive proof came from SN observations (Riess et al., 1998; Perlmutter et al., 1999); a discovery which was subsequently awarded with the Nobel Prize for Physics in 2011. Approximately 70% of the total energy budget of the Universe is found to be in the form of dark energy (e. g. Planck Collaboration et al., 2013a).

1.2 Describing space-time with General Relativity

We start our description of the cosmos by considering the expanding background that obeys the Cosmological Principle; perturbations will be added in Chapter

---

[1] In cosmology, the term "baryonic" matter encompasses the standard model particles: electrons, protons and all higher elements. Neutrinos are usually treated separately, as they are standard model particles but could also be considered a form of relativistic dark matter.

[2] DArk MAtter Large sodium Iodide Bulk for RAre processes

2. For the sake of brevity, we will only collect the most important definitions and results in this section. For detailed derivations and proofs we refer to e. g. Dodelson (2003) or Weinberg (2008).

The first objective in this context is to define a suitable set of coordinates. For a general metric $g_{\mu\nu}$, the line element is given by

$$\mathrm{d}s^2 = -g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu}, \tag{1.2}$$

where repeated indices are summed over. We follow the usual convention that the 0 coordinate denotes time and our choice for the metric signature is $\{-,+,+,+\}$. One can define a *comoving* observer, who always sees the Universe as homogeneous and isotropic, by considering hyper-surfaces of constant time t and 3D spatial coordinate $\boldsymbol{x}$. This coordinate choice leads to the Friedmann-Robertson-Walker (FRW) metric

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right] . \tag{1.3}$$

The spatial part reduces to spherical coordinates for $k = 0$, and k therefore describes the 3D *curvature*, which can be positive ($k > 0$), negative ($k < 0$) or Euclidean ($k = 0$), and has units of length^{-2}. The *scale factor* $a(t)$ describes the radial expansion (or contraction) of the Universe. Conventionally, $a(t)$ is fixed to be unity today: $a(t_0) \equiv 1$. While the comoving observer follows the expansion of the background, Equation (1.3) tells us that *physical* coordinates of an object are time-dependent:

$$\boldsymbol{r}(t) = a(t)\boldsymbol{x}. \tag{1.4}$$

By introducing a new radial coordinate χ, the metric can be rewritten as

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\left[\mathrm{d}\chi^2 + f_{\mathrm{k}}^2(\chi)(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right] , \tag{1.5}$$

where

$$f_k(\chi) = \begin{cases} \sin(\sqrt{k}\chi)/\sqrt{k} & \mathrm{k} > 0 \\ \chi & \mathrm{k} = 0 \\ \sinh(\sqrt{-k}\chi)/\sqrt{-k} & \mathrm{k} < 0 \end{cases} . \tag{1.6}$$

Independently of the exact expressions for $a(t)$ and the value of k, we can derive several quantities that follow from the FRW metric alone. The time evolution of the scale factor can be expressed by the *expansion rate* of the Universe, also called the Hubble parameter,

$$H(t) = \frac{\dot{a}}{a}, \tag{1.7}$$

where the dot denotes a derivative with respect to t. The value of $H(t)$ today, H_0, is called the *Hubble constant*, and it is often parameterised by $H_0 = 100\,h\,\mathrm{km\,s^{-1}Mpc^{-1}}$ with $h \sim 0.7$.

An important consequence of the expansion of the Universe is that photons are redshifted due to time dilation. Light rays travel on null-geodesics of space-time, which implies $\mathrm{d}s^2 = 0$ in Equation (1.3). For radial light rays (assuming without loss of generality that $\mathrm{d}\theta = \mathrm{d}\phi = 0$), this leads to

$$\mathrm{d}t = \pm a(t)\,\mathrm{d}\chi, \tag{1.8}$$

where the sign depends on the direction of the light ray (away from or towards the observer). For a light ray emitted at t_1 and reaching the observer at the coordinate origin at t_0, this can be integrated to yield the comoving radial separation:

$$\chi = \int_{t_1}^{t_0} \frac{\mathrm{d}t'}{a(t')}. \tag{1.9}$$

Since χ is a comoving coordinate, which is independent of time, we immediately see that a time interval at time t_1 is related to time interval at t_0 via

$$\frac{\mathrm{d}t_1}{a(t_1)} = \frac{\mathrm{d}t_0}{a(t_0)}, \tag{1.10}$$

and since a time interval is the inverse of the frequency we can define the *redshift* z by

$$(1+z) \equiv \frac{\nu_1}{\nu_0} = \frac{a(t_0)}{a(t_1)} = \frac{1}{a(t_1)}, \tag{1.11}$$

where in the last step we have used $a(t_0) = 1$. By definition, $z = 0$ today and the sign of z determines whether we see photons with a redshift or a blueshift. If the emission frequency of the photon can be determined (i. e. when the physical process is known), the redshift can also be used as a distance indicator.

The fact that all but a few nearby galaxies appear to be redshifted tells us that the scale factor was smaller at earlier times. The outliers are due to an additional Doppler shift from the galaxy's peculiar velocity (of the order of a few hundred km/s), which can be larger than the cosmological redshift for nearby objects. For an object with velocity $\boldsymbol{u}(\boldsymbol{x},t) \equiv \dot{\boldsymbol{r}}(t)$, the peculiar velocity $\boldsymbol{v}(\boldsymbol{x},t)$ is therefore defined by

$$\boldsymbol{u}(\boldsymbol{x},t) = \dot{a}\boldsymbol{x} + \boldsymbol{v}(\boldsymbol{x},t), \tag{1.12}$$

and the first term is the recession velocity due to expansion. For nearby sources, we recover the famous Hubble law $v_{\mathrm{rec}} = H_0|\boldsymbol{r}|$.

Equation (1.9) also allows us to define the comoving particle horizon, i. e. the maximum distance that light can travel in a time interval between t_i and t:

$$\chi_{\mathrm{h}} = \int_{t_{\mathrm{i}}}^{t} \frac{\mathrm{d}t}{a(t)} = \int_{a_i}^{a} \frac{\mathrm{d}\ln a}{aH(a)} \tag{1.13}$$

If we choose the initial time $t_{\mathrm{i}} \to 0$, $d_{\mathrm{h}} = a(t)\,\chi_{\mathrm{h}}$ gives us the maximum physical separation of two regions in space that have been in causal contact between the Big Bang and time t. As we will explain in section 2.3, this causality condition is apparently violated by the uniformity of the Cosmic Microwave Background. A solution for this *Horizon problem* is provided by the theory of Inflation.

Apart from the radial separation, we are also interested in the distance between two points measured perpendicular to the line of sight, e. g. for two sources at the same redshift (or the diameter of an extended source), given an angle $\mathrm{d}\theta$ on the sky. The length of this line element $\mathrm{d}s$ can be obtained from Equation (1.5) by setting $\mathrm{d}t = \mathrm{d}\chi = \mathrm{d}\phi = 0$, and leads to the definition the angular diameter distance

$$D_\mathrm{A} = \frac{\mathrm{d}s}{\mathrm{d}\theta} = a(t) f_k(\chi). \tag{1.14}$$

Although it is not explicitly used in this thesis, we will mention another distance measure for completeness, the luminosity distance. Since the observed flux F of a source decreases with the square of the distance to the observer, we can define the luminosity distance as

$$D_\mathrm{L} = \sqrt{\frac{L}{4\pi F}} \tag{1.15}$$

where L is the total luminosity over all frequencies. The luminosity distance is employed for example when using Supernovae of Type Ia to constrain the expansion history; their intrinsic brightness is related to the decay time of their light emission (Branch & Tammann, 1992).

The distance estimates are related by the distance-duality relation

$$D_L = (1+z)^2 D_A = (1+z) f_k(\chi), \tag{1.16}$$

which holds independently on the curvature k. Testing the distance-duality relation by combining different probes provides constraints on the metric and even exotic physics like the nature of dark energy (Bassett & Kunz, 2004).

1.3 The energy content of the Universe

The expressions derived in the previous section follow directly from the FRW metric without any assumptions on the exact form of scale factor or the value of the spatial curvature. Enter General Relativity (GR), which relates the metric $g_{\mu\nu}$ to the energy content of the Universe – parameterised by the energy-momentum tensor $T_{\mu\nu}$ – in the Einstein Field Equations (EFE)

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu}, \tag{1.17}$$

where the Ricci tensor $R_{\mu\nu}$ and Ricci scalar R are uniquely determined by first and second derivatives of the metric, and G is the gravitational constant. The cosmological constant Λ was originally introduced by Einstein to allow for a static Universe, but is used today to explain the observed accelerated expansion (see below). Note that the signs on the right-hand side of Equation (1.17) depend on the sign conventions used in the derivation (e. g. Misner et al., 1973).

Applying the Cosmological Principle, it can be shown that $T_{\mu\nu}$ has to be the energy-momentum tensor of a perfect fluid[3], $T = \mathrm{diag}(\rho, p, p, p)$, where the mass

---

[3] A perfect fluid is characterised by its rest frame mass density and isotropic pressure; shear

density ρ and the pressure p are related via an equation of state (EoS)

$$p = w\rho. \tag{1.18}$$

Inserting the metric and T into Equation (1.17) leads (after some algebra) to the Friedmann equations (FE)

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3} \tag{1.19}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} \tag{1.20}$$

which describe the expansion rate and the acceleration as functions of the density, pressure and spatial curvature. The two Friedmann equations can be combined to yield the adiabatic equation[4]

$$\frac{\mathrm{d}(a^3\rho)}{\mathrm{d}t} = -p\frac{\mathrm{d}a^3}{\mathrm{d}t}\ , \tag{1.21}$$

which is valid for each fluid component individually if there is no exchange of energy. Assuming a constant EoS parameter w, the adiabatic equation leads to

$$\rho(a) = a^{-3(w+1)}\rho_0, \tag{1.22}$$

where ρ_0 is the density today, at $a(t_0) = 1$.

The time evolution is therefore related to the EoS parameter of the fluid. For non-relativistic matter (CDM, baryons) $w \approx 0$, while for relativistic particles (photons, neutrinos) $w = 1/3$. The cosmological constant Λ can also be included by defining $\rho_\Lambda = (8\pi G)^{-1}\Lambda$, and $\rho_\Lambda = \text{const.}$ then requires $w_\Lambda = -1$. This leads to the peculiar property $p_\Lambda = -\rho_\Lambda$, and further emphasizes the significance of the observed accelerated expansion at late times: The acceleration $\ddot{a}$ depends on $\rho + 3p$ which is positive for ordinary matter, but negative for a dominating cosmological constant Λ, or any other species with $w < -1/3$. However, this phenomenological description does not offer much in the way of explaining the physics behind a negative EoS parameter.

Assuming a non-interacting fluid with time-independent w for each component, the total energy density and pressure can then be written as

$$\rho(a) = \rho_\mathrm{m}(a) + \rho_\mathrm{r}(a) + \rho_\Lambda, \qquad p(a) = \frac{\rho_\mathrm{r}(a)}{3} - \rho_\Lambda, \tag{1.23}$$

where we have used subscripts to distinguish “matter” (non-relativistic particles) and “radiation” (relativistic particles).

---

stress and viscosity vanishes. This property is reminiscent of the definition of the comoving observer, and is therefore a sensible choice. A more formal proof can be found e. g. in Weinberg (2008).

[4] Any two equations from the set (1.19)-(1.21) can be used to derive the third, and Equation (1.21) is also a direct consequence of local energy-momentum conservation in the EFE.

The final ingredient for our model is the curvature parameter k. We first define the *critical density* $\rho_{\mathrm{crit}}(a) = 3H^2(a)/(8\pi G)$, which is the total energy density for a flat Universe ($k = 0$). The *density parameters* are then given by normalizing ρ_i with ρ_{crit},

$$\Omega_{\mathrm{m}}(a) = \frac{\rho_{\mathrm{m}}(a)}{\rho_{\mathrm{crit}}(a)}, \qquad \Omega_{\mathrm{r}}(a) = \frac{\rho_{\mathrm{r}}(a)}{\rho_{\mathrm{crit}}(a)}, \qquad \Omega_{\Lambda}(a) = \frac{\rho_{\Lambda}}{\rho_{\mathrm{crit}}(a)}, \tag{1.24}$$

$$\Omega_{\mathrm{tot}}(a) = \Omega_{\mathrm{m}}(a) + \Omega_{\mathrm{r}}(a) + \Omega_{\Lambda}(a) = 1 + k/(aH)^2 \tag{1.25}$$

where the last equality on the second line follows from Equation (1.19). A universe with $k > 0$ has a total density higher than ρ_{crit}, while a universe with negative curvature has a total density less than ρ_{crit}. The curvature can thus be determined by the deviation of Ω_{tot} from unity. Note that when the density parameters are measured from observations, their value is generally given at $z = 0$, which is denoted by $\Omega_{i,0}$ or simply Ω_i without an explicit time-dependence. In that case, Equation (1.19) can be rewritten as

$$H^2(a) = H_0^2\left[\frac{\Omega_{\mathrm{r}}}{a^4} + \frac{\Omega_{\mathrm{m}}}{a^3} + \frac{(1-\Omega_{\mathrm{tot}})}{a^2} + \Omega_{\Lambda}\right], \tag{1.26}$$

where the time-dependence of the individual components follows directly from Equation (1.22). In general, $H(a)$ contains contributions from all components of the fluid and the solution for $a(t)$ has to be found numerically. However, the different dependence on $a(t)$ means we can obtain approximate analytical solutions for certain epochs where a single component dominates the energy density. For a flat Universe dominated by Ω_{r}, it is straightforward to show that $a(t) \sim t^{1/2}$, whereas matter domination is characterised by $a(t) \sim t^{2/3}$.

Current constraints on the density parameters at $z = 0$ are $\Omega_{\mathrm{m}} \sim 0.3$ (baryons and dark matter), $\Omega_{\Lambda} \sim 0.7$, $\Omega_{\mathrm{r}} < 10^{-4}$ and $|1 - \Omega_{\mathrm{tot}}| \lesssim \mathcal{O}(10^{-3})$, whereas the Hubble constant is $H_0 \sim 70$ km s^{-1} Mpc^{-1}. For a more complete list of cosmological parameters (including error bars) see Table 2.1 at the end of Chapter 2.

1.4 Thermal history

In this section, we will give a brief description of the thermal history of our Universe. As we go back in time ($a \to 0$), the Universe becomes hotter and denser, and the interaction rates increase as well. This implies that at some point all particle species were in thermal equilibrium, and thus the most important quantity in the following discussion will be the temperature T. Since $E = k_{\mathrm{B}}T$, we can express the temperature in eV, where $1\mathrm{eV} \sim 1.2 \cdot 10^4\,\mathrm{K}$. With our choice of $c = 1$, mass and energy are equivalent as well and we can use temperature, energy and mass interchangeably.

We will start our discussion at energy scales of a few hundred MeV, where baryons, photons and neutrinos are tightly coupled, and the temperature of the

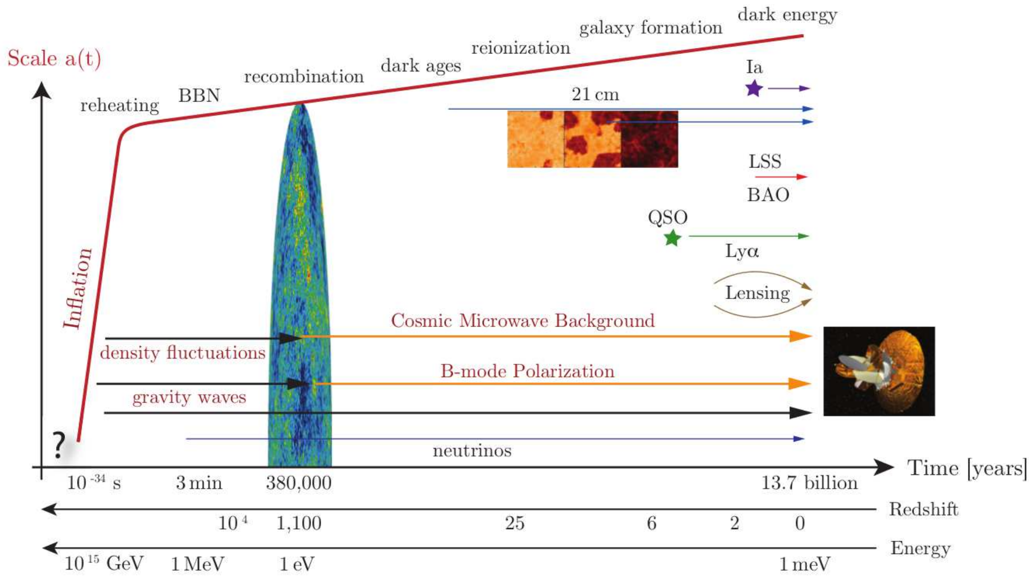

Figure 1.1: Evolution of the cosmological scale factor as a function of time, redshift, and the corresponding energy scale. The different epochs and observational techniques are explained in detail in the text. [Figure taken from Baumann (2009).]

Universe is thus identical to that of the plasma, which implies $T \sim (1+z)$.[5] Note that we do not know a priori if the physical laws that apply on Earth are applicable also in the early Universe, since current particle physics experiments are only able to probe physics up to energy scales of a few TeV. However, this also means that the early Universe is a test bed for new physics on energy scales that are out of reach of current (and even future) particle colliders. A qualitative illustration of the different epochs along with a selection of observational probes used to study them is shown in Figure 1.1 (from Baumann 2009), and will be described in the following.

The early Universe The initial conditions of the Universe are a highly speculative phase that is often based on effective theories that are able to reproduce the late-time amplitude and evolution of fluctuations. An important ingredient is a period of early exponential expansion, called *Inflation* (section 2.3), caused by a scalar field that dominates the energy density of the Universe. It is very difficult to put direct observational constraints on the physics of the early Universe as no photons can from before the epoch of recombination (see below) can reach us. Some models of Inflation predict an observable background of gravitational waves, but the sensitivity of current instruments is still too low to observe it. At the end of the inflationary phase the scalar field decays and its energy is used to produce all known particles species (“reheating”), and the Universe enters into the radiation domination phase.

---

[5] This follows immediately for photons from the redshift and $E_\gamma \sim k_\mathrm{B}T$, and can be extended to the tightly coupled baryon-photon plasma because the ratio of the number density of baryons to photons is $\eta \sim 10^{-8}$.

Radiation domination After reheating, the energy contents of the Universe can be described by a hot plasma. Baryons and radiation (photons and neutrinos) are tightly coupled to each other and are thus in thermal equilibrium. The scale factor grows as $a(t) \sim t^{1/2}$, so the expansion is no longer accelerated. Initial density fluctuations begin to grow under the effects of gravity, but the photon pressure in the plasma temporarily prevents the condensation of *baryonic* overdensities. As the temperature drops, interaction rates between particles decrease and begin to slowly move away form thermodynamical equilibrium. The neutrinos are the first to decouple from the plasma since they only interact via the weak force. In principle, this neutrino background should present a homogeneous signal comparable to the CMB (with slightly lower temperature), but neutrinos are of course notoriously difficult to detect.

Big Bang Nucleosynthesis (BBN) Once the temperature drops to a few MeV, higher elements can form out of protons and neutrons. This is limited to Helium isotopes and traces of Lithium and Beryllium because of the lack of tightly bound isotopes with mass numbers between 5 and 8, and the large number of dissociating high-energy photons. Since three-body interactions are rare, the formation of higher elements is strongly suppressed until the first generations of stars provide more favourable conditions for their production. Using standard interaction rates and the expansion predicted by the Friedmann equations, BBN predicts the primordial abundance of light elements in excellent agreement with observations (e. g. Alpher et al., 1948; Walker et al., 1991; Spergel et al., 2007).

Matter domination At around $z \sim 4800$ the Universe enters into the matter dominated phase. A flat Universe dominated by cold dark matter and without a cosmological constant, i. e. $\Omega_{\mathrm{tot}} = \Omega_{\mathrm{m}} = 1$, is called an Einstein-deSitter (EdS) Universe, and the expansion is described by $a(t) \sim t^{2/3}$. Even after the paradigm shift to ΛCDM, the EdS model is still often used to simplify analytical calculations. Recombination and reionization (see below), and most of the actual structure growth happen during this epoch, which eventually ends when the cosmological constant Λ begins to dominate around $z = 0.6$.

Recombination and the CMB Around $z \sim 1100$, free electrons and protons can combine to neutral hydrogen, and the photons are released as a homogeneous background radiation, the Cosmic Microwave Background. After recombination, the baryons can follow gravity and fall into the potential wells created by the dark matter which was not prevented from collapsing by the pressure of the photons. However, there will also be a small but significant influence of the baryons on the dark matter distribution, which manifests itself in the so-called Baryonic Acoustic Oscillations (BAO) imprinted on the CMB and on galaxy clustering.

The recombination processes are so efficient that the Universe is opaque to high-energy photons for a while, because their mean-free path is very small and their temperature keeps decreasing through the expansion. This epoch is therefore called the “dark ages”, because it can not be probed by electromagnetic radiation. Perturbations continue to grow and in the highest overdensities, the clumps of neutral hydrogen become so dense that they can ignite fusion processes, leading to the first generation of stars. The light from the CMB is also polarized through anisotropic Thompson scattering during its propagation through the large-scale structure. This polarisation contains additional information about structure formation in the early Universe and will be probed in great detail by the *Planck* satellite.

Reionization and galaxy formation Once the first stars and galaxies have formed, their radiation can re-ionize the surrounding neutral medium. This means that from this point on the Universe will become transparent to photons again. On small scales, this is expected to happen in patches around the strongest ionizing sources[6], but macroscopically it is still very difficult to determine the exact physical processes and even the starting and end point of this epoch. Apart from the strongly ionizing radiation from quasars, dwarf galaxies may also play a significant role – even though their luminosity is much lower – because they are much more numerous (Stark et al., 2007). The epoch of reionization is still poorly constrained, but current estimates place it between $z = 12$ and $z = 6$ (Stark et al., 2010, and references therein). Since the Universe at those times still consists mainly of neutral hydrogen, the 21-cm spin-flip transition can be used to study this phase.

Late-time accelerated expansion The final transition occurs at $z \sim 0.6$, about 5.6 Gyrs ago. The Universe turns from decelerating into accelerating as it becomes dominated by Dark Energy/the cosmological constant Λ. Several ways of probing the late-time evolution of structures exist: as mentioned above, Supernovae Type Ia can be used to derive the expansion history because their intrinsic luminosity is known. The statistics of the clustering of galaxies in the large-scale structure (LSS). The Lyman-α emission of quasars (QSOs) can be seen to high redshifts because these objects are intrinsically very luminous. In addition, gravitational deflection of photons (Gravitational Lensing, see e. g. Schneider 2006) by the intervening matter can be used to constrain the statistics of the density field directly.

---

[6] Quasars or quasi-stellar objects (QSOs) are galaxies with a supermassive black hole in their center which dumps a lot of energy into the surrounding medium when it is fed by mass accretion.

2 Structure formation

A universe described by the Friedmann-Robertson-Walker (FRW) metric is a rather boring place. Since the matter distribution is homogeneous and isotropic by construction, it would simply remain in this equilibrium state forever – apart from expansion and contraction. In order to form any of the structures we see today (stars, galaxies, . . .), the primordial universe has to allow for some sort of perturbations about the background given by the FRW metric. Note that in this chapter we will often make qualitative comparisons to the results of numerical simulations, which are formally introduced in Chapter 3.

The earliest picture we have of the inhomogeneous Universe is the Cosmic Microwave Background (CMB), showing relative temperature differences of the order 10^{-5}. Temperature perturbations correspond to density perturbations, and initially overdense regions will attract more matter than regions with less than average density (section 2.1). These overdense regions will eventually stop expanding with the background and collapse into bound structures, called *dark matter halos* (section 2.2). These halos serve as condensation points for galaxy formation by providing the potential wells into which baryons can fall after decoupling. The spatial clustering as well as the mass function[1] of halos contain a wealth of cosmological information. However, parameter constraints from large-scale structure are severely hampered by the existence of *biasing* (section 2.2.1).

The origin and amplitude of the initial perturbations is a puzzle by itself, because at early times we have to rely on effective theories to make predictions that are in agreement with the properties of the CMB. In section 2.3 we will discuss how the theory of *Inflation* can be used to predict the amplitude of the initial density perturbations as well solve several apparent fine-tuning problems in the early Universe. Most models of Inflation depend on applying quantum field theory to one or more *scalar* fields, which had been a purely theoretical concept until the recent discovery of the Higgs Boson (ATLAS Collaboration et al., 2012), therefore providing an intimate connection between the early Universe and (extensions of) the standard model of particle physics.

In section 2.4 we introduce the concept of Primordial non-Gaussianity (PNG), i. e. the deviation of the distribution function of the perturbations from a Gaussian. Any detection of PNG is a direct probe for inflationary physics, as well as an opportunity to find new physics. While current constraints on PNG comes primarily from the CMB, future large-scale-structure (LSS) observations will be able to obtain similar precision by taking advantage of a scale-dependence in the bias parameter which is absent in the Gaussian case.

---

[1] The number density of objects in a given mass range, see section 2.1.4

2.1 Describing structure growth

We have described the energy content of the Universe by a perfect fluid consisting of dark matter, "baryons" like electrons and protons, and "radiation" in the form of neutrinos and photons. The most general equation to describe a (possibly) relativistic fluid with phase-space distribution function $f(\boldsymbol{x}, \boldsymbol{v}, t)$ is the collisional Boltzmann Equation (BE), which in its most basic form looks quite innocent,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = C\left[f\right], \tag{2.1}$$

and simply enforces particle conservation in phase-space. Here, $\boldsymbol{v}$ is the peculiar velocity defined in Equation (1.12), and the term on the right-hand side of Equation (2.1) describes particle *collisions*, or scattering processes. However, describing the scattering processes and including superhorizon[2] effects using GR is highly non-trivial. For a full treatment we refer to e. g. Ma & Bertschinger (1995) or Dodelson (2003).

Here, we are mainly interested in the growth of matter perturbations which means we can focus on cold dark matter (CDM). For the case of a fluid dominated by CDM, we can make several simplifying assumptions which also provide the basis of Standard Perturbation Theory used in Chapter 4. CDM is collisionless and thus the right-hand side of Equation (2.1) vanishes. Secondly, on scales smaller than the horizon, we can apply the weak-field limit and describe the dark matter fluid with Newtonian dynamics only. Solutions to the BE can then be derived by taking moments of the peculiar velocity, i. e. $\int \mathrm{d}^3v\, \boldsymbol{v}^n f(\boldsymbol{x}, \boldsymbol{v}, t)$ with $n = 0, 1, 2$, and so on.

Furthermore, we only consider the growth of fluctuations during the matter dominated epoch, which allows us to use the Friedman equation for a pure Einstein-deSitter (EdS) Universe thereby simplifying the calculations. This assumption is justified because it can be shown that any subhorizon growth is dominated by that during the matter dominated epoch. Finally, we consider the dark matter fluid in the single-stream approximation (i. e. neglecting velocity dispersion), which implies that any moments with $n > 2$ vanish. Even after all these simplifying assumptions, the resulting analytic description is quite accurate as long as the perturbations stay small. Other effects can be included in the transfer function (section 2.1.2) which is usually calculated numerically.

In addition to the peculiar velocity, it is useful to define the *density contrast*

$$\delta(\boldsymbol{x}, t) \equiv \frac{\rho(\boldsymbol{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \tag{2.2}$$

where $\bar{\rho}(t)$ is the time-dependent mean density of the Universe. Note that the foregoing equations are already expressed in comoving coordinates. The 3 equations describing a Newtonian fluid in terms of $\delta(\boldsymbol{x}, t)$, $\boldsymbol{v}(\boldsymbol{x}, t)$ and the scale factor $a(t)$

---

[2] A perturbation is said to be "inside the horizon" at time t if its wavelength in Fourier space is $k \gg aH$. Conversely, $k \ll aH$ implies superhorizon perturbations.

are the Continuity Equation (describing mass conservation)

$$\frac{\partial \delta(\boldsymbol{x},t)}{\partial t} + \frac{1}{a}\boldsymbol{\nabla}_{\boldsymbol{x}} \cdot [(1+\delta(\boldsymbol{x},t)) \cdot \boldsymbol{v}(\boldsymbol{x},\boldsymbol{t})] = 0, \tag{2.3}$$

the Euler Equation (describing momentum conservation)

$$\frac{\partial \boldsymbol{v}(\boldsymbol{x},t)}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v}(\boldsymbol{x},t) + \frac{1}{a}[\boldsymbol{v}(\boldsymbol{x},t) \cdot \nabla_{\boldsymbol{x}}]\boldsymbol{v}(\boldsymbol{x},t) = -\frac{1}{a}\nabla_{\boldsymbol{x}}\Phi(\boldsymbol{x},t), \tag{2.4}$$

and the Poisson equation

$$\nabla_{\boldsymbol{x}}^2\Phi(\boldsymbol{x},t) = \frac{3H_0^2\Omega_{\mathrm{m}}}{2a}\delta(\boldsymbol{x},t), \tag{2.5}$$

which relates the gravitational potential and the density contrast in an expanding Universe. Since these equations form a set of non-linear, coupled partial differential equations, a general analytical solution does not exist. Numerical solutions are considered in Chapter 3, but it is also possible to find approximate analytical solutions via perturbation theory. Finding the linear solution is quite straightforward and will be discussed briefly in the next section, while Chapter 4 deals with perturbation theory up to third order.

2.1.1 Linear theory

If we only consider small perturbations in the density contrast, the peculiar velocity and the gravitational potential: $\delta \ll 1$, $|\boldsymbol{v}| \ll 1$ and $\Phi \ll 1$, we can linearise Equations (2.3)-(2.5) in these quantities. Neglecting higher-order terms leads to

$$\frac{\partial \delta(\boldsymbol{x},t)}{\partial t} + \frac{1}{a(t)}\nabla_{\boldsymbol{x}} \cdot \boldsymbol{v}(\boldsymbol{x},t) = 0, \tag{2.6}$$

and

$$\frac{\partial \boldsymbol{v}(\boldsymbol{x},t)}{\partial t} + H(t)\boldsymbol{v}(\boldsymbol{x},t) = -\frac{1}{a(t)}\nabla_{\boldsymbol{x}}\Phi(\boldsymbol{x},t), \tag{2.7}$$

while the Poisson equation is already linear in δ, $|\boldsymbol{v}|$ and Φ.

We can now combine these equations by taking the time derivative of Equation (2.6), the divergence of Equation (2.7) and inserting the Poisson Equation to obtain

$$\frac{\partial^2 \delta(\boldsymbol{x},t)}{\partial t^2} + 2H(t)\frac{\partial \delta(\boldsymbol{x},t)}{\partial t} - \frac{3H_0^2\Omega_{\mathrm{m}}}{2a^3(t)}\delta(\boldsymbol{x},t) = 0. \tag{2.8}$$

This equation does not explicitly contain $\boldsymbol{x}$ or derivatives of $\boldsymbol{x}$. Therefore the solution factorises and can be written as a product of functions of $\boldsymbol{x}$ and t:

$$\delta(\boldsymbol{x},t) = D_+(t)\Delta_+(\boldsymbol{x}) + D_-(t)\Delta_-(\boldsymbol{x}), \tag{2.9}$$

where D_+ and D_- are the two independent solutions of the time dependent part of Equation (2.8). D_+ describes the time evolution of growing perturbations and is thus called the *linear growth-factor*, while the second term describes a solution that decays with time, which we will therefore neglect.

One can show by insertion that $H(t)$ is a solution to Equation (2.9) (albeit the decaying part, and only for a time-independent EoS parameter), but it can be used to construct the growing solution (Heath, 1977). Assuming a constant w_Λ for the vacuum energy, $D(a) \equiv D_+(a)$ is given by:

$$D(a) \propto H(a) \int_0^a \frac{\mathrm{d}a'}{[\Omega_\mathrm{m}/a' + \Omega_\Lambda a'^2 + \Omega_\mathrm{r}/a'^2 + (\Omega_\mathrm{tot} - 1)]^{3/2}}, \tag{2.10}$$

which needs to be evaluated numerically except in a few limiting cases. For example, in an Einstein-deSitter (EdS) Universe, where $\Omega_{tot} = \Omega_m = 1$, we have $D(t) \sim a(t)$ and structures simply grow proportional to the scale factor. For a ΛCDM Universe, $D(t)$ instead grows slower than $a(t)$ at late times, leading to a suppression the density contrast relative to the EdS case.

A note of caution: Two different conventions can be used to determine the constant of proportionality in Equation (2.10), either by defining $D \equiv 1$ today (convenient when dealing mostly with event at $z = 0$), or by defining $D \equiv a$ during matter domination (sensible because we know the amplitude of the fluctuations at that time from the CMB). When using the latter definition, structure formation in a ΛCDM Universe is suppressed by a factor of ~ 1.3 at $z = 0$ with respect to the EdS solution. For consistency with the existing literature we are forced to use both normalizations in this thesis: the former in Chapter 4 and the latter in Chapter 5.

2.1.2 Transfer function

We need to keep in mind the assumptions that we made in order to obtain the solution for $D(t)$: we have only considered matter perturbations inside the horizon during the epoch of matter domination. This means that the “initial” density fluctuation Δ_+ in Equation (2.9) is not equivalent to the actual primordial density perturbations that are generated during Inflation (section 2.3). Finding out what happens to perturbations e. g. on superhorizon scales or during the earlier epoch of radiation domination unfortunately does require extended calculations with the Boltzmann Equation. While some parts can be done analytically, the final steps generally involve numerical integration, and the results depend on the cosmological parameters and the detailed physical properties of the components (e. g. “hot” vs. “cold” dark matter, or the small but important influence of baryons on dark matter).

The main result from the full treatment for matter perturbations is the following (e. g. Dodelson, 2003): on superhorizon scales, perturbations are always allowed to grow, both during radiation and matter domination. However, on scales smaller than the horizon, matter perturbations suffer suppressed growth during radiation domination. This means that a perturbation which enters the horizon early, during radiation domination, will be suppressed at late times with respect to a matter perturbation that entered the horizon only during matter domination. The difference in growth can be expressed through the *transfer function* $T(k)$, defined by

$$\frac{\tilde{\delta}(k, z_0)}{\tilde{\delta}(k, z_i)} = T(k)\frac{\tilde{\delta}(k_s, z_0)}{\tilde{\delta}(k_s, z_i)}. \tag{2.11}$$

The transfer function describes the growth of a density perturbation on scale k today, $\tilde{\delta}(k, z_0)$ relative to the growth of a perturbation that never suffered from suppressed growth during radiation domination $\tilde{\delta}(k_\mathrm{s} \ll k, z_0)$, normalized with respect to the initial density perturbations $\tilde{\delta}_\mathrm{i}(k)$ and $\tilde{\delta}_\mathrm{i}(k_\mathrm{s})$.

By definition, $T(k) \to 1$ on the largest scales, but the small scale behaviour depends sensitively on the exact nature of the DM particles. For CDM, $T(k) \sim \ln{(k)}/k^2$ on small scales, while the free streaming of warm or hot dark matter is able to erase structures for large k, which translates into an exponential cutoff for $T(k)$. Approximate fitting functions were provided for example by Bardeen et al. (1986) and Eisenstein & Hu (1998), but higher accuracy can be obtained using optimized numerical algorithms like CAMB (Code for Anisotropies in the Microwave Background, Lewis et al. 2000).

2.1.3 Statistics of spatial clustering

As we will discuss in more detail in section 2.3, the theory of Inflation predicts that the perturbations in the primordial density field are sourced by quantum fluctuations which are blown up to classical scales by a period of exponential expansion in the very early Universe. This means that we merely observe a *realisation* of a random field with certain statistical properties, which can be expressed by a (formally infinite) series of *moments*. In the case of the underlying density field $\delta(\boldsymbol{x}, t)$, we already know from the definition that the first moment, the ensemble average $\langle\delta(\boldsymbol{x}, t)\rangle = 0$ at any time t.

We can also consider the spatial clustering in Fourier space, which leads to the definition of the *power spectrum* $P(k)$ as the connected part of the second moment:

$$\langle\tilde{\delta}(\boldsymbol{k}_1)\tilde{\delta}(\boldsymbol{k}_2)\rangle_c = (2\pi)^3\delta_D(\boldsymbol{k} - \boldsymbol{k}_1 - \boldsymbol{k}_2)P(k), \tag{2.12}$$

where the fact that $P(k)$ only depends on the modulus of $\boldsymbol{k}$ is due to the assumed isotropy, and the Dirac delta function δ_D ensures the translation invariance required by homogeneity. Defined in this way, the power spectrum is simply the Fourier transform of the 2-point correlation function, which describes the excess clustering with respect to a random field generated by a Poisson process (Peebles, 1980). The connected, or irreducible part is defined by subtracting any lower order contribution:

$$\langle\tilde{\delta}(\boldsymbol{k}_1)\tilde{\delta}(\boldsymbol{k}_2)\rangle_c = \langle\tilde{\delta}(\boldsymbol{k}_1)\tilde{\delta}(\boldsymbol{k}_2)\rangle - \langle\tilde{\delta}(\boldsymbol{k}_1)\rangle\langle\tilde{\delta}(\boldsymbol{k}_2)\rangle, \tag{2.13}$$

so in the case of a zero-mean field $\langle\delta_1\delta_2\rangle_c = \langle\delta_1\delta_2\rangle$. The normalisation of $P(k)$ is often expressed through the variance of the power spectrum calculated from linear theory, $P_{\rm lin}(k,z)$,

$$\sigma^2(R,z) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\, P_{\rm lin}(k,z)|W(R,k)|^2, \tag{2.14}$$

where $W(k,R)$ is the Fourier transform of the filter function used to smooth the density field on scale R.[3] For historical reasons, a common choice is to normalise the power spectrum by $\sigma_8 \equiv \sigma(R = 8\,h^{-1}\,\mathrm{Mpc}, z=0) \sim 0.8$.

Analogously, we can define the bispectrum via the connected part of the 3-point correlator:

$$\langle \tilde{\delta}(\boldsymbol{k}_1)\tilde{\delta}(\boldsymbol{k}_2)\tilde{\delta}(\boldsymbol{k}_3)\rangle_{\rm c} = (2\pi)^3 \delta_{\rm D}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3). \tag{2.15}$$

In this case, the Dirac delta function implies that only closed triangle configurations contribute to the signal.

In general, an infinite number of moments is required to describe the density field. In the case of a Gaussian random field, however, all the information is contained in the 2-point function. Higher order connected moments either vanish (for n odd), or can be written as products of the power spectrum using Wick's theorem. The temperature fluctuations in the CMB tell us that the primordial Universe is indeed very nearly Gaussian (e. g. Planck Collaboration et al., 2013b), which is also predicted by the simplest models of Inflation (section 2.3). However, even a small deviation from Gaussianity in the initial conditions, so-called primordial non-Gaussianity (PNG) would constitute a fundamental discovery. We will return to this topic in section 2.4 and Chapter 5. Additionally, non-linear structure formation introduces non-Gaussian features at late times by mixing different k-modes (Fry, 1984). While the amplitude of this gravitationally induced bispectrum is generally much higher than that expected from PNG, their different scale-dependence can be used to disentangle the primordial contribution by considering only specific triangle configurations.

2.1.4 Non-linear structure formation

Linear structure formation is only valid if the density contrast is small, which applies to early times, or sufficiently smoothed large scales modes. It is therefore desirable to develop analytical techniques that extend the validity of linear theory towards smaller scales. The full non-linear evolution of Equations (2.3) - (2.5) may only be followed using numerical simulations (Chapter 3). Figure 2.1 shows the matter power spectrum at $z = 0$ from linear theory and the non-linear prediction using a fit to simulations (Smith et al., 2003). While both predictions agree well on large scales, the small-scale clustering is strongly enhanced in the non-linear case. Why not always use a fitting function instead of trying to find analytical descriptions? Unfortunately, it is difficult to predict how accurate such a fit is

---

[3] Assuming ergodicity and averaging over a large enough region, ensemble and spatial average are equivalent.

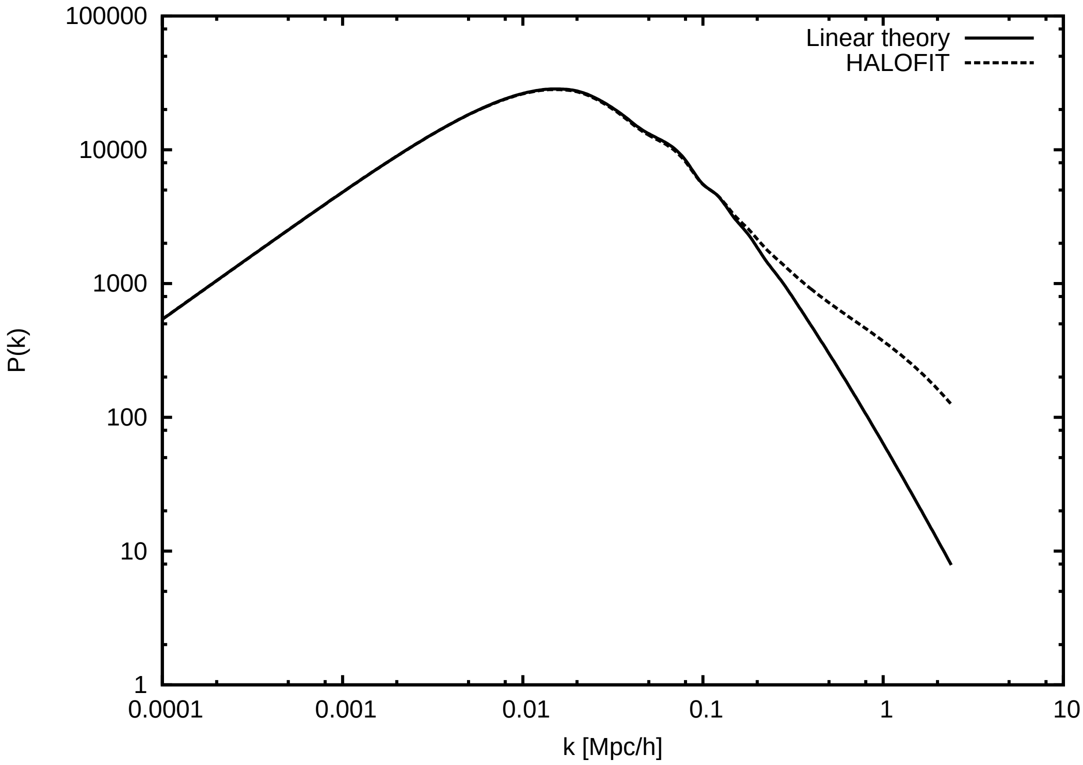

Figure 2.1: Matter power spectrum at $z = 0$ calculated from linear theory, and a non-linear prediction based on fits to numerical simulations (Smith et al., 2003). On large scales (small k), the power spectrum is essentially "primordial" and thus proportional to k (see section 2.3), while the small-scale behavior is dictated by the transfer function, so $P_{\rm lin}(k) \sim k^{-3}$.

when cosmological parameters used in the simulation are changed, and Running simulations for every possible parameter combinations is prohibitively expensive in terms of computing time.

In this thesis, we will focus on Standard Perturbation Theory (SPT) which is probably the most widely used extension to linear theory (see e. g. the review by Bernardeau et al. 2002). The successes and failures of SPT will be discussed in detail in Chapter 4, where we additionally show a new way of testing its accuracy against numerical simulations. Here we just give a brief description of the basic idea, which is to expand the density contrast in a series of higher-order terms, which are related to the linear solution by convolving $\delta_{\rm lin}$ with a set of kernel functions. As shown in Appendix A, these kernels can be calculated recursively from the Euler and Continuity Equation. SPT can extend the validity of linear theory to "mildly non-linear scales" ($k \lesssim 0.5\,h\,{\rm Mpc}^{-1}$), but suffers from convergence issues especially at late times. Other more complicated theories are able to improve the small-scale behavior using techniques like resummation, renormalization, path integrals, or by including additional short-range terms in the equations of motion (Carlson et al. 2009, Carrasco et al. 2013; and references therein).

2.2 Dark matter halos

Eventually, gravitational collapse leads to the formation of bound structures that decouple from the Hubble expansion; these objects are called *dark matter halos*.

A simple analytical model to describe halo formation is the spherical collapse model (Peebles, 1967; Gunn & Gott, 1972). Consider a spherical region which initially grows with the expanding universe. If the region was initially overdense, it will begin to collapse under its own gravity, provided that its density contrast exceeds a critical threshold value. The spherical collapse model predicts that regions with an initial overdensity $(\Delta\rho/\bar{\rho})_\mathrm{i}$ collapse into objects with final density $(\Delta\rho/\bar{\rho})_\mathrm{f} \sim 200$, after the Universe has expanded by a factor $a_f/a_i \sim 1.68\,(\Delta\rho/\bar{\rho})_\mathrm{i}$. While these values are derived assuming an EdS Universe ($\Omega_\mathrm{m} = \Omega_\mathrm{tot} = 1$), they are still reasonably valid for ΛCDM as well. During the collapse, the halo will eventually reach an equilibrium state through violent relaxation. Note that energy conservation dictates that the halo will stop collapsing at a minimum radius, called the *virial radius*, which is about 1/2 of the maximum radius before collapse.

The spherical collapse model thus relates the density contrast to the epoch of collapse, and this can be used to calculate the fraction of mass in collapsed objects at a certain time t, based on a given probability distribution for the density field (Press & Schechter, 1974). This probability can be transformed into a halo *mass function*, i. e. the mean number density of halos in a given mass interval. Independently on the exact values of the cosmological parameters (as long as the dark matter is "cold", Bhattacharya et al. 2011), the mass function can be parameterised as

$$\frac{\mathrm{d}\bar{n}_\mathrm{h}}{\mathrm{d}M}(M,z) = f(\sigma)\frac{\bar{\rho}_\mathrm{m}}{M}\frac{\mathrm{d}\ln\sigma^{-1}}{\mathrm{d}M}, \tag{2.16}$$

where $\sigma = \sigma(M,z)$ is the root mean square of the linear density field on mass scale M, e. g. $M = 4\pi\bar{\rho}_\mathrm{m}R^3/3$ for a spherical top hat filter. The multiplicity function $f(\sigma)$ can either be predicted by theory or by fitting the results of a simulation (reaching an accuracy of $< 5\%$, Tinker et al. 2008). The result from Press & Schechter (1974) is

$$f(\sigma) = \sqrt{\frac{2}{\pi}}\frac{\delta_c}{\sigma}\exp\left(\frac{-\delta_\mathrm{c}^2}{2\sigma^2}\right), \tag{2.17}$$

where $\delta_\mathrm{c} \sim 1.68$ is the density threshold from the spherical collapse model.

The simple Press-Schechter model is in reasonable agreement with numerical simulations, but overpredicts (underpredicts) the abundance of low-mass (high-mass) halos. Theoretical extensions include e. g. treating the crossing of the density threshold as a random walk (Bond et al., 1991; Lacey & Cole, 1993), considering an ellipsoidal collapse (Sheth et al., 2001), or allowing for non-Gaussian initial conditions (Maggiore & Riotto, 2010). Popular fitting functions from simulations include those by Sheth & Tormen (1999), Jenkins et al. (2001), Tinker et al. (2008) and Pillepich et al. (2010).

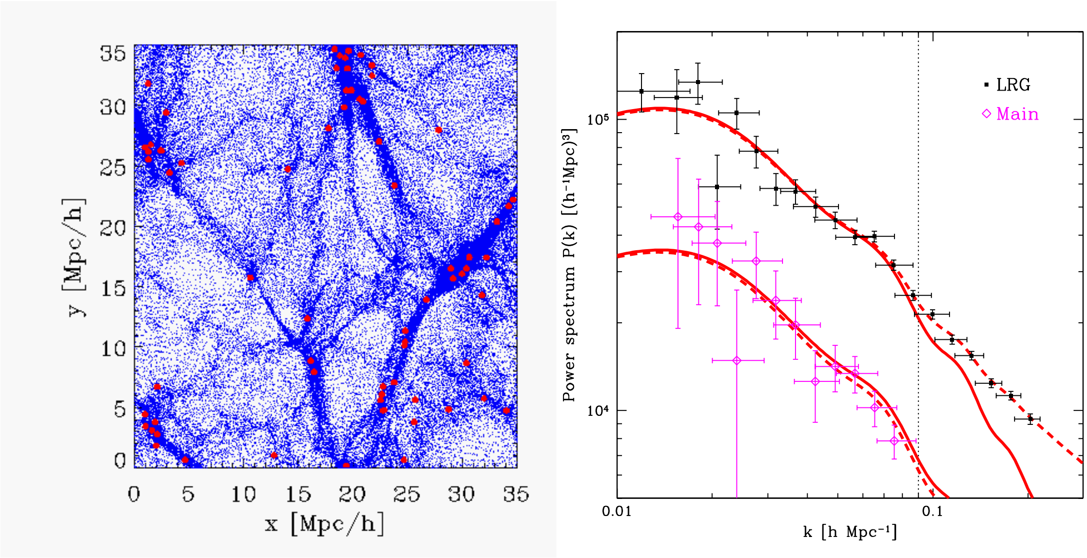

Figure 2.2: Left: Illustration of the biasing effect. Shown in blue is the dark matter distribution from a numerical simulation (2D projection of a 3D slice of thickness $10\,h^{-1}$ Mpc), while dark matter halos above a certain mass are shown as red points. The correlation between matter overdensity and the number of dark matter halos is clearly non-trivial. Right: Galaxy power spectra measured for two different populations (points with errorbars), while lines show fits to the predicted matter power spectrum when biasing is taken into account (solid: linear theory, dashed: including phenomenological non-linear corrections). [Right panel taken from Tegmark et al. (2006).]

2.2.1 Halo and galaxy bias

Dark matter overdensities can already grow prior to recombination while the baryons are still coupled to the photons. After recombination, the baryonic components will eventually fall into the dark matter potential wells and galaxy formation will be triggered inside halos. In the hierarchical structure formation scenario, collapse happens on small scales first, and larger structures like groups and clusters form through the merging of smaller objects. Since dark matter halos are the result of collapsing overdensities, the spatial clustering of halos should be related to the spatial clustering of the matter density field. This relationship is referred to as *biasing*, and halos are called *biased tracers* of the dark matter field.

However, most observations will actually measure the photons emitted by the baryonic component (stars, galaxies, etc.), and one large dark matter halo may contain several galaxies. A distinction has thus to be made between *halo bias* and *galaxy bias*, and we will concentrate on halo bias because it is more straightforward to deal with analytically and in the context of DM simulations. Given a model of halo bias, the effect of galaxy formation can for example be included by additional models of the *halo occupation distribution* (e. g. Kauffmann et al., 1997; Benson et al., 2000; Kravtsov et al., 2004, and references therein).

Examples of the effects of biasing are illustrated in Figure 2.2. The left panel shows the matter field of a dark matter-only numerical simulation[4] (blue dots) and

---

[4] For details on numerical simulations of structure formation and halo finding see Chapter 3.

the position of collapsed halos above a certain mass threshold in the same volume (red points). As expected, the number of halos is higher in overdense regions at the intersection of filaments, and the overall structure of the dark matter field is traced by collapsed objects. However, it is obvious that only observing the point distribution of halos leads to an incomplete picture of the density field, especially inside voids (underdense regions). The right panel shows the power spectrum measured for two different galaxy populations in the Sloan Digital Sky Survey (points with error bars, Tegmark et al., 2006). Note that the different lines represent fits to the data using the matter power spectrum from linear theory (solid), and including non-linear corrections to both the power spectrum and bias model on small scales (dashed, Cole et al. 2005). One sample includes only red galaxies with high luminosities (LRG=Luminous Red Galaxies), which are often found inside galaxy clusters, while the “Main” sample consists of a mix of galaxies. These examples illustrate that bias can depend on mass, redshift, galaxy type or environment, as well as the detailed formation history of the halo (e. g. Sheth & Tormen, 2004; Gao et al., 2005; Zehavi et al., 2011).

Historically, bias was first discussed in detail by Kaiser (1984) to explain the observed difference in correlation length between galaxies and galaxy clusters. Starting from the assumption that galaxy clusters form at maxima of the (Gaussian) density field, Kaiser was able to calculate the 2-point correlation function $\xi_{>\nu}$ for regions above a density threshold $\nu\sigma$ (σ being the r. m. s. density of a region with radius R, and ν a positive number)

$$\xi_{>\nu}(r) \approx \frac{\nu^2}{\sigma^2}\xi(r), \tag{2.18}$$

which is related to the 2-point correlation function of the underlying density field, $\xi(r)$. The amplitude of the bias $b_\nu^2 = \nu^2/\sigma^2$ can be predicted by smoothing the density field on a given mass scale (e. g. based on the spherical collapse model discussed above), and the results were in reasonable agreement with observations at that time, finding $b_\nu \sim 5$ for massive clusters.

Predictions for the bias can also be derived from the mass function. By splitting the density field in a long-range and a short-range part $\delta(\boldsymbol{x}) = \delta_\mathrm{L}(\boldsymbol{x}) + \delta_\mathrm{S}(\boldsymbol{x})$ (the *peak-background split*, Bardeen et al. 1986), the bias function can be related to the halo mass function (Cole & Kaiser, 1989)

$$b_\mathrm{pb} \equiv \frac{\partial \ln \bar{n}_\mathrm{h}(M,z)}{\partial \delta_\mathrm{L}} + 1, \tag{2.19}$$

where $\delta_\mathrm{L} = \nu\sigma(M,z)$ in the peak-background split. The physical interpretation is that a large-scale mode with e. g. $\delta_\mathrm{L} > 1$ lowers the threshold for collapse, leading to a change in the local number density of halos at a given mass scale. Unfortunately, neither the mass function nor δ_L is easily accessible from observations, but Equation (2.19) can be applied to one of the theoretical or simulated mass functions discussed in the previous section. Note that one can either consider the density contrast in Lagrangian space or in Eulerian space, where the former is applied in the initial conditions (section 2.4) and the latter at later times (Chapter 4). On large scales the Eulerian and Lagrangian bias are related by $b_{\mathrm{E}} = b_{\mathrm{L}} + 1$ (Mo & White, 1996; Catelan et al., 1998, 2000).

A more phenomenological approach is to relate the density field of the tracers (i. e. a smoothed representation of the spatial distribution of halos or galaxies) to the dark matter by a general functional B of the matter density contrast δ,

$$\delta_{\mathrm{tracer}}(\boldsymbol{x}) = B\left[\delta\right]. \tag{2.20}$$

In principle, the functional B can take many forms, but the most common case is that of *local* Eulerian bias (Dekel & Rees, 1987; Fry & Gaztanaga, 1993), where the tracer field depends only on the matter field at the same location, i. e. $\delta(\boldsymbol{x})$. A natural extension of this model is non-deterministic bias which includes some scatter around a mean bias relation (Scherrer & Weinberg, 1998; Dekel & Lahav, 1999). Expressing the bias through Equation (2.20) allows for another interpretation of B: any bias function could be expressed by a set of bias parameters that should be *measured*, not predicted from theory (e. g. McDonald, 2006). This interpretation does not require the knowledge of whichever physical processes are responsible for the exact form of halo or galaxy bias. We will return to this point in Chapter 4.

When the bias is simply a constant factor b, the correlation function for δ_{tracer} is just a scaled version of the corresponding dark matter statistic (the same is true for Equation (2.18) on large scales). Transforming to Fourier space, this *linear* biasing results in a tracer power spectrum

$$P_{\mathrm{tracer}}(k) = b^2 P(k). \tag{2.21}$$

Linear biasing is often used when extracting cosmological parameters from the observed galaxy power spectrum or correlation function, but this simple relation is only valid in the large-scale limit (Bardeen et al., 1986; Scherrer & Weinberg, 1998). We stress again that b generally depends on e. g. halo mass, galaxy type, or redshift.

Analytical models for the bias function give reasonable predictions on large scales, but often fail when confronted with the details of small-scale clustering. This is especially troubling because the constraining power resides in the small scales of the power spectrum where the statistical error bars are much smaller. Paired with the fact that smaller scales are more difficult to model even for the dark matter field alone due to non-linearities, biasing constitutes a severe systematic effect that stands in the way of utilising the full statistical potential of current and future large field galaxy surveys.

2.3 Initial conditions: Inflation

The idea of an early epoch of exponential expansion is usually attributed to Guth (1981), followed up by a flurry of related publications, e. g. Linde (1982); Albrecht

& Steinhardt (1982); Hawking et al. (1982); Starobinsky (1982) and Wald (1983). For a summary of detailed calculations we refer to Mukhanov et al. (1992); Maldacena (2003) and Baumann (2009).

Flatness problem From the first FE we can describe the curvature as a function of time and this can be expressed by re-arranging Equation (1.25):

$$1 - \Omega_{\mathrm{tot}}(a) = -\frac{k}{(aH)^2}, \tag{2.22}$$

where aH is called the comoving Hubble radius. This equation has two immediate implications: The sign of the curvature is conserved, so an open Universe can not turn into a closed Universe and vice versa. Second, any initial deviation from flatness will grow with time if $(Ha)^{-1}$ increases, and this can easily be shown to be the case during both matter and radiation domination. The fact that Ω_{tot} is very close to 1 at the current epoch (see Table 2.1) then requires that the Universe has to be almost perfectly flat in the initial conditions. However, an early epoch where $(Ha)^{-1}$ actually decreases (meaning that the Hubble radius increases) would solve this apparent fine-tuning problem.

Horizon problem The Cosmic Microwave Background radiation is isotropic up to 1 part in 10^5 over the whole sky. The particle horizon at recombination – representing the maximum distance that particles could have interacted on until that moment – can be translated into an angle on the sky by using the angular diameter distance (Equations (1.13) - (1.14)). The resulting angle of $\sim 2^\circ$ contradicts the observed isotropy unless causality is violated or if the Universe was somehow initially homogeneous on larger scales. This presents the second fine-tuning problem of the initial conditions. However, as with the Flatness problem, an increasing Hubble horizon towards early times would allow for causal interaction on much larger scales. Put in another way, the homogeneity of the largest scales accessible to us today constrains the minimal amount of expansion during the inflationary period.

The origin of structures With Inflation, we also obtain a natural explanation for the origin of the primordial perturbations: initially tiny quantum fluctuations are blown up by the rapidly increasing scale factor until they become “classical” perturbations when their fluctuation time-scale exceeds the Hubble time. The simplest models of Inflation result in fluctuations that are nearly perfectly Gaussian (e. g. Gangui et al., 1994), in concordance with the statistics of the CMB, and any deviation from exact Gaussianity can put constraints on the type of inflationary physics (section 2.4 and Chapter 5). Inflation also explains the absence of observational evidence for topological defects like magnetic monopoles, which could be produced at temperatures above the Grand Unified Theory (GUT) scale (Preskill, 1979). Their density will be diluted by the exponential expansion, making it essentially impossible to observe them.

General requirement for Inflation The time derivative of the comoving Hubble horizon is directly related to the acceleration,

$$\frac{\mathrm{d}}{\mathrm{d}t}(aH)^{-1} = -\frac{\ddot{a}}{(aH)^2}, \tag{2.23}$$

so if the left hand side is < 0 the acceleration has to be positive. Using the identity $\ddot{a}/a = \dot{H} + H^2$, we can write the acceleration as

$$\frac{\ddot{a}}{a} = H^2(1-\epsilon), \tag{2.24}$$

where we have defined $\epsilon = -\dot{H}/H^2$. For positive acceleration, $\epsilon < 1$. By inserting both Friedmann equations into Equation (2.24), we can write ϵ in terms of the EoS parameter w:

$$\epsilon = \frac{3}{2}(w+1). \tag{2.25}$$

This means we have found 3 equivalent conditions for Inflation: a decreasing Hubble radius, accelerated expansion and an EoS parameter $w < -1/3$, which can all be satisfied by requiring $\epsilon < 1$. In the following, we will attempt to connect these conditions to the physics of the early Universe.

2.3.1 The action principle

In this section we revisit the underlying principles of field theory, avoiding however any lengthy derivations. The basis of the Einstein field equations (Equation (1.17)) is the requirement that the Einstein-Hilbert action,

$$S_{\mathrm{EH}} = -\frac{1}{16\pi G}\int \mathrm{d}^4x\,\sqrt{-g}\,(R + \mathcal{L}_{\mathrm{M}}), \tag{2.26}$$

is invariant under an infinitesimal variation of the metric. The different contributions to the action are $g = \det g_{\mu\nu}$, the Ricci scalar R describing the curvature, and the *Lagrangian* density $\mathcal{L}_{\mathrm{M}}$ which describes any matter field present at early times (this could also include a cosmological constant). The field equations can then be derived by requiring $\delta[S_{\mathrm{EH}}] = 0$ under $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$.

In general, any Lagrangian can be described by a kinetic term and a potential term which determine the dynamical evolution of the field, and the resulting field equations depend on its functional form. Even without knowing the detailed physical processes involved, we can make a few statements about the Lagrangian at early times: We have found that the conditions for Inflation are $\ddot{a} > 0$, or equivalently $w < -1/3$, and this criterion could be fulfilled for example by a cosmological constant with $w = -1$ (this was the original proposition by Guth 1981). A more general idea is to consider a new scalar field[5] $\phi(\boldsymbol{x}, t)$ called the *Inflaton*, that provides the dominant energy content in the primordial Universe.

If this field initially resides in a non-vacuum state inside the potential, it will

---

[5] We briefly discuss multi-field models and other non-standard contributions in section 2.4.

"roll" towards its true minimum, while driving the accelerated expansion of the Universe (see Figure 2.3). Once it reaches the potential minimum, its energy has to be transformed into standard model particles during an epoch called reheating. Since the energy scales at such early times are far beyond those accessible to current particle physics experiments, the exact process of reheating is unknown, but likely also depends on the details of the previous inflationary period (e. g. Bassett et al., 2006, and references therein). The whole process of Inflation and reheating should therefore be considered an effective theory, with a scalar field designed such that it matches the observations of the CMB temperature fluctuations. However, we will see later that inflationary models also make *predictions* that can be used to constrain the almost infinite set of models that are able to solve the 3 puzzles explained at the beginning of this section.

2.3.2 Single field slow-roll Inflation

The Inflaton field can be incorporated in the Einstein-Hilbert action by adding the Lagrangian

$$\mathcal{L}_\phi = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi), \tag{2.27}$$

which consists of a "standard" kinetic term, and a potential term $V(\phi)$ which describes the self-interaction of the field. This Lagrangian is said to be "minimally coupled" to gravity because it is a simple additive term to the Einstein-Hilbert action and does not directly multiply R. The energy momentum tensor is given by

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - g^{\mu\nu}\mathcal{L}_\phi, \tag{2.28}$$

which allows us to derive the density and pressure of ϕ:

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi) + \frac{1}{2}(\nabla\phi)^2, \tag{2.29}$$

$$p = \frac{1}{2}\dot{\phi}^2 - V(\phi) + \frac{1}{6}(\nabla\phi)^2. \tag{2.30}$$

Since the Universe should be homogeneous within the small patch that is inflated, the assumption that ϕ only depends on t and not on $\boldsymbol{x}$ is justified, and all terms with ∇ in the previous equations vanish. In addition, we also neglect all contributions from spatial curvature (justified by Equation (2.22)). With ρ and p, we can write down the Friedmann Equations for ϕ:

$$H^2 = \frac{8\pi G}{3}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right), \tag{2.31}$$

and the equation of motion follows from the general requirement that $T^{\mu\nu}$ is conserved

$$\ddot{\phi} + 3H\dot{\phi} + V_{,\phi} = 0, \tag{2.32}$$

where $V_{,\phi}$ denotes the derivative of V with respect to ϕ.

By calculating the time derivative of (2.31), inserting (2.32), and using (2.24)

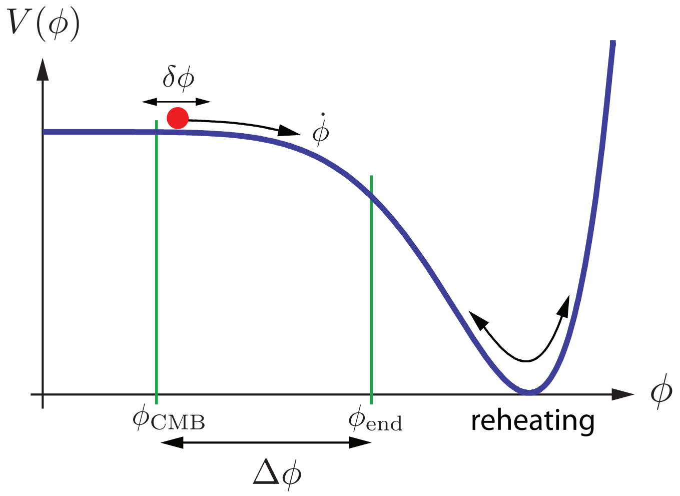

Figure 2.3: Potential for slow-roll Inflation, where the value of the field ϕ is used instead a time coordinate. Inflation happens while the scalar field rolls down towards its true minimum, where its energy is transformed into standard model particles. ϕ_{CMB} marks the creation of the CMB fluctuations by quantum fluctuations $\delta\phi$ around the mean value, and ϕ_{end} marks the endpoint of the slow-roll period. The time between ϕ_{CMB} and ϕ_{end} has to be large enough to allow the Universe to expand by about 60 e-folds in order to solve the Horizon Problem. [Figure taken from Baumann (2009).]

we obtain an expression for ϵ, which depends on the time derivative of ϕ,

$$\epsilon = \frac{1}{2}\frac{\dot{\phi}^2}{H^2}. \tag{2.33}$$

Since we required $\epsilon < 1$, this gives us the first condition on the Inflaton field in order to have an early epoch of accelerated expansion. If we also want Inflation to last a sufficiently long time (e. g. to solve the Horizon problem), we need a condition on $\dot{\epsilon}$ as well. Taking the time derivative of the previous equation, a little algebra yields

$$\eta \equiv \epsilon - \frac{\dot{\epsilon}}{2\epsilon H} = -\frac{\ddot{\phi}}{H\dot{\phi}}, \tag{2.34}$$

which relates the change in ϵ to the second time derivative of ϕ by defining a second parameter $|\eta| < 1$.

Now consider the case where the potential dominates over the kinetic term, $\dot{\phi} \ll V(\phi)$, which immediately implies $\rho \approx -p$, reminiscent of Λ at late times. Since $H^2 \sim V$ as well, Equation (2.33) tells us that in this case $\epsilon \ll 1$ and $|\eta| \ll 1$. Since ϵ is by definition also proportional to $\dot{H}$, we find $\rho \sim H^2 \sim V(\phi) \sim$ const., which is referred to as "slow-roll" Inflation (illustrated in Figure 2.3). In this model, inflationary expansion occurs as long as the field is rolling down the slowly changing potential slope. Eventually, the slow-roll phase has to transition into the reheating phase, and H will no longer be constant.

A constant Hubble parameter implies exponential expansion with $a(t) \sim \mathrm{e}^{Ht}$. In order to explain the amplitude of the initial fluctuations and to ensure that the largest scales accessible to current observations have been in causal contact at some time in the past, $N \sim 60$ e-folds of inflation are required (Liddle & Lyth, 1993).

2.3.3 Seeds for structure formation

During Inflation, the Inflaton field has local quantum fluctuations $\Delta\phi(\boldsymbol{x}, t)$ around its mean value $\phi(t)$. These fluctuations translate into a locally different time evolution, where each patch evolves faster or slower depending on the sign of the fluctuation. This creates local density perturbations, which become classical perturbations after horizon exit where the fluctuation time-scale is longer than the Hubble time. In general, a change in the density will induce a perturbation in the metric $g^{\mu\nu}$ because both are coupled through the field equations.

When performing any calculation involving inhomogeneities, care must be taken when defining a coordinate system (gauge).[6] In the unperturbed Universe the choice of the comoving observer was straightforward, as it preserved the conditions for homogeneity and isotropy. In a perturbed Universe that choice is no longer unique, but often calculations can be simplified by selecting the right gauge. Note however that GR implies that any *physical* perturbation must be gauge independent. Instead of choosing a specific gauge, it is therefore also sensible to define gauge-independent quantities and try to find their relation to e. g. density perturbations that we are interested in.

Furthermore, it can be shown that to linear order, scalar, vector and tensor perturbations[7] evolve independently from each other (e. g. Baumann 2009), so we can consider each in turn. In most models, vector perturbations are not generated with large amplitudes, and they will also decay with the subsequent expansion, so they can be neglected. We are interested in the creation of classical density perturbations, which are sourced by scalar perturbations, while primordial gravitational waves are sourced by tensor perturbations (e. g. Bardeen, 1980).

Starting from the Einstein-Hilbert action, the amplitude of scalar perturbations can be calculated as follows: We consider the comoving curvature perturbation $\mathcal{R}(\boldsymbol{x}, t)$, which measures the spatial curvature of constant-ϕ hypersurfaces. In addition to being gauge-independent, it can be shown that $\mathcal{R}$ remains constant outside of the horizon, making it sufficient to calculate the amplitude at the time of horizon crossing where $k = Ha$. The curvature perturbations act like a potential that sources the density perturbations (see below). Once we have determined their amplitude, the subsequent evolution is given by the transfer function and the growth factor as described in section 2.1.

Treating the metric perturbations as small allows us to expand the action up to second order in $\mathcal{R}$. The requirement that the action vanishes leads to the Mukhanov Equation (Mukhanov et al., 1992), which describes the time evolution of the curvature perturbations in a form similar to a damped harmonic oscillator. In general, the solutions to this equation have to be found numerically, because

---

[6] See Kodama & Sasaki (1984) for a discussion of the different gauge choices.

[7] The different classes are defined by the way they behave under coordinate transformations.

they depend on the time evolution of both ϕ and H. Additionally, the final result depends on the vacuum state which sources the fluctuations. A common choice is that of a thermal state without particle production (Christensen, 1976; Bunch & Davies, 1978).

By defining the dimensionless power spectrum

$$\Delta^2(k) = \frac{k^3}{(2\pi)^3} P(k), \tag{2.35}$$

and evaluating the solutions of $\mathcal{R}$ at horizon crossing (denoted by $*$), the steps outlined above lead to an expression for the curvature power spectrum

$$\Delta^2_{\mathcal{R}}(k) = \frac{1}{8\pi^2} \frac{H_*^2}{M_{\mathrm{pl}}^2} \frac{1}{\epsilon_*}, \tag{2.36}$$

where $M_{\mathrm{pl}} = (8\pi G)^{-1/2}$ is the Planck mass in natural units.[8]

This rather simple expression contains a few interesting properties. If we write $\Delta^2_{\mathcal{R}}(k)$ as a power law

$$\Delta^2_{\mathcal{R}}(k) \sim k^{n_s - 1}, \tag{2.37}$$

slow-roll Inflation *predicts* a scalar index $n_{\mathrm{s}} = 1 + 2\eta_* - 4\epsilon_* \approx 1$. This scale-independence of the initial curvature perturbations is often referred to as a Harrison-Zeldovich power spectrum, predicted by Harrison (1970) and Zel'dovich (1972) based on arguments about entropy and the initial amplitude required for galaxy formation, roughly a decade before the work of Guth (1981). The fact that n_{s} deviates from 1 can be qualitatively understood by remembering that H and ϵ are nearly – but not exactly – constant during Inflation. Since the amplitude of each mode is determined by its amplitude at horizon crossing, changes in H translate into a scale-dependence in the power spectrum. The deviation of n_s from unity has already been confirmed to an accuracy of 5σ, providing strong evidence for the inflationary paradigm (Planck Collaboration et al., 2013a).

Roughly speaking[9], the relation between $\mathcal{R}$ and the density perturbations δ is given by the Poisson equation and additional factors of order unity. In Fourier space this means $\tilde{\delta} \sim \tilde{\mathcal{R}} k^2$, so we finally find for the power spectrum of *primordial* density perturbations

$$P_\delta(k) = A(k_{\mathrm{p}}) \left(\frac{k}{k_{\mathrm{p}}} \right)^{n_{\mathrm{s}}}, \tag{2.38}$$

where the amplitude $A \sim H_{\mathrm{p}}^2/(8\pi G)$ at an arbitrary pivot scale $k_{\mathrm{p}} = aH_p$ can be measured from the CMB. The primordial dependence on k can still be observed in the matter power spectrum on the largest scales where the transfer function is constant (see Figure 2.1).

---

[8] For a detailed derivation see e. g. Baumann (2009), Knobel (2012), and references therein.

[9] The full justification extends over several pages in Weinberg (2008) and Knobel (2012).

2.4 Primordial non-Gaussianity

In the standard slow-roll Inflation model discussed above, the probability distribution of the initial perturbations is extremely close to a Gaussian (e. g. Gangui et al., 1994). Any deviation of the initial perturbations from a Gaussian distribution would therefore invalidate one or more of the assumptions of the Inflation model. To distinguish this effect from the late-time non-Gaussianity introduced by non-linear evolution, any non-Gaussianity present in the initial conditions is called Primordial non-Gaussianity (PNG).

While it is easy to quantify "Gaussianity", the possibilities for a probability distribution to be "non-Gaussian" are infinite. As already discussed in section 2.1.3, higher order statistics like the bispectrum vanish only in the Gaussian case. It is therefore useful to distinguish between different classes of PNG, commonly expressed through bispectrum *templates*. These templates are defined by the triangle configurations that contribute most of the signal to the bispectrum (parameterised by three modes k_1, k_2 and k_3 shown in Figure 2.4). The tightest constraints on PNG currently come from the CMB (e. g. Bennett et al., 2012), but certain types of PNG should also leave a detectable signal in the halo distribution at late times. This signal comes in the form of a *scale-dependent bias* on large scales, in contrast to the Gaussian case where the bias asymptotes to a constant value (see section 2.2.1). For a recent review on the effects of PNG on large-scale structure see Desjacques & Seljak (2010a).

The bispectrum for single-field slow-roll models peaks in the *squeezed* configuration, characterised by $k_1 \approx k_2 \gg k_3$, however the predicted amplitude is orders of magnitude below the sensitivity of current experiments (Acquaviva et al., 2003; Maldacena, 2003). A large amplitude in the squeezed configuration is predicted for example when multiple fields are present during Inflation (Linde & Mukhanov, 1997). The expected signal in these models only depends on the local amplitude of the initial perturbations, and is therefore also called the *local* model.

Additional classes of PNG models can be constructed by relaxing one (or more) of the other assumptions made in section 2.3.2. Kinetic terms with higher-order derivatives in the Lagrangian generate the largest signal for *equilateral* triangles (Creminelli et al., 2006). Finally, deviations from the standard Bunch-Davies vacuum assumption give rise to *folded* type PNG (Meerburg et al., 2009). The so-called *orthogonal* template can be constructed from a combination of equilateral and folded shapes (Senatore et al., 2010). Of course once several effects are combined, the resulting bispectrum will also be a combination of the different templates. In that case no analytical expression for the bispectrum exists and numerical predictions can become quite involved (Wagner et al., 2010).

In the case of PNG it is common to express the primordial fluctuations by the Bardeen curvature potential Φ_B, which is related to the comoving curvature perturbations $\mathcal{R}$ and the Newtonian gravitational potential Φ_G after reheating by $\Phi_\mathrm{B} = 3\mathcal{R}/5 = -\Phi_\mathrm{G}$ (Bardeen, 1980). In terms of Φ_B, the local model can be

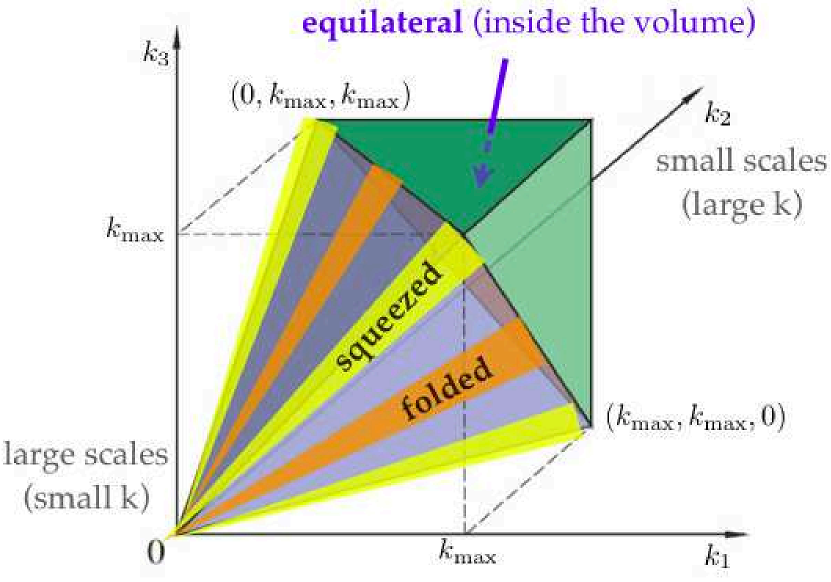

Figure 2.4: Representation of different triangle configurations for the bispectrum. The squeezed and folded configurations lie on the faces of the "tetrapid", while equilateral triangles are located inside the volume. [Figure taken from Schmittfull et al. (2013).]

expressed by

$$\Phi_{\rm B}(\boldsymbol{x}) = \phi(\boldsymbol{x}) + f_{\rm NL}\left[\phi(\boldsymbol{x})^2 - \langle\phi^2\rangle\right] + g_{\rm NL}\left[\phi(\boldsymbol{x})^3 - 3\langle\phi^2\rangle\phi(\boldsymbol{x})\right] + \ldots, \tag{2.39}$$

where the non-Gaussian potential is written as a local expansion of the Gaussian potential $\phi(\boldsymbol{x})$ with coefficients $f_{\rm NL}$, $g_{\rm NL}$, ...and the terms proportional to the ensemble averages ensure that the resulting non-Gaussian density field still has a zero mean. In linear theory, ϕ is of order 10^{-5}, while current constraints predict $f_{\rm NL} \ll 100$, emphasizing the smallness of the effect. The coefficient $f_{\rm NL}$, $g_{\rm NL}$ are usually assumed to be real-valued constants, but models with scale-dependent $f_{\rm NL}$ have also been discussed, and in that case the multiplications in Equation (2.39) have to be replaced by convolutions (e. g. Becker et al., 2012, and references therein).

To first order in $f_{\rm NL}$ ($g_{\rm NL}$) the higher order terms in Equation (2.39) generate a bispectrum (trispectrum) that is proportional to $f_{\rm NL}$ ($g_{\rm NL}$). In addition, local PNG leads to a scale-dependent bias on large scales, both for $f_{\rm NL}$ (Dalal et al., 2008; Matarrese & Verde, 2008; Slosar et al., 2008) and $g_{\rm NL}$ (Desjacques & Seljak, 2010b; Smith et al., 2012), which we will discuss in the following.

The origin of the scale-dependence can be qualitatively understood by applying the Laplacian to Equation (2.39). For the $f_{\rm NL}$-term, the change in the potential translates into a change in the density contrast

$$\delta_{\rm PNG} \approx \delta\left[1 + 2f_{\rm NL}\phi\right], \tag{2.40}$$

where the approximate sign indicates that we have neglected any contributions from the constant term in Equation (2.39), because we are interested only in the peaks of the density field where halos form (Dalal et al., 2008). The second term in Equation (2.40) is proportional to ϕ, and thus to the density contrast by k^{-2}. By assuming a local Lagrangian bias $\delta_\mathrm{h} = b_\mathrm{L}\,\delta$ (suitable for the initial conditions considered here), the change in the bias induced by f_NL is[10]

$$\Delta b \sim 2 f_\mathrm{NL}\, b_\mathrm{L}\, k^{-2}. \tag{2.41}$$

The full derivation, including the correct normalisation factors and the transfer function to account for the evolution of the density perturbations, then leads to

$$\Delta b(k,z) = f_\mathrm{NL}(b_\mathrm{E}-1)\frac{3\,\Omega_\mathrm{m} H_0^2 \delta_\mathrm{c}}{D(z)T(k)k^2}, \tag{2.42}$$

where we have additionally switched to the Eulerian bias (see section 2.2.1) because we want to measure the effect of PNG from the evolved large-scale structure.

For the g_NL term, the contribution to the scale-dependent bias actually comes from the term proportional to $\langle\phi^2\rangle\phi$. Just like for f_NL, this term modifies the density contrast by adding a constant and therefore the scale dependence in the bias is the same. However, both terms are expected to show a different mass and redshift dependence (for a full discussion we refer to Smith et al. 2012 and Chapter 5). While most Inflation models predict either a dominant f_NL or a dominant g_NL term, there are theories where both contributions are important (e. g. the Curvaton model, Lyth et al., 2003). Current constraints on PNG parameters from the large-scale structure usually assume that only one of the parameters is different from 0. In Chapter 5 we investigate whether the degeneracy between the two contributions can be broken in the ideal case where we know the underlying values of f_NL and g_NL in a numerical simulation, and show that neglecting either term can lead to a highly biased estimate of the local PNG parameter.

As a closing remark, folded-type PNG also leads to a scale-dependent bias, $b(k) \sim k^{-1}$ on large scales, but the effect is orders of magnitude smaller than for the local type. For the equilateral template the bias is only changed by a small constant, which is therefore indistinguishable from the Gaussian case (Verde & Matarrese, 2009; Wagner & Verde, 2012).

2.5 Cosmological parameters

When work on this thesis was begun, the standard cosmological model was mainly constrained by the measurements of the WMAP satellite (in conjunction with estimates from other probes like Supernovae and galaxy clustering, see below). In March 2013, the results of the first year analysis of the *Planck* collaboration were published, which had measured the cosmological parameters with even greater

---

[10] Note that there is no minus sign because ϕ=-ϕ_G.

precision. While most of the values are consistent between the two probes, the increase in precision leads to slight tension in a few cases. This shows that while the combinations of different probes leads to an overall concordance for ΛCDM, they may still be affected by systematic effects. In Table 2.1, WMAP++ refers to using the CMB temperature and polarisation from the WMAP satellite, combined with estimates for H_0 and the BAO scale from large-scale structure (Hinshaw et al., 2012). The *Planck*++ estimate combined the CMB temperature from *Planck* with the polarisation of WMAP and the previous BAO estimate (Planck Collaboration et al., 2013a). Constraints on the primordial non-Gaussianity parameters will be discussed in Chapter 5.

Table 2.1: Selection of cosmological parameters from WMAP++ (Hinshaw et al., 2012) and *Planck*++ (Planck Collaboration et al., 2013a). The first set of 6 parameters can be determined directly from the data, while the others are derived using additional assumptions. All values are given at 68% confidence unless explicitly stated otherwise.

Symbol	WMAP++ (68%)	*Planck*++ (68%)	Description
$\Omega_\mathrm{c}h^2$	0.1153 ± 0.0019	0.1187 ± 0.0017	Cold dark matter
$\Omega_\mathrm{b}h^2$	0.02223 ± 0.00033	0.02214 ± 0.00024	Baryons
Ω_Λ	$0.7135^{+0.0095}_{-0.0096}$	0.692 ± 0.010	Dark energy
$10^9\Delta^2_\mathfrak{R}$	2.464 ± 0.072	$2.200 \pm 0.1(03)$	Scalar amplitude
n_s	0.9608 ± 0.0080	0.9608 ± 0.0054	Scalar index
$r_{0.002}$	< 0.13 (95%)	< 0.111 (95%)	Tensor amplitude
w	$-1.073^{+0.090}_{-0.089}$	$-1.13^{+0.24}_{-0.25}$ (95%)	EoS parameter for Λ
$100\,\Omega_\mathrm{k}$	$-0.27^{+0.39}_{-0.38}$	$-0.05^{+0.65}_{-0.66}$ (95%)	Curvature
H_0	69.32 ± 0.80	67.80 ± 0.77	Hubble rate [$\mathrm{km\,s^{-1}Mpc^{-1}}$]
σ_8	$0.820^{+0.013}_{-0.014}$	0.826 ± 0.012	Power spectrum amplitude
t_0	13.772 ± 0.059	13.798 ± 0.037	Age of the Universe [Gyr]

3 Numerical simulations

So far we have discussed linear structure formation and Standard Perturbation Theory as an extension to the slightly non-linear regime. In order to follow the fully non-linear evolution while sampling the entire phase-space volume, one has to resort to numerical techniques. During the last two decades, numerical simulations have become a powerful tool to study astronomical objects on a large range of scales. They are used extensively in all areas of astronomy and astrophysics, like the formation of single stars and planets, analysing the dynamics of star clusters, modelling hydrodynamical processes in galaxies like our Milky Way, and last but not least studying the large-scale structure of cosmological volumes. Each of these problems has to take into account different physical properties and make assumptions to simplify the problem at hand.

The following discussion will focus on cosmological N-body simulations, because these have been extensively used to obtain the results presented in this thesis. We will give a general description of the N-body technique and its implementation by the GADGET-2 and SUBFIND codes. The basic principle of an N-body simulation is that the object of interest is traced by a set of *particles* which interact with each other in a well-defined way. In this context, the term particle refers to the minimal unit of the simulation, not a fundamental particle in the usual physical sense. In fact, N-body particles are many orders of magnitudes heavier than fundamental particles, single stars, or even galaxies in case of cosmological volumes (see below).

In the case of a ΛCDM universe, structure formation is dominated by dark matter, which interacts only via gravity. This allows for very simple simulations of structure formation using only dark matter particles, without including the complicated physics of baryons. While baryons can significantly affect the dark matter distribution , these effects are important mostly on small scales (e. g. Stinson et al., 2006; Sales et al., 2007; Sawala et al., 2013, see also Chapter 6). While early attempts at simulating cosmologically relevant volumes were limited by very poor resolution, the rapid increase in computing power coupled with efficient algorithms has lead to a tremendous improvement in accuracy and predictive power. Current state-of-the art N-body simulations require the use of the largest high-performance computing facilities with millions of CPU hours[1], and total memory consumption of a few to several hundred terabyte (Kuhlen et al., 2012).

---

[1] The total number of CPU hours is the product of the actual code runtime with the number of cores used in parallel, where N_{core} is of the order several hundred to 10,000 for the largest simulations.

3.1 The GADGET-2 code

GADGET-2 is a publicly available, highly parallel code for numerical simulation of both collisionless dark matter particles and hydrodynamics to study gas and star formation. However, we will only use it to perform dark matter simulations and thus restrict the description to this case.

Given a set of initial conditions and cosmological parameters, it takes a density distribution represented by a system of *particles* and evolves this system forward through gravitational interaction. GADGET-2 is parallelised using MPI[2] routines, and it has been used extensively within the astronomical community to obtain dark matter simulations of single galaxies (e. g. the *Aquarius* project, Springel et al. 2008) up to cosmological simulations of the Hubble volume (e. g. the *Horizon Run*, Kim et al. 2011). Chapter 6 is based on simulations performed with a more recent version, GADGET-3, which differs from the previous version mainly in the efficiency of implementation, but not in the underlying concepts.

3.1.1 Description of the problem

The basic ingredients for an N-body simulation are the size of the (comoving) volume V that should be studied, the number of particles N used to trace the matter distribution and the underlying cosmological model, expressed by a set of parameters. The ensemble of particles is then evolved forward in time under the effects of gravity. The cosmological parameters determine the expansion history and the mass content of the virtual universe, while the box size $L = V^{1/3}$ and particle number N determine the accessible scales and the mass resolution.

The mass of the particles can be calculated by

$$m_{\mathrm{p}} = \rho_{\mathrm{crit}}\,\Omega_{\mathrm{m}}\,(L/N)^3, \tag{3.1}$$

where $\rho_{\mathrm{crit}} \sim 27.78 \cdot 10^{10}\,h^{-2}\,\mathrm{M}_\odot\,\mathrm{Mpc}^{-3}$ in cosmological units. For example, for a simulation with $L = 1000\,h^{-1}$Mpc and $N = 1024$, the particle mass is of the order of $7 \cdot 10^{10}\,h^{-1}\,\mathrm{M}_\odot$, only about 15 times lower than the mass of our own galaxy. This simulation could therefore only be used to study the clustering of massive objects on large scales. These types of simulations are used in Chapters 4 and 5. If one is instead interested in the internal dynamics of galaxy-sized objects, a much better mass resolution is required, achievable by using a larger number of particles (significantly increasing run-time, memory and storage requirements) and/or by reducing the volume (resulting in fewer objects). This type of high-resolution simulation is discussed in section 3.1.6 and used in Chapter 6.

Note on large box sizes The interaction between the dark matter particles is calculated using Newtonian gravity, instead of General Relativity (GR). This is straightforward when the boxes are small compared to the comoving horizon size. However, even when box sizes approach the size of the observable Universe, it

---

[2] Message Passing Interface, a commonly used parallelisation technique

appears that GR effects cancel out and so Newtonian gravity can be safely used (e. g. Chisari & Zaldarriaga, 2011; Green & Wald, 2012, and references therein); see e. g. Zhang (2011) or Flender & Schwarz (2012) for the opposing point of view.

For small N, the gravitational forces could be calculated directly for each pair of particles, but this direct approach scales as N^2, so it will be prohibitive for the large number of particles required to accurately trace cosmological volumes. Instead, the particle distribution is extrapolated to a uniform grid. From this density field, the gravitational potential and thus the forces at each grid point can then be calculated by solving the Poisson equation in comoving coordinates,

$$\nabla^2 \Phi(\boldsymbol{x}, t) = 4\pi G a^2(t) \left[\rho(\boldsymbol{x}, t) - \bar{\rho}(t)\right]. \tag{3.2}$$

Based on the cosmological principle, periodic boundary conditions are employed. These boundary conditions allow the use of Fourier transformations, which allows us to write an integral or gradient as a simple multiplication. Taking advantage of sophisticated numerical libraries like FFTW (Fastest Fourier Transform of the West; Frigo & Johnson 2005) leads to very efficient algorithms to solve the Poisson equation.

3.1.2 TreePM

GADGET-2 uses a TreePM code, a hybrid algorithm employing different ways of calculating the forces between particles on large and small scales. On large scales, the potential is calculated via the fast Particle-Mesh (PM) method, and on small scales via the slow but precise Barnes-Hut tree method.

Particle-mesh (PM) Method Starting from an initial particle distribution, the density field ρ is obtained by assigning the particles to a grid via interpolation. Then, the potential Φ can be calculated by inverting the Poisson equation (3.2), and the force is obtained by calculating the gradient of Φ, either by finite differencing in real-space, or multiplying by $\mathrm{i}\boldsymbol{k}$ in Fourier space (which is computationally more expensive but possibly more accurate, Ferrell & Bertschinger 1994). Finally, the force is applied to each particle using the same grid interpolation scheme to ensure that any artefacts like self-forces on particles vanish (Hockney & Eastwood, 1988). The PM algorithm is fast and low in memory consumption. However, the numerical accuracy quickly degrades for particle separations smaller than a few grid cells.

Tree method This method is also called the Barnes-Hut algorithm (Barnes & Hut, 1986). The main idea is to subdivide the particle distribution into a hierarchy of 3D cells called tree nodes. Each subnode, or *leaf* is constructed out of the parent node by subdividing it into 8 cells with each 1/2 of the length of the parent. Each leaf is subdivided again, until the lowest level of the tree is reached, where each leaf contains one or zero particles. In case of a homogeneous particle distribution, this would simply correspond to a cubic grid. For inhomogeneous distributions, the tree will be coarser in underdense regions and finer in overdense regions.

The force acting on a particle is then calculated by summing the partial forces from neighbouring tree nodes, which are obtained by multipole expansions. To lowest order, this means treating an ensemble of particles in an individual node (or leaf) as a single particle located at the center of mass. A tree node is "opened", i. e. the level of refinement is increased, when the separation between it and the particle for which the force needs be computed is less than a threshold value.

This threshold value needs to be chosen small enough to reduce force errors, but large enough so that the simulation is not slowed down too much. Typically, force errors of 1-2 percent are acceptable for cosmological simulations, and GADGET-2 stays below this limit even in the regime between the two methods (Springel, 2005). If the particles positions only change slightly between two timesteps, it is also not necessary to re-calculate the full tree decomposition, and small dynamical adjustments suffice to achieve the desired force accuracy (McMillan & Aarseth, 1993). Since the Tree algorithm is spatially adaptive, the force accuracy is not limited by the initial choice of a grid spacing as in the case of the PM method. A drawback of this method is that it becomes very memory intensive, so the TreePM hybrid algorithm combines the best of both worlds.

3.1.3 Time integration

Once the forces are computed, one can perform the time integration of the equations of motion (in comoving coordinates)

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \frac{\boldsymbol{v}}{a}, \qquad \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} + H\boldsymbol{v} = -\frac{\nabla\Phi(\boldsymbol{x},t)}{a}. \tag{3.3}$$

However, these equations are coupled: $\boldsymbol{x}$ depends on $\boldsymbol{v}$ and $\boldsymbol{v}$ depends on the force at position $\boldsymbol{x}$. Time integration is therefore done with a leap-frog[3] integration scheme, consisting of a series of *kick* and *drift* operations for each timestep. A common example would be the kick-drift-kick scheme, where first the velocity is advanced for half a time step (kick), then the position is updated with this new velocity (drift), and finally the second half of the velocity integration is performed (kick).

The choice of time step is crucial in a practical numerical implementation. If it is chosen too large, the forces will not be accurate enough, and inaccuracies will build up over time until the result no longer resembles the truth. Instead, if the time step is chosen too small, the code will slow down significantly. The latter case is especially problematic for high density regions. Modern codes therefore often allow for an adaptive time stepping criterion which evolves depending on the local density. In this way, only a subset of particles have very small timesteps, which significantly reduces the computational cost.

---

[3] In contrast to the commonly used Runge-Kutta integrator, the leap-frog scheme is symplectic, which is a requirement for a Hamiltonian system.

3.1.4 Force softening and resolution

If two simulation particles get close to each other, they may actually experience two-body relaxation, which is however only an artefact of the limited mass resolution. This behaviour can be avoided by introducing a *softening parameter* ϵ that modifies the small-scale gravitational force law, e. g.

$$\boldsymbol{F}(\boldsymbol{r}) = -G \frac{m_i m_j}{(r^2 + \epsilon^2)^{3/2}} \boldsymbol{r}. \tag{3.4}$$

This so-called Plummer softening law is both intuitive and easy to implement, but it is not optimal (Dehnen, 2001). Nowadays a cubic spline is generally used; nevertheless the value of the softening parameter ϵ is still often quoted as its Plummer-equivalent. The force softening prevents unphysical “collisions” but also sets a minimal spatial scale below which the forces are no longer accurate. Power et al. (2003) studied this for the GADGET-2 code and found that the optimal (in terms of accuracy and computing time) softening scale for a galaxy-sized halo is given by

$$\epsilon \approx 4 \frac{r_{200}}{\sqrt{N_{200}}}, \tag{3.5}$$

where N_{200} is the number of particles within the halo radius r_{200} (more thoroughly defined in section 3.2). For current high-resolution cosmological simulations, this softening scale is on the order of a few hundred parsecs, but for single objects can be as small as a few parsecs. In case of GADGET-2 the softening scale also enters into the adaptive time stepping criterion for each particle i

$$\Delta t_i \propto \sqrt{\epsilon |\boldsymbol{a}_i|^{-1}}, \tag{3.6}$$

where $|\boldsymbol{a}_i|$ is the acceleration experienced by particle i, so a small value for ϵ will negatively impact the runtime of the simulation.

The force resolution for small separations also depends on the number of particles used to trace the density field. According to Power et al. (2003), numerical convergence can be expected for radii where

$$\frac{\sqrt{200}}{8} \frac{N(r)}{\ln N(r)} \left[\frac{\bar{\rho}(r)}{\rho_{\text{crit}}} \right]^{-1/2} \geq 1, \tag{3.7}$$

where $N(r)$ and $\bar{\rho}(r)$ is the number of particles and the mean density inside the radius r, respectively. In summary, the particle number affects both the minimum mass scale (Equation (3.1)), as well as the minimum separation that can be accurately probed, and this needs to be considered when deciding on the specifics of a simulation run.

3.1.5 Initial conditions

In most cases, the initial density field of a cosmological simulation is a realisation of a Gaussian random field with the desired clustering dictated by the late-time power spectrum, which can be calculated numerically by Boltzmann solvers like CAMB (Code for Anisotropies in the Microwave Background; Lewis et al., 2000). Starting from Gaussian white noise, the density field is scaled by the primordial power spectrum and the transfer function (defined in section 2.1.2) at the starting redshift. The starting redshift for a simulation is usually chosen as $z \sim 30 - 100$, high enough so that the fluctuations are still in the linear regime but at the same time already well inside the matter dominated era. The appropriate value depends on the exact cosmology and the box size, which determines the r. m. s. (root mean square) value of the density fluctuations. Starting at significantly higher redshift is generally not advisable because it will waste valuable CPU hours on purely linear evolution, as well as enhance noise from numerical inaccuracies like the imprint of the initial grid (see below) and round-off errors.

As already discussed in section 2.3, only the simplest inflationary model predicts that the initial fluctuations are practically Gaussian. If the underlying model is not that of standard slow-roll Inflation, the density field has to be modified. For the case of local primordial non-Gaussianity (PNG) this is quite straightforward: According to Equation (2.39), local PNG can be expressed by adding higher-order terms to the initial potential (related to the density contrast via the Poisson equation). Since we are considering *primordial* non-Gaussianity, it is important to modify the potential *before* applying the transfer function to the density field, because $T(k)$ encodes the evolution of structures after horizon crossing. For other PNG templates, setting up initial conditions is not as simple, and many of them require numerical integrations of the desired bispectrum, which is computationally expensive (Wagner et al., 2010; Regan et al., 2012). But once the initial conditions have been suitably modified, the code behaves exactly like in the Gaussian case, as the interaction process (gravity) is unchanged.

Sampling of the density field by the N-body particles often starts with these particles distributed on a regular grid.[4] The particles are then moved slightly away from the grid nodes to create density perturbations. The simplest way to calculate the initial displacements and velocities is to apply the Zel'dovich Approximation (ZA; Zel'dovich, 1970) in Lagrangian coordinates $\boldsymbol{q}$ which are related to the Eulerian position $\boldsymbol{x}$ by a displacement field $\boldsymbol{\Psi}$, so $\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\Psi}(\boldsymbol{q}, t)$. The ZA then leads to

$$D(t)\,\delta(\boldsymbol{q}) = -\nabla\boldsymbol{\Psi}^{(1)}(\boldsymbol{q}, t), \tag{3.8}$$

$$\boldsymbol{v}(t) = a(t)\frac{\mathrm{d}D}{\mathrm{d}t}\boldsymbol{\Psi}^{(1)}(\boldsymbol{q}, t), \tag{3.9}$$

which relates the linear displacement field $\boldsymbol{\Psi}^{(1)}$ to the density contrast in the initial conditions, and $D(t)$ is the linear growth factor defined in Equation (2.10). The displacement field is then used to shift the particles away from their position on the nodes of the regular grid, and assign their initial velocities. One problem with

---

[4] Another option is to start from a *glass*, which is the stable end-state of a short simulation with repulsive forces.

this approach is that the ZA introduces transient effects that can leave spurious signatures in the evolved density field at low redshift (corresponding to per cent effects in the power spectrum). This can be avoided by starting at a redshift slightly above that required by the density r. m. s., or by using higher-order techniques like second-order Lagrangian perturbation theory (2LPT; Crocce et al. 2006).

3.1.6 Zoom-in simulations

Zoom-in simulations are useful when one wants to study a smaller region within a cosmological volume, for example an individual halo in a ΛCDM background (e. g. Katz & White, 1993). This technique is used for the Milky-Way-like *Aquarius* simulations (Springel et al., 2008) discussed in Chapter 6, where the particles cluster strongly in a small spatial region, but the influence of the large-scale environment is not negligible. In the first step, a large cosmological volume is evolved until $z = 0$ (or less, depending on what kind of objects are to be studied). Within this parent simulation, the object of interest is selected, and its corresponding Lagrangian region in the initial conditions is identified. Then that region is resampled at higher particle resolution, while the surrounding large-scale structure is sampled at low resolution to ensure the correct large-scale density modulation. In practice, this means defining two sets of particles with different mass that are evolved together. The high-res region has to be chosen sufficiently large (a few times the size of the object of interest), so that there is no undesired contamination with low-res particles in the volume of interest.

3.2 Dark matter halos

Once the density contrast is high enough, the gravitationally bound dark matter particles can decouple from the expansion and collapse to form structures, called dark matter halos, which eventually host galaxies (see also section 2.2). In order to test the analytical predictions for collapsed objects (e. g. the Press-Schechter mass function, Equation (2.17)), we need to be able to identify these halos in our simulation.

3.2.1 Halo finding

Historically, there have been two major classes of halo finding algorithms: Spherical Overdensity (SO) and Friends-of-Friends (FoF), which we will briefly describe below. Recent years have seen increased activity in developing new algorithms, often using the full phase-space information of the particle distribution. For a recent, detailed comparison of 17 different halo finding algorithms, see Knebe et al. (2011, 2013).

Spherical Overdensity (SO) As a first step, the locations of local overdensities (or *peaks*) in the matter distribution are identified. Then a sphere is drawn around each peak with increasing radius. Once the sphere encloses a certain overdensity Δ with respect to the background, the region inside the sphere is called a dark matter halo. Several conventions exist in the literature and care has to be taken when comparing results. For example, the overdensity Δ can be defined relative to the mean density, or the critical density of the background cosmology. The boundary of the SO halo reflects the choice of Δ, where common values are $\Delta = 200$ or $\Delta = 500$, and the corresponding halo radius is then called r_{200} or r_{500}, respectively.

A somewhat misleading, but also very commonly used name for the halo radius defined in this way is the *virial* radius. However, the overdensity threshold has nothing to do with whether the halo is in virial equilibrium or not. For the remainder of this thesis, we will always refer to the halo boundary of SO halos as r_{200}, where $\Delta = 200$ is taken with respect to the *critical* density of the Universe at any given redshift. The mass of the halo, M_{200}, is then calculated by integrating the mass distribution out to r_{200}, assuming spherical symmetry.

Friends-of-Friends (FoF) The Friends-of-Friends algorithm (Davis et al., 1985) connects particles within a certain *linking length*, typically chosen to be 0.2 times the mean particle separation. This value was originally chosen to match the theoretical predictions in the spherical collapse model (Efstathiou et al., 1988). Each FoF halo consists of all particles that have been uniquely linked together. The shape of a FoF halo is thus arbitrary, in contrast to a SO halo. Its center and bulk velocity can either be defined as the one of the particle with the largest binding energy – the most-bound particle – or as the center of mass computed from all linked particles, and these two centres are often identical unless the FoF algorithm created a spurious link between two large structures. The mass of a FoF halo is given by the sum of all its associated particles. An example of a commonly used FoF halo finder is the SUBFIND code (Springel et al., 2001), discussed in section 3.2.3. If required, the “radius” of a FoF halo can be estimated by fitting a density profile (see below) to the halo and identifying the radius where e. g. M_{200} corresponds to the FoF mass.

Based on the differences in the definitions of FoF and SO halos, it is not surprising that the resulting statistics (mass function, clustering) can be discrepant by up to 10% (Jenkins et al., 2001; Knebe et al., 2011, 2013).

3.2.2 NFW density profile

One would expect that given the chaotic nature of structure formation, halos came in all shapes and sizes. The latter at least is true because simulations show a hierarchical scenario where almost all of the mass is enclosed in halos of smaller and smaller size (expected to continue until the free-streaming scale of WIMP dark matter, $M \sim M_{\rm Earth}$). However, on average, the functional form of halo density profiles in dark matter simulations was found to be *universal*. In the so-called Navarro-Frenk-White profile (NFW; Navarro et al., 1997), the density at radius r

is given by

$$\rho(r) = \frac{4\rho_\mathrm{s}}{(r/r_\mathrm{s})(1+r/r_\mathrm{s})^2}, \tag{3.10}$$

which depends only on two parameters: the scale radius r_s where the logarithmic slope of the profile is -2, and $\rho_\mathrm{s} \equiv \rho(r_\mathrm{s})$. On large scales, the logarithmic slope approaches -3 and in the central regions of the halo it approaches -1. It is still unclear why the chaotic processes that are involved in the collapse of halos generate such a universal density profile (e. g. Frenk & White, 2012; Ludlow et al., 2013, and references therein).

Note that the mass of an NFW halo, given by the integral over Equation (3.10), diverges unless a cut-off radius is fixed. If we chose the cut-off radius as r_{200}, the enclosed mass can be calculated analytically and is given by

$$M_{200} = 4\pi\rho_\mathrm{s} r_\mathrm{s}^3 \left[\ln(1+c) - \frac{c}{1+c}\right] \tag{3.11}$$

where we have defined the *concentration* parameter $c = r_{200}/r_\mathrm{s}$. A large value for c therefore implies that a halo is more centrally concentrated. Note that c is in general dependent on both mass and redshift (e. g. Bullock et al., 2001).

While the NFW profile (and alternatives like the Einasto profile, e. g. Gao et al. 2008) is a good fit to numerical simulations over several orders of magnitude in mass, observations suggest that the actual halo mass profile is shallower in the central regions (de Blok, 2010), but the discrepancy can be alleviated by including baryonic feedback (e. g. Oh et al., 2011), see also the discussion in section 6.1.

3.2.3 Substructure

As expected from hierarchical structure formation, simulated halos also contain substructure, i. e. self-bound halos within halos. In the real world, these substructure are supposed to host e. g. galaxies within clusters of galaxies, or dwarf galaxies in the local group surrounding the Milky Way. The radial density profile of subhalos can be described by truncated NFW profiles, because material is stripped from them by interactions with the main halo (Diemand et al., 2007; Springel et al., 2008). The maximum level of nested structures within (sub-)structures is ultimately limited by the DM free streaming scales ($\sim M_\mathrm{Earth}$ for standard CDM particles), and more practically by the mass resolution and the minimum number of particles used to identify a halo (at least 20 for simple statistics like the mass function, more if the internal properties like the density profile are to be studied).

There are also different ways to define subhalos in simulations. Here, we briefly describe the substructure finder used for our work, the SUBFIND code (Springel et al., 2001): After the first selection of FoF halos is complete, each particle within such a top level structure is considered to be a possible center for a substructure. For that, the local density is calculated (as usual on a grid), and isodensity contours are drawn around local overdensities by continually lowering the density threshold. If contours drawn around different density peaks come into contact, this corresponds to a saddle point in the density profile, and determines the region that each substructure encompasses. Note that this process up to now allows for one particle to be part of several subhalos within subhalos. Thus, for each substructure, the potential and kinetic energy of its particles is calculated, and all particles with positive total energy are removed. In the final step, a particle is uniquely assigned to the lowest level of substructure that it is bound to. This whole procedure leads to a hierarchical system of structures within larger structures, and an additional subset of particles that are not bound to any halo.

After repeating this step for every simulation output, it is possible to generate merger trees for both FoF halos and substructures. A merger tree is simply a linked list of progenitors and descendants for each halo at a given redshift. Since halos can and will merge during the course of the simulation, each will have several progenitors in many cases. The major branches of the merger tree are then defined by following the most massive progenitor, and this can be used to study the mass accretion history of a given structure. In Chapter 6, we employ merger trees and another similar technique to track the time evolution of substructures in several high-resolution simulations of Milky-Way-like galaxies.

4 Standard Perturbation Theory and the Local Eulerian halo bias model

This chapter represents a modified version of Roth & Porciani (2011) with small changes in notation and additional references to literature and other chapters of this thesis.

4.1 Introduction

Redshift surveys have revealed the existence of large-scale structures in the Universe: the galaxy distribution is organized in a complex network of filaments surrounding underdense regions and crossing at density peaks which host galaxy clusters. These structures are believed to form through gravitational instability starting from (practically) Gaussian fluctuations characterized by a nearly scale-invariant power spectrum. Mathematically, we can follow the growth of dark-matter density (and velocity) perturbations with respect to a smooth background in terms of a set of Eulerian fluid equations coupled with the Poisson equation. As long as the density and the velocity deviate only slightly from their unperturbed values, the evolutionary equations can be linearised and solved analytically. In this case, each fluid property can be written as the superposition of a term that grows with time and a second one that decays.

At later times, however, non-linear terms in the fluid equations become important and it is no longer possible to derive an exact solution for realistic initial conditions. A widespread technique to compute approximate solutions is to use a perturbative approach and write the full solution as a series expansion in powers of the linear perturbation amplitude. This goes under the name of Standard Perturbation Theory (SPT; Juszkiewicz 1981, Vishniac 1983, Goroff et al. 1986, Makino et al. 1992, Bernardeau et al. 2002). In linear theory, each Fourier mode of the fluid properties evolves independently of the others. Non-linearities in the dynamics correspond to couplings between modes of different wavelengths. When the statistical properties of the initial fluctuations are known, a diagrammatic technique analogous to the Feynman diagrams can be developed to compute ensemble averaged statistics. For a given statistic (e. g. the power spectrum), the lowest-order, non-vanishing terms in the perturbative expansion give the leading expression while higher- order contributions provide additional corrections. Basically, this gives an expansion in powers of the variance of the density contrast.

Since it does not account for multi-streaming, SPT should break down in the highly non-linear configurations that lead to the formation of collapsed structures.

In hierarchical scenarios for structure formation this happens at increasingly larger scales with time. An interesting question is whether, and to what degree, this breakdown affects our ability to predict the evolution of structures on the largest scales. It is generally expected that SPT is meaningful in the so called mildly non-linear regime, i. e. for Fourier modes with wavenumber $k <$ a few $\times\, 0.1\, h\,\mathrm{Mpc}^{-1}$.

Comparison with N-body simulations shows that, in our currently favoured ΛCDM model and for statistics like the power spectrum and the bispectrum, SPT is rather accurate at redshifts $z > 1$ (for $k < 0.2\, h\,\mathrm{Mpc}^{-1}$) while it becomes increasingly imprecise as $z \to 0$ due to the fact that the variance of the density contrast approaches unity (e. g. Carlson et al. 2009; Nishimichi et al. 2009). When this happens, all terms in the perturbative series are of the same importance and it does not make sense to truncate the expansion at finite order. A number of techniques have been proposed to resum or truncate the series in a more meaningful way (Crocce & Scoccimarro 2006, Pietroni 2008, Taruya & Hiramatsu 2008).

Another complication arises if we want to model the large-scale structure seen in galaxy redshift surveys, namely we have to account for the fact that galaxies (and dark-matter halos in general) are biased tracers of the underlying mass distribution. This is clearly seen observationally as different galaxy types show different clustering amplitudes. Consistently, N-body simulations show that the distribution of dark-matter halos depends on their mass and also on other characteristics. The most common way to account for galaxy biasing is to assume that the relation between the fluctuation in galaxy counts within a characteristic volume centred at a given location can be written as a power series of the corresponding volume-averaged mass-density contrast (Fry & Gaztanaga, 1993). This Eulerian local bias (ELB) scheme seems to give an accurate description of the large-scale clustering of dark-matter halos in simulations (Manera & Gaztañaga, 2011). However, it is difficult to accurately predict the bias coefficients based on models for the collapse of fluctuations (Manera et al., 2010).

Heavens et al. (1998) have shown how the ELB model can be combined with SPT to compute the power spectrum of biased tracers of the large-scale structure up to next-to-leading order. Similarly, the dependence of higher order statistics on cosmological parameters and bias coefficients can be evaluated following the same approach (e. g. Gaztanaga & Frieman 1994; Matarrese et al. 1997; Buchalter & Kamionkowski 1999; Sefusatti & Scoccimarro 2005).

This method has to face a technical difficulty: the ELB scheme only makes sense when density fields are averaged over a finite volume while SPT applies to unsmoothed fluctuations. Since the bias parameters depend on the actual scale used for this spatial averaging procedure the model does not make unique predictions. Also, the resulting spectra and multi-spectra show unwanted features at wavenumbers corresponding to the smoothing scale. Renormalization of the bias parameters has been proposed to alleviate the problems (Heavens et al. 1998, McDonald 2006).

Irrespective of these difficulties, the ELB+SPT model has been used to extract the bias parameter that best fit observations (e. g. Verde et al. 2002). Moreover, it is often used to discuss how biasing and non-linearities modify baryonic acoustic oscillations in the galaxy power spectrum (Smith et al. 2007; Jeong & Komatsu 2006, 2007, 2009a) or to quantify scale-dependent biasing in the presence of non-Gaussian initial conditions (Taruya et al. 2008; Jeong & Komatsu 2009b; Sefusatti 2009; Giannantonio & Porciani 2010; Baldauf et al. 2011). Many forecasts about the constraining power of cosmological parameters from future observational campaigns are based on these calculations.

Even though each ingredient of the model has been separately tested against numerical simulations, it still is unclear what level of accuracy can be achieved by combining SPT with the ELB scheme to describe the distribution of biased tracers of the cosmic mass distribution. Future galaxy surveys aiming at determining the origin of cosmic acceleration will require models of the galaxy power spectrum with per cent accuracy. Can the ELB+SPT computational scheme satisfy such a requirement?

We present a direct comparison of model predictions against the outcome of state-of-the-art high-resolution N-body simulations. We follow a novel approach where we evaluate the SPT expansion of the mass density and velocity fields (up to third order) starting from the same realisation of the linear density field that has been used to generate the initial conditions of the simulations[1]. This allows us to make a point-by-point comparison between the non-linear mass and halo overdensities while past studies have only focussed on two- or three-point statistics (of either SPT or ELB). Our analysis sheds new light on the interpretation of the bias parameters in ELB and on the effect of the smoothing procedure intrinsic to the ELB scheme.

This chapter is organized as follows: In section 4.2 we present the principles of SPT and describe how we calculate the matter density field up to third order. The numerical simulations used are described briefly in section 4.3. In section 4.4 we compare the SPT matter density field to the simulations on a point-by-point basis and using one- and two-point statistics. Section 4.5 provides an introduction into galaxy bias, and then describes the estimation of the bias parameters and their dependence on the halo mass and smoothing scales. Using the bias parameters estimated from the simulation allows us to generate a biased dark matter density field. The resulting halo density field is compared to the simulations, and the accuracy of the halo-halo power and halo-mass cross spectra is investigated in detail.

4.2 Standard Perturbation Theory

Linear structure formation is only valid if the density contrast is small, which applies to large scales and/or early times. SPT presents an analytical description to extend linear theory to smaller scales, by finding perturbative solutions to the coupled non-linearised fluid equations (2.3) - (2.5). For convenience, we switch

---

[1] We do not consider resummed theories for which the inclusion of biasing has only recently been discussed (Elia et al., 2011).

to conformal time τ for the following calculations. The basic idea of SPT is that the density contrast in Fourier space $\tilde{\delta}(\boldsymbol{k},\tau)$ can be written as a sum of separable functions of τ and $\boldsymbol{k}$

$$\tilde{\delta}(\boldsymbol{k},\tau) = \sum_{n=1}^{\infty} D_n(\tau)\tilde{\delta}_n(\boldsymbol{k}), \tag{4.1}$$

$$\tilde{\theta}(\boldsymbol{k},\tau) = -\mathcal{H}(\tau)\sum_{n=1}^{\infty} D_n(\tau)\tilde{\theta}_n(\boldsymbol{k}), \tag{4.2}$$

where $D_n(\tau) \approx [D(\tau)]^n$ with $D(\tau)$ being the linear growth factor for ΛCDM models (Bernardeau 1994a; Scoccimarro et al. 1998), and $D = 1$ today. The linearised Continuity equation (2.6) then immediately implies $\tilde{\delta}_1 = \tilde{\theta}_1$. In general, $\tilde{\delta}_n(\boldsymbol{k})$ is of n-th order in the linear density contrast field $\tilde{\delta}_1(\boldsymbol{k})$:

$$\begin{aligned}\tilde{\delta}_n(\boldsymbol{k}) = \int \frac{\mathrm{d}^3 q_1}{(2\pi)^3} \cdots \int \frac{\mathrm{d}^3 q_n}{(2\pi)^3} \Big[\tilde{\delta}_1(\boldsymbol{q}_1),\ldots,\tilde{\delta}_1(\boldsymbol{q}_n) \\ \times F_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n)\,\delta_{\mathrm{D}}(\boldsymbol{k}-\boldsymbol{q_1}-\cdots-\boldsymbol{q_n})\Big].\end{aligned} \tag{4.3}$$

The kernels $F_n(\boldsymbol{q}_1\ldots\boldsymbol{q}_n)$ describe the mode coupling in Fourier space due to the dynamical non-linearities. They can be derived easily for an EdS universe (see Appendix A). The numerical corrections for a flat ΛCDM universe are of the percent order (Bernardeau et al., 2002) and we will therefore neglect them.

The mode-coupling kernels also follow a hierarchy, which can be written as a recursion relation for $n \geq 2$ (Goroff et al., 1986):

$$\begin{aligned}F_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m)}{(2n+3)(n-1)}\,&[(2n+1)\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2) \\ &\times F_{n-m}(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n) + 2\beta(\boldsymbol{k}_1,\boldsymbol{k}_2) \\ &\times G_{n-m}(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n)\Big], \\ G_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m)}{(2n+3)(n-1)}\,&[3\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2) \\ &\times F_{n-m}(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n) + 2n\beta(\boldsymbol{k}_1,\boldsymbol{k}_2) \\ &\times G_{n-m}(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n)\Big],\end{aligned} \tag{4.4}$$

where $\boldsymbol{k}_1 \equiv \boldsymbol{q}_1 + \cdots + \boldsymbol{q}_m$, $\boldsymbol{k}_2 \equiv \boldsymbol{q}_{m+1} + \cdots + \boldsymbol{q}_n$ and $F_1 = G_1 \equiv 1$. The kernels $\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)$ and $\beta(\boldsymbol{k}_1,\boldsymbol{k}_2)$ are given by

$$\begin{aligned}\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2) &\equiv \frac{(\boldsymbol{k}_1+\boldsymbol{k}_2)\cdot\boldsymbol{k}_1}{k_1^2}, \\ \beta(\boldsymbol{k}_1,\boldsymbol{k}_2) &\equiv \frac{(\boldsymbol{k}_1+\boldsymbol{k}_2)^2(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)}{2k_1^2k_2^2}.\end{aligned} \tag{4.5}$$

Statistically, the SPT density field can be described by the power spectrum

$$(2\pi)^3\,P_{mn}(k)\,\delta_{\mathrm{D}}(\boldsymbol{k}+\boldsymbol{k'}) = \langle\tilde{\delta}_m(\boldsymbol{k})\;\tilde{\delta}_n(\boldsymbol{k'})\rangle, \tag{4.6}$$

where $\langle\cdot\rangle$ denotes the ensemble average and $P_{11}(k)$ is called the linear power spectrum. The next-to-leading-order corrections (also called the one-loop corrections) to the power spectrum are of the order of $\tilde{\delta}_1^4$, because the correlations of third order in $\tilde{\delta}_1$ are odd moments and vanish for the case of Gaussian initial conditions. This means that the following terms contribute at the one-loop level:

$$P_{\text{1-loop}}(k) \equiv P_{22}(k) + 2\,P_{13}(k), \tag{4.7}$$

and so we have to calculate the density contrast field up to third order. In order to meaningfully truncate the expansion after $n = 3$, and to preserve the perturbative hierarchy, we need to make sure that $|\delta(\boldsymbol{x})| \ll 1$. This is only true for very early times and/or if the field is "smoothed" (averaged) over a sufficiently large scale. We adopt a Gaussian smoothing kernel, which corresponds to a simple multiplication in Fourier space, and denote the smoothed fields with a superscript s:

$$\tilde{\delta}^{\mathrm{s}}(\boldsymbol{k}) \equiv \mathrm{e}^{\frac{-(|\boldsymbol{k}|R)^2}{2}}\,\tilde{\delta}(\boldsymbol{k}). \tag{4.8}$$

The time evolution of the density contrast is given by

$$\delta^{\mathrm{s}}_{\mathrm{SPT}}(\boldsymbol{k},\tau) = D(\tau)\delta^{\mathrm{s}}_1(\boldsymbol{k}) + D^2(\tau)\delta^{\mathrm{s}}_2(\boldsymbol{k}) + D^3(\tau)\delta^{\mathrm{s}}_3(\boldsymbol{k}), \tag{4.9}$$

and we can obtain the real-space counterparts $\delta^{\mathrm{s}}_{\mathrm{SPT}}(\boldsymbol{x},\tau)$ and $\delta^{\mathrm{s}}_n(\boldsymbol{x})$ using inverse Fourier transformation.

The one-loop power spectra can also be directly calculated from the linear power spectrum (Makino et al. 1992, Jain & Bertschinger 1994) by inserting Equation (4.3) into Equation (4.6) and making extensive use of Wick's theorem:

$$P_{22}(k) = \; 2\int \frac{\mathrm{d}^3q}{(2\pi)^3} F_2^2\left[\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}\right] P_{11}(|\boldsymbol{k}-\boldsymbol{q}|)P_{11}(q), \tag{4.10}$$

$$P_{13}(k) = \; 3\,P_{11}(k)\int \frac{d^3q}{(2\pi)^3} P_{11}(q) F_3^{(s)}\left[\boldsymbol{k},\boldsymbol{q},-\boldsymbol{q}\right], \tag{4.11}$$

where $F_3^{(s)}\left[\boldsymbol{k},\boldsymbol{q},-\boldsymbol{q}\right]$ denotes $F_3\left[\boldsymbol{k},\boldsymbol{q},-\boldsymbol{q}\right]$ symmetrized with respect to its arguments.

In previous studies, the one-loop corrections to the power spectrum were calculated using Equation (4.10) and Equation (4.11). In this chapter, we explicitly calculate the third-order density contrast field *point-by-point* on a 3D grid and compare it to a simulated non-linear density contrast field to test the validity of SPT at low redshifts. We start from an initial density field with N^3 grid points and calculate the corresponding higher order density fields, which can then be evolved to any redshift according to Equation (4.1).

While Equation (4.3) is an elegant analytical description of how the higher order density contrast field depends on the linear one, there are other formulations which are more suited for numerical integrations, because they do not require $(n-1)$ integrations over k-space (the integral over the Dirac delta function can be done

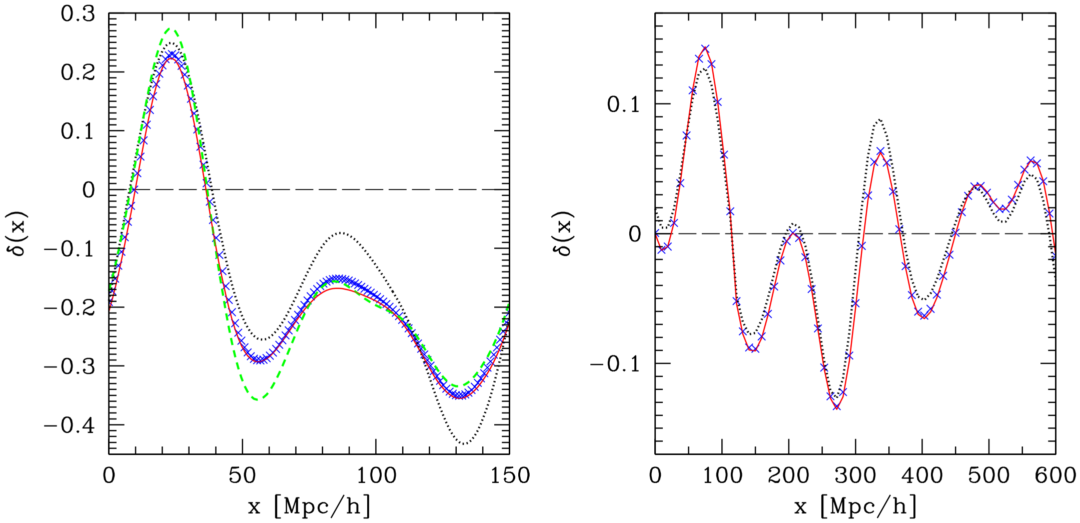

Figure 4.1: Left panel: $\delta^{\mathrm{s}}_{\mathrm{m}}(\boldsymbol{x})$ (blue crosses) compared to the different SPT orders (black dotted: $\delta^{\mathrm{s}}_1(\boldsymbol{x})$, green dashed: $\delta^{\mathrm{s}}_1(\boldsymbol{x}) + \delta^{\mathrm{s}}_2(\boldsymbol{x})$, red solid: $\delta^{\mathrm{s}}_{\mathrm{SPT}}(\boldsymbol{x})$) at redshift 0 for the small box with $R = 12\,h^{-1}\,\mathrm{Mpc}$. Right panel: Same for the large box and $R = 28\,h^{-1}\,\mathrm{Mpc}$. (Second order is not shown here because it is very similar to the third order for this smoothing scale.)

analytically). Makino et al. (1992) showed that Equation (4.3) can be rewritten as a recursion relation (their Eqs. 2.15a and 2.15b), expressed schematically like:

$$\begin{aligned}\tilde{\delta}_n &= K\left[\tilde{\delta}_1, \dots, \tilde{\delta}_{n-1}, \tilde{\theta}_1, \dots, \tilde{\theta}_{n-1}\right],\\ \tilde{\theta}_n &= J\left[\tilde{\delta}_1, \dots, \tilde{\delta}_{n-1}, \tilde{\theta}_1, \dots, \tilde{\theta}_{n-1}\right],\end{aligned} \tag{4.12}$$

where K and J are *single* integrals over k-space with (different) mode-coupling kernels (see Appendix A). The starting point of this integration is given by the linearised Continuity equation: $\tilde{\theta}_1(\boldsymbol{k}) = \tilde{\delta}_1(\boldsymbol{k})$. This method is numerically faster because it takes significantly less time to calculate two single integrals (e. g. $\tilde{\delta}_2$ and $\tilde{\theta}_2$ which are needed for $\tilde{\delta}_3$) than to perform a double-integral for $\tilde{\delta}_3$ using Equation (4.3). This also means that even higher-order corrections could be calculated in the same amount of time and one obtains the velocity divergence field $\tilde{\theta}(\boldsymbol{k})$ up to the $(n-1)$-th order as a side product. We find that this method is also more stable against numerical effects such as the discretization of the mode-coupling kernels.

However, even with this hierarchical method, the integration time still scales with the square of the number of grid points, because if we increase the number of grid points by, say, a factor of 8 (doubling the resolution in each dimension), for each of these points we also have 8 additional grid points that contribute to each integral. This means we are limited by the resolution of our grid, i. e. the minimum separation between grid points. The highest resolution that can be achieved with a common workstation is $(128)^3$ grid points.

4.3 N-body simulation

The simulations that are used in this chapter are described in detail in Pillepich et al. (2010). They were run using the TreePM code GADGET-2 (Springel 2005) in a flat ΛCDM cosmology starting from a Gaussian random field. The cosmological parameters are $\Omega_\mathrm{m} = 0.279$, $\Omega_\mathrm{b} = 0.0462$, $\Omega_\Lambda = 0.721$, $h = 0.701$, $\sigma_8 = 0.817$ and $n_\mathrm{s} = 0.96$. The output contains the 3D positions and velocities of the dark matter particles, which are converted into a density contrast field on a grid using the cloud-in-cell (CIC) algorithm. Outputs for 30 time steps between $z = 10$ and $z = 0$ logarithmically spaced in $(1+z)^{-1}$ are available. The linear density contrast field $\delta_1(\boldsymbol{x})$, which is the basis of our calculations, is obtained from the initial conditions of the simulation. The non-linear matter field from the simulation output at $z = 0$ we denote as $\delta_\mathrm{m}(\boldsymbol{x})$.

The simulations consist of two cubic volumes with different side lengths L, which are used to explore different length scales and halo masses. The small box has $L = 150\,h^{-1}\,\mathrm{Mpc}$ and contains $(1024)^3$ dark matter particles with mass $M_\mathrm{part} = 2.433 \times 10^8\ h^{-1}\,\mathrm{M}_\odot$. The large box has $L = 1200\,h^{-1}\,\mathrm{Mpc}$ and $(1024)^3$ particles with $M_\mathrm{part} = 1.246 \times 10^{11}\,h^{-1}\,\mathrm{M}_\odot$. There are a total of 1,051,230 (1,953,437) halos at redshift 0 identified with the Friends-of-Friends algorithm in the small (large) simulation volume. The linking length is 0.2 times the mean inter-particle separation and each halo contains at least 100 particles. This gives a total mass range of $2.433 \times 10^{10}\,h^{-1}\,\mathrm{M}_\odot \leq M_\mathrm{h} \leq 1.2 \times 10^{15}\,h^{-1}\,\mathrm{M}_\odot$. The initial redshifts are $z = 70$ and $z = 50$, respectively.

To compare the SPT results with the simulations, we need to smooth the density field. The resolution limit of $(128)^3$ grid points for the SPT calculation restricts the scales that can be probed for the large simulation box, because each grid cell has a minimal size of $1200/128\,h^{-1}\,\mathrm{Mpc} \approx 9.4\ \ h^{-1}\,\mathrm{Mpc}$. In order for the smoothing kernel to extend over several grid cells, we have to choose $R \approx 28\,h^{-1}\,\mathrm{Mpc}$ for the large simulation volume. For the small box, the resolution restriction is not significant, but we find that SPT does not give accurate results if $\delta^\mathrm{s}(\boldsymbol{x})$ is of order unity at redshift 0. For this reason, we adopt $R = 12\,h^{-1}\,\mathrm{Mpc}$ which corresponds to an root mean square (r. m. s.) density contrast of 0.6.

We will use the output of the simulations for both dark matter and dark matter halos to compare with our SPT calculation and investigate different length scales, halo masses and redshifts. The simulation results are of course not limited by the SPT grid resolution, but if we want to make point-by-point comparisons, we have to use the same grid and smoothing scale to the simulated mass and halo distributions.

4.4 Dark matter

In the following sections, all density fields are considered to be at redshift 0, unless specifically stated otherwise.

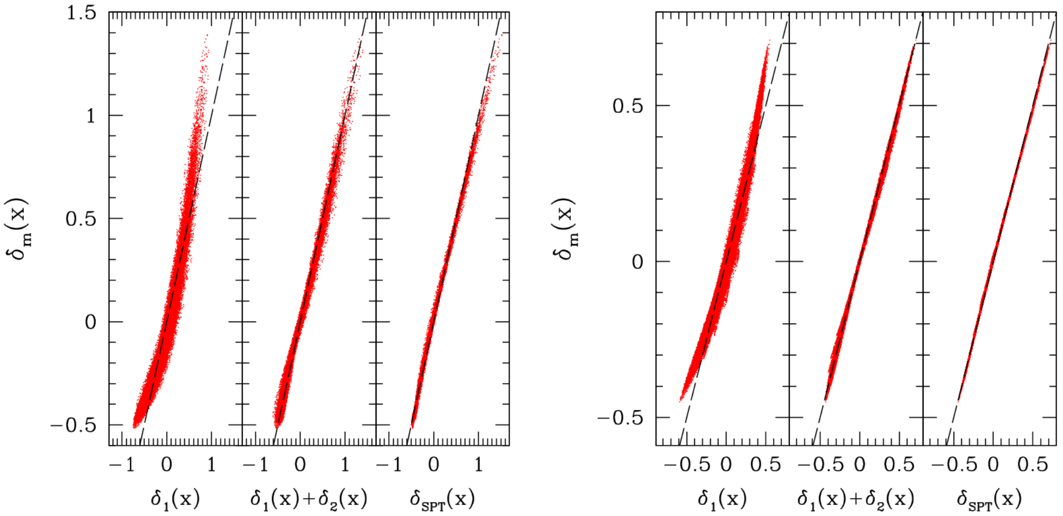

Figure 4.2: Left: Scatter plot of $\delta^{\rm s}_{\rm m}(\boldsymbol{x})$ vs. linear order SPT (left panel), second-order SPT (middle panel), third-order SPT (right panel). Black dashed lines show $\delta^{\rm s}_{\rm m}/\delta^{\rm s}_{\rm SPT} = 1$. The smoothing scale is $R = 12\,h^{-1}\,$Mpc. Note that we only show one out of eight points to improve readability. Right: Same for the large box with $R = 28\,h^{-1}\,$Mpc.

4.4.1 SPT density contrast field

We want to approximate the non-linear matter density contrast field $\delta^{\rm s}_{\rm m}(\boldsymbol{x})$ with the third-order SPT field $\delta^{\rm s}_{\rm SPT}(\boldsymbol{x})$. We also have to investigate if the third-order expansion is actually more accurate than the lower orders. Figure 4.1 shows the different approximations to $\delta^{\rm s}_{\rm m}(\boldsymbol{x})$ along a line parallel to one of the coordinate axis in the simulation over the box-length L (half a box-length for $L = 1200\,h^{-1}\,$Mpc to make the lines more distinguishable). For both simulation volumes, the third-order density contrast gives the best approximation to the simulated matter density contrast $\delta^{\rm s}_{\rm m}(\boldsymbol{x})$. The agreement is already pretty good for the small box but excellent for the large box, owing to the larger smoothing scale in the latter case. While the linear density contrast $\delta^{\rm s}_{1}(\boldsymbol{x})$ traces the overall structure of the non-linear field, the higher order expansion is more accurate.

Figure 4.2 shows a scatter plot of the non-linear field in the simulation and the different SPT orders for both volumes. Starting from the left panel, we see again that the linear density contrast is not a very good approximation to the non-linear density contrast. Although this result is not really surprising, the plot once again shows that the linear approximation is especially wrong when $|\delta^{\rm s}(\boldsymbol{x})| \not\ll 1$. The middle panel shows the density contrast up to second order and the right panel shows the density contrast up to third order. Note that the third-order over corrects the linear density around $\delta^{\rm s} \approx 1$. From left to right the scatter around the black line given by $\delta^{\rm s}_{\rm m}/\delta^{\rm s}_{\rm SPT} = 1$ is reduced. In all three panels the difference to $\delta^{\rm s}_{\rm m}(\boldsymbol{x})$ is largest when $|\delta^{\rm s}(\boldsymbol{x})| \not\ll 1$. Again, the difference in scatter between the large and small simulation volumes is due to the different smoothing scales used (12 and $28\,h^{-1}\,$Mpc).

Figure 4.3 shows the probability distribution function (PDF) for the linear, SPT and non-linear matter density contrast for the small box, with a smoothing scale

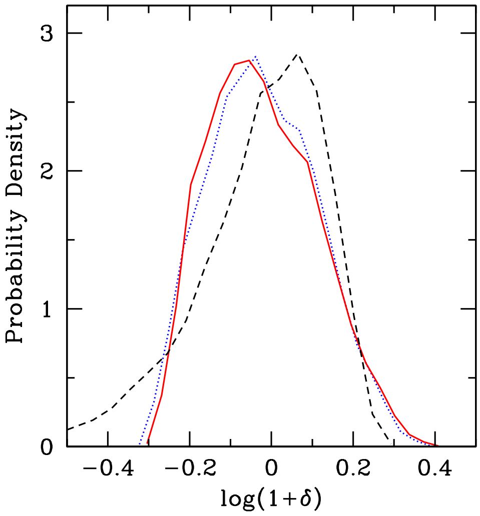

Figure 4.3: Probability distribution function (PDF) for $\delta^{\rm s}_{\rm SPT}(\boldsymbol{x})$ (red solid), $\delta^{\rm s}_{\rm m}(\boldsymbol{x})$ (blue dotted) and $\delta^{\rm s}_{\rm l}(\boldsymbol{x})$ (black dashed) at redshift 0 (small box, $R = 12\,h^{-1}\,$Mpc).

$R = 12\,h^{-1}\,$Mpc. This also shows that SPT provides a good approximation for the non-linear density contrast even at redshift 0.

Redshift Dependence

So far we have only compared our SPT calculation at redshift 0. Now we want to extend the comparison to higher redshifts. Figure 4.4 shows the conditional mean of the non-linear density contrast given its linear counterpart at redshift 0 (left panel) and redshift 1.6 (right panel). The lines show the conditional mean density corresponding to: the linear evolution (black short-dashed), the SPT density (red solid), the non-linear evolution from the simulation (blue dotted) and an analytic expectation from the spherical collapse model (Bernardeau 1994b, Mo & White 1996, green long-dashed), all smoothed with $R = 12\,h^{-1}\,$Mpc. The hatched areas show the 1-σ scatter contours for SPT and simulation. SPT and non-linear matter density agree well if $\delta^{\rm s}_{\rm lin}$ is not too large, as expected, but the scatter around the mean SPT density is generally larger than around $\delta^{\rm s}_{\rm m}$. The spherical collapse model (which was derived for the Einstein-de Sitter cosmological model) always overpredicts the non-linear density contrast, and gets close only when $\delta^{\rm s} < 0$. The linear approximation overpredicts the simulation at low densities, and underpredicts it at high densities. The right panel shows the same for redshift 1.61. At this redshift, SPT and non-linear matter density are in almost perfect agreement, but the scatter in SPT is still slightly larger. The spherical collapse model and the linear evolution are now closer to the non-linear evolution, but still show the systematics seen at redshift 0.

The spherical collapse model is commonly used to compute Eulerian bias coefficients for dark matter halos (e. g. Mo & White 1996, Giannantonio & Porciani 2010 and references therein), but it does not describe our data well. We therefore provide a new and more accurate fitting formula for the relation between $\delta^{\rm s}_{\rm m}$ and $\delta^{\rm s}_{\rm l}$. The redshift range we can study is $0 \le z \le 10$ in 30 bins, logarithmically

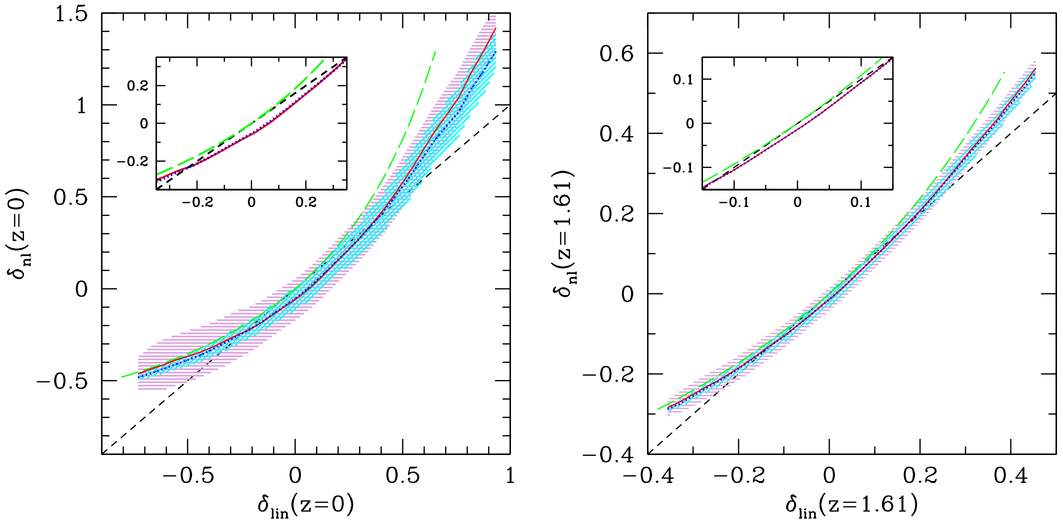

Figure 4.4: Left panel: Mean non-linear matter density contrast from the simulation (blue dotted), SPT (red solid) and an analytic prediction from the spherical collapse model (green long-dashed) as a function of the linear density contrast at redshift 0, for the small box with $R = 12\,h^{-1}\,\mathrm{Mpc}$. The hatched areas show the 1-σ scatter around the mean for $\delta^{\rm s}_{\rm SPT}$ (horizontal hatching) and $\delta^{\rm s}_{\rm m}$ (45°), and the black short-dashed line corresponds to linear evolution. The inset shows the region around $\delta^{\rm s} = 0$, where linear theory actually overpredicts the non-linear evolution. Right panel: Same at $z = 1.61$.

spaced in $(1+z)^{-1}$. For each redshift bin, we find that the parameterisation

$$\delta_{\rm m}(z) = A(z) + B(z) \cdot \delta^{\rm s}_1(z) + C(z) \cdot [\delta^{\rm s}_1(z)]^2 \tag{4.13}$$

gives a very good fit for all values of $\delta^{\rm s}_1$: the ratio of the fit residuals to the true value is less than 5 per cent for redshift 0 and less than 1 per cent for $z > 6$. The functions A, B and C depend on z in the following way:

$$\begin{aligned} A(z) &= 0.006 - \frac{0.054}{1+z}, \\ B(z) &= 1.000 - \frac{0.045}{1+z}, \\ C(z) &= 0.643 - 0.012\,z - \frac{0.122}{1+z}. \end{aligned} \tag{4.14}$$

In the limiting case $z \to \infty$, some of the fit coefficients do not show the right asymptotic behaviour ($A \to 0$, $B \to 1$, $C \to 0$) as one would expect because $\delta^{\rm s}_{\rm m} \to \delta^{\rm s}_1$, so this fit should *not* be used for redshifts outside the fitted range. Both density fields were smoothed with $R = 12\,h^{-1}\,\mathrm{Mpc}$. The choice of R does not significantly influence the fitting parameters, but the agreement is worse when large values of $\delta^{\rm s}_1$ and $\delta^{\rm s}_{\rm m}$ are allowed (i. e. R is small at low redshifts).

Another relation between the linear and non-linear density fields that is sometimes used is the lognormal transformation (e. g. Coles & Jones 1991, Bernardeau & Kofman 1995). This has proven to yield a PDF which agrees well with simulations. However, this does not imply that it can be used as a point-by-point relation between the linear and non-linear density fields, as pointed out by Kayo et al. (2001). We follow their parameterisation, which can be rewritten as

$$\delta_{\rm m}^{\rm s}(z) = e^{\gamma(\delta_1^{\rm s}(z) - \frac{1}{2}\gamma\sigma^2)} - 1, \tag{4.15}$$

with

$$\gamma^2\sigma^2 = \ln(1 + \sigma_{\rm m}^2). \tag{4.16}$$

Here, σ^2 ($\sigma_{\rm m}^2$) is the variance of the smoothed linear (non-linear) density field. We calculate σ^2 and leave γ as the fit parameter. This fit yields values for $\delta_{\rm m}$ which are systematically lower by ≈ 30 per cent at redshift 0 and still 20 per cent at $z > 6$. So indeed the lognormal transformation can not be used to accurately predict the evolution of an initial Gaussian field into a non-linear field. Again, this does not mean that the statistics of a non-linear density field can not be described by a lognormal field, just that one can not expect that, for a point-by-point comparison, the linear field used in the transformation corresponds to the initial conditions of that non-linear field.

4.4.2 Matter power spectra

After calculating the higher order density contrast, we now calculate the matter power spectra up to the 1-loop order. We can do this in two ways: calculating the volume average of the SPT density contrast via Equation (4.6) or integrating a given linear power spectrum $P_{11}(k)$ via Equation (4.10)), where $P_{11}(k)$ is the same power spectrum that was used to set up the initial conditions of the simulation. This also serves as a consistency check for our grid-based calculation. The results for the large box are shown in Figure 4.5. The small box is not shown because it is very similar. The lines show the results from Equation (4.10) and the points show the numerical result, without smoothing, i. e. $R = 0$. The errors of the power spectra are calculated assuming that our fields are Gaussian (Feldman et al., 1994):

$$\sigma_P^2 = \frac{2}{n_{\rm B}} P^2(k), \tag{4.17}$$

where $n_{\rm B}$ is the number of modes in each k-bin[2].

The linear power spectrum from the simulation agrees very well with the input power spectrum until the scales which are affected by the resolution of the grid of the matter field. Once we apply smoothing, this small-scale effect will disappear. The P_{22}-term from the integration and the SPT calculation fits the simulation very well on all applicable scales. Finally, the P_{13}-term from the SPT calculation is slightly lower than the power spectrum integration (a factor of ≈ 1.1). The lines that are shown in the left panel of Figure 4.5 are the integration result with limits corresponding to the scales which are available in our simulation box, but since the box only contains discrete modes and the power spectrum integration assumes a continuous frequency spectrum, choosing the same upper and lower bound for

---

[2] The factor of 2 reflects the fact that we should only consider independent Fourier modes in each bin, and for the case of a real-valued field, the modes $\boldsymbol{k}$ and $-\boldsymbol{k}$ are not independent.

Table 4.1: Halo mass bins: I-III for the small box and IV-VI for the large box.

Mass bin	Mass range [$10^{12}\,h^{-1}\,\mathrm{M}_\odot$]	Number of halos
I	0.025 - 0.035	91,511
II	0.5 - 2	17,030
III	5 - 10	2,693
IV	12.5 - 30	376,292
V	60 - 150	71,581
VI	300 - 1200	7,673

both methods does not necessarily give the same result. A consistency check for both methods is if the power spectra overlap in the k-region covered by both boxes. We find that the grid-based method gives very consistent results in this region, while there is an offset for the power spectrum integration. By slightly modifying the integration limits for $P_{13}(k)$, the discrepancy in the latter can be fixed, and both methods agree with each other.

The right panel of Figure 4.5 shows the full third-order SPT matter power spectrum compared to the simulation. The lines show the power spectra directly calculated from the linear and non-linear density contrast in the simulation, which have been “glued together” to show both volumes at the same time. The points show the result from our grid-based calculation. For the large box, linear theory, SPT and non-linear evolution agree until $k \approx 0.1\,h\,\mathrm{Mpc}^{-1}$, but linear theory seems to be the better approximation for slightly larger k. For the small box, the SPT matter power spectrum describes the non-linear evolution better than the linear power spectrum for scales $0.2\,h\,\mathrm{Mpc}^{-1} < k < 0.3\,h\,\mathrm{Mpc}^{-1}$, while they are equivalent for larger scales. We stress again that these power spectra (or the density contrast used to compute them) are not smoothed at all in both panels of Figure 4.5. This explains the discrepancy of the SPT power spectra on small scales with respect to the good agreement of the density contrasts shown in Figure 4.1. The fact that the points deviate from each other in the overlap region between the large and small box is due to $P_{13}(k)$, as discussed above. There is excess power on these scales for the full SPT power spectrum because $P_{13}(k) < 0$.

4.5 Halos

In this section we study dark matter halos and how they relate to the underlying matter density field. The halos are split up in 3 mass bins for each box because the bias parameters are very sensitive to the mass of the objects (see also Figure 4.11). Table 4.1 shows the selection criteria for the binning. Halo positions are assigned to the grid using the CIC algorithm to obtain a continuous halo density contrast field $\delta_\mathrm{h}(\boldsymbol{x})$. The bins were chosen such that they are separated in mass, but still contain a large enough number of halos to reduce the effect of shot noise.

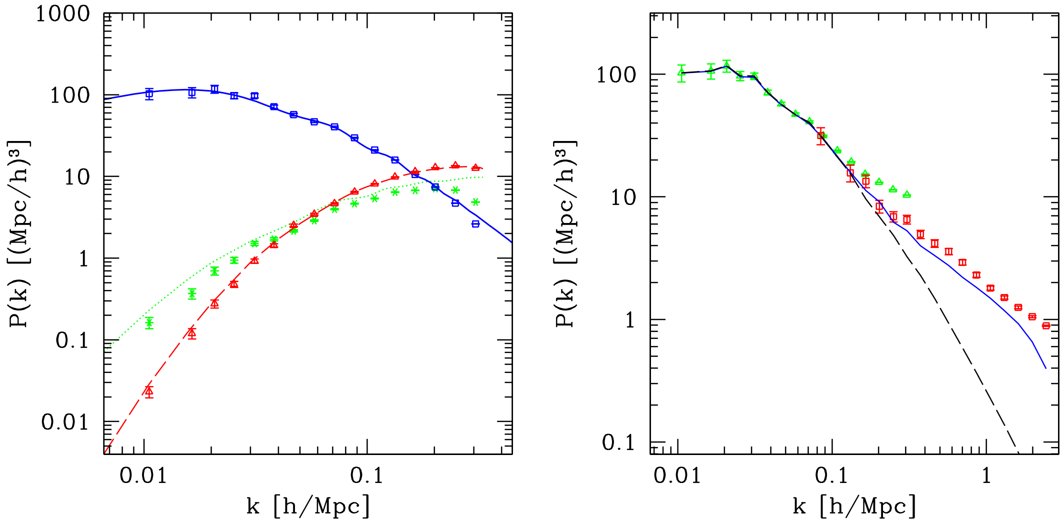

Figure 4.5: Left panel: Contributions to the SPT matter power spectrum for the big box: $P_{11}(k)$ (blue solid line and squares), $P_{22}(k)$ (red dashed line and triangles) and $|2\,P_{13}(k)|$ (green dotted line and stars). The points are the results from our calculations, while the lines are obtained from (integrating) a linear power spectrum according to Equation (4.10). Right panel: The third-order SPT power spectrum (points with error bars) compared to the non-linear power spectrum from the simulation (blue solid line) and the linear power spectrum evolved to redshift 0 (black dashed line) for both boxes. Green triangles and red squares are used to distinguish the large and small box.

4.5.1 Bias estimation

Many different bias models exist in the literature (see section 2.2.1), but they can be broadly classified by the choice of coordinates (Eulerian vs. Lagrangian), and their functional form (local vs. non-local). Here, we use the *local Eulerian* bias model

$$\delta_{\rm h}(\boldsymbol{x}) = F\,[\,\delta_{\rm m}(\boldsymbol{x})\,]\,, \tag{4.18}$$

where the term “local” implies that the halo density $\delta_{\rm h}(\boldsymbol{x})$ depends only on the matter density $\delta_{\rm m}(\boldsymbol{x})$ at the same point in space through a general function F. A choice of F which is widely used is that of a linear relation between $\delta_{\rm h}(\boldsymbol{x})$ and $\delta_{\rm m}(\boldsymbol{x})$ with a constant bias factor b:

$$\delta_{\rm h}(\boldsymbol{x}) = b\,\delta_{\rm m}(\boldsymbol{x}). \tag{4.19}$$

This so-called linear biasing only changes the amplitude but not the shape of the power spectrum, since $P_{\rm h}(\boldsymbol{k}) = b^2\,P_{\rm m}(\boldsymbol{k})$. It has been shown that this is not in agreement with the power spectra derived from numerical simulations (e. g. Smith et al. 2003). Also, for $|b| > 1$, it can lead to $\delta_{\rm h}(\boldsymbol{x}) < -1$, which of course does not make physical sense. However, on *large scales* the linear approximation is valid as will also be shown in section section 4.5.2.

The next step would be to consider a *non-linear* relation between $\delta_{\rm h}(\boldsymbol{x})$ and $\delta_{\rm m}(\boldsymbol{x})$. If we assume that $\delta_{\rm m}(\boldsymbol{x})$ is a smooth field, we can Taylor-expand the right hand side of Equation (4.18) around $\delta_{\rm m}(\boldsymbol{x}) \approx 0$ and get an expression for the higher-order *bias coefficients* b_j

$$\delta_{\rm h}(\boldsymbol{x}) = \sum_{j=0}^{\infty} \frac{b_j}{j!}\, \delta_{\rm m}^j(\boldsymbol{x}), \tag{4.20}$$

where b_1 is the linear bias, and b_0 is determined by the requirement $\langle \delta_{\rm h}(\boldsymbol{x}) \rangle = 0$. To be consistent with our third-order SPT matter field, we first smooth $\delta_{\rm h}(\boldsymbol{x})$ according to Equation (4.8) and truncate Equation (4.20) after b_3:

$$\delta_{\rm h}^{\rm s}(\boldsymbol{x}) = \; b_1 \delta_{\rm m}^{\rm s}(\boldsymbol{x}) + \frac{b_2}{2}\left(\delta_{\rm m}^{\rm s}(\boldsymbol{x})\right)^2 + \frac{b_3}{6}\left(\delta_{\rm m}^{\rm s}(\boldsymbol{x})\right)^3 . \tag{4.21}$$

This means we obtain the b_i by fitting a polynomial to the scatter plot of $\delta_{\rm h}^{\rm s}(\boldsymbol{x})$ and $\delta_{\rm m}^{\rm s}(\boldsymbol{x})$, which is shown for all six mass bins in Figure 4.6.

However, there is a subtlety in the way the local bias model is applied. While the input mass density contrast field is smoothed on a certain scale R, and also the halo density field in the scatter plot is smoothed on the same scale, the *reconstructed* halo density from the fits also contains contributions from smaller scales. The reason for this is the choice of the filter function in Equation (4.8). If we used a top-hat filter in k-space, this effect would vanish, because then all modes with $k > R$ are set to zero (however this filter oscillates in real-space). But with the Gaussian filter, modes with, say, $kR = 2$ are only damped by a factor of ≈ 0.14. When applying the ELB model, taking the square and cube of the smoothed field in real-space corresponds to a convolution in k-space. This *unsmoothing effect* can be seen especially in the halo power spectra (section subsection 4.5.2), and limits the scales where we can trust the calculated halo power spectra even more than the validity of the SPT approximation discussed in section subsection 4.4.2.

Our fitting process allows for three free parameters: b_1, b_2 and b_3 (b_0 is not a free parameter and always very close to 0 so it can be neglected). We also fit the point cloud with fewer bias parameters and use the Akaike information criterion[3] (AIC) to determine the best fit without overfitting the data with too many parameters. We define

$$\begin{aligned}
\delta_{\rm h,f_1}(\boldsymbol{x}) &\equiv b_1 \delta_{\rm m}^{\rm s}(\boldsymbol{x}), \\
\delta_{\rm h,f_2}(\boldsymbol{x}) &\equiv b_1 \delta_{\rm m}^{\rm s}(\boldsymbol{x}) + \frac{b_2}{2}\left(\delta_{\rm m}^{\rm s}(\boldsymbol{x})\right)^2 , \\
\delta_{\rm h,f_3}(\boldsymbol{x}) &\equiv b_1 \delta_{\rm m}^{\rm s}(\boldsymbol{x}) + \frac{b_2}{2}\left(\delta_{\rm m}^{\rm s}(\boldsymbol{x})\right)^2 + \frac{b_3}{6}\left(\delta_{\rm m}^{\rm s}(\boldsymbol{x})\right)^3 ,
\end{aligned} \tag{4.22}$$

which describes the parameterisation used for fitting and the notation to denote the halo density contrast $\delta_{\rm h,f_i}(\boldsymbol{x})$ obtained from the i-th order fit. Note that $\delta_{\rm h,f_i}(\boldsymbol{x})$ is not written with the superscript s due to the unsmoothing effect discussed in the previous paragraph. The errors on the parameters are calculated using the jackknife method with 8 subsamples of the density contrast fields. Scatter plots of $\delta_{\rm h}^{\rm s}(\boldsymbol{x})$ vs. $\delta_{\rm m}^{\rm s}(\boldsymbol{x})$ and the fitted polynomials are shown in Figure 4.6 for all six

---

[3] Basically a χ^2 method which penalizes extra parameters.

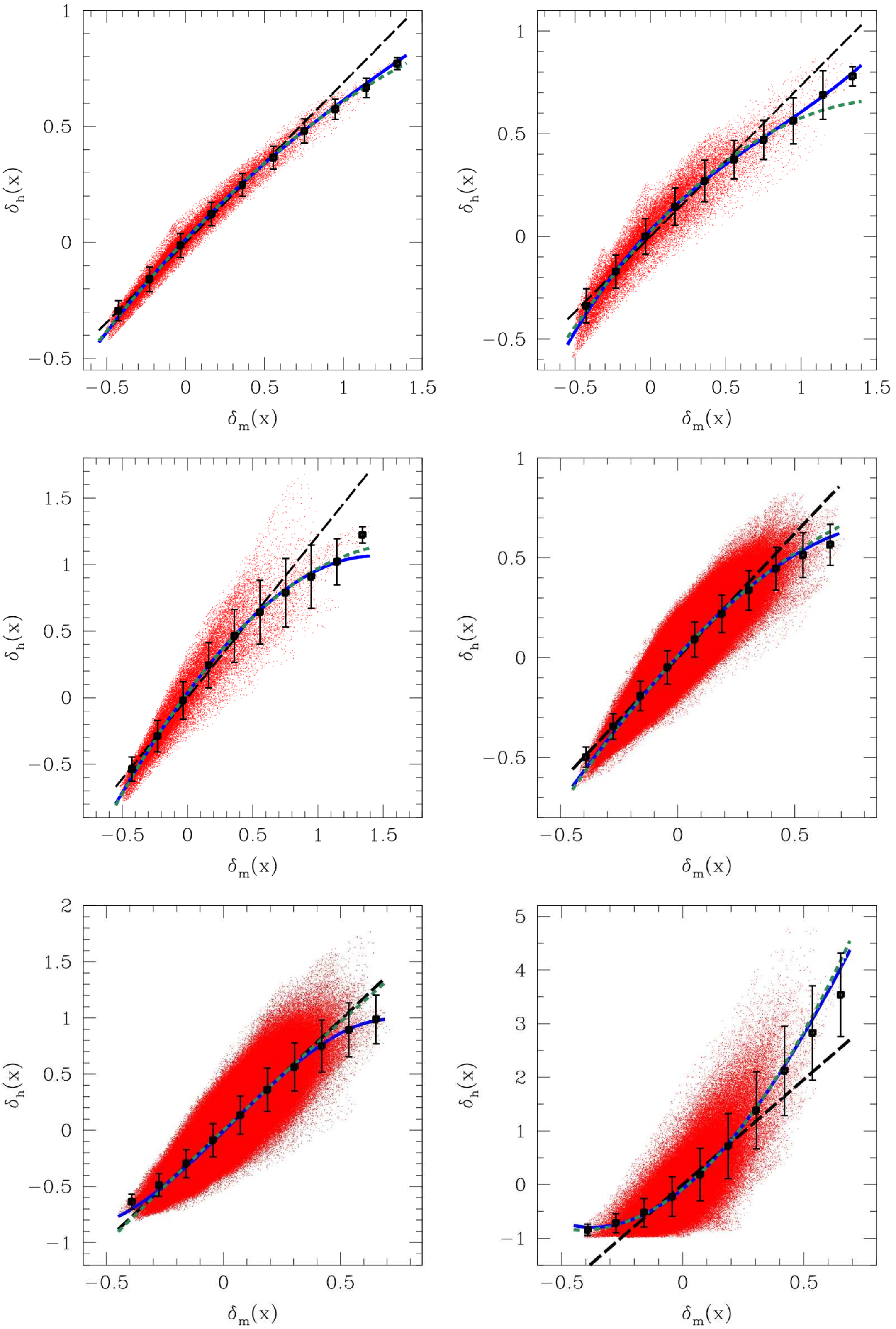

Figure 4.6: Scatter plot of $\delta_{\mathrm{h}}^{\mathrm{s}}(\boldsymbol{x})$ vs. $\delta_{\mathrm{m}}^{\mathrm{s}}(\boldsymbol{x})$. Top row: mass bins I and II ($R = 12\,h^{-1}$ Mpc), middle row: bins III ($R = 12\,h^{-1}$ Mpc) and IV ($R = 28\,h^{-1}$ Mpc), bottom row: bins V and VI ($R = 28\,h^{-1}$ Mpc). The lines show the polynomials of different order fitted to determine the bias parameters b_i (black long-dashed: linear, green short-dashed: second order, solid blue: third order). The black squares with errorbars show the mean of $\delta_{\mathrm{h}}^{\mathrm{s}}$ in bins of $\delta_{\mathrm{m}}^{\mathrm{s}}$ and the 1-σ scatter around it as visual guidance. Note that to increase readability, we only show one of every eight points used in the fitting procedure.

Table 4.2: Bias parameters for both third-order fits for different mass bins and smoothing scales R [h^{-1} Mpc].

	Mass bin	R	$\delta_{\mathrm{h}}^{\mathrm{s}}$ vs. $\delta_{\mathrm{m}}^{\mathrm{s}}$	$\delta_{\mathrm{h}}^{\mathrm{s}}$ vs. $\delta_{\mathrm{SPT}}^{\mathrm{s}}$
b_1	I	8	0.747 ± 0.006	0.666 ± 0.008
b_2	I	8	-0.341 ± 0.022	-0.554 ± 0.021
b_3	I	8	0.171 ± 0.034	1.542 ± 0.048
b_1	I	12	0.719 ± 0.009	0.711 ± 0.010
b_2	I	12	-0.300 ± 0.034	-0.438 ± 0.017
b_3	I	12	0.171 ± 0.125	0.620 ± 0.111
b_1	III	12	1.333 ± 0.035	1.323 ± 0.034
b_2	III	12	-0.716 ± 0.140	-1.086 ± 0.093
b_3	III	12	-0.288 ± 0.612	0.066 ± 0.532
b_1	V	28	2.039 ± 0.023	2.037 ± 0.023
b_2	V	28	0.270 ± 0.124	-0.346 ± 0.136
b_3	V	28	-8.682 ± 1.476	-8.556 ± 1.446

mass bins. The mean and $1-\sigma$ scatter for $\delta_{\mathrm{h}}^{\mathrm{s}}$ in bins of $\delta_{\mathrm{m}}^{\mathrm{s}}$ are indicated by points with errorbars. One can see that the local bias model seems to be a rather good approximation, although there is some scatter in the relation. A possible source of this could be shot noise from sampling discrete halos. Assuming Poissonian shot noise, the amplitude can be estimated by calculating the number of halos $N_{\mathrm{s}} = \rho_{\mathrm{h}}\, V_{\mathrm{s}}$ that are in each smoothing volume $V_{\mathrm{s}} = 6\pi^2 R^3$, and the mean number of halos in that mass bin, $\overline{N}$, from the third column in Table 4.1:

$$\Delta\delta_{\mathrm{h}}^{\mathrm{s}} = \frac{\Delta N_{\mathrm{s}}}{\overline{N}} = \frac{\sqrt{\overline{N}\,(\delta_{\mathrm{h}}^{\mathrm{s}}+1)}}{\overline{N}}. \tag{4.23}$$

where $\Delta N_{\mathrm{s}} = \sqrt{N_{\mathrm{s}}}$ (Poissonian noise) was used in the second step. For mass bin I with $R = 12\, h^{-1}$ Mpc, $|\delta_{\mathrm{h}}^{\mathrm{s}}| \leq 0.8$. The maximum error from shot noise would be $\Delta(\delta_{\mathrm{h}}^{\mathrm{s}} = 0.8) \approx 0.03$. But one can see from Figure 4.6 that the scatter does not increase as $|\delta_{\mathrm{h}}^{\mathrm{s}}|$ increases, and that it is much larger than the maximum shot noise error. This means that there is an *intrinsic scatter* around the local bias model, i. e. the bias is not deterministic (see also Somerville et al. 2001). For the higher mass bins with fewer halos (right panel), the effect of shot noise is more prominent. There have been suggestions in the literature how to deal with stochasticity (e. g. Dekel & Lahav 1999; Cai et al. 2011), but we will not include this in our modelling. The AIC values can be converted into a relative probability w_i for each model (see e. g. the appendix of Porciani & Norberg 2006), with $\sum_i w_i = 1$. Independent of halo masses and smoothing scale R, the third-order bias model always provides the best fit, with $w_3 \approx 1$ while w_1, w_2 are lower by several orders of magnitudes. This means that although the second- and third-order fits in Figure 4.6 differ only in regions with relatively few points, the AIC criterion clearly favours a model with 3 bias parameters.

In the previous sections, we showed that $\delta_{\mathrm{SPT}}^{\mathrm{s}} \approx \delta_{\mathrm{m}}^{\mathrm{s}}$. So instead of fitting to the simulation matter density contrast, we can also obtain the bias parameters by

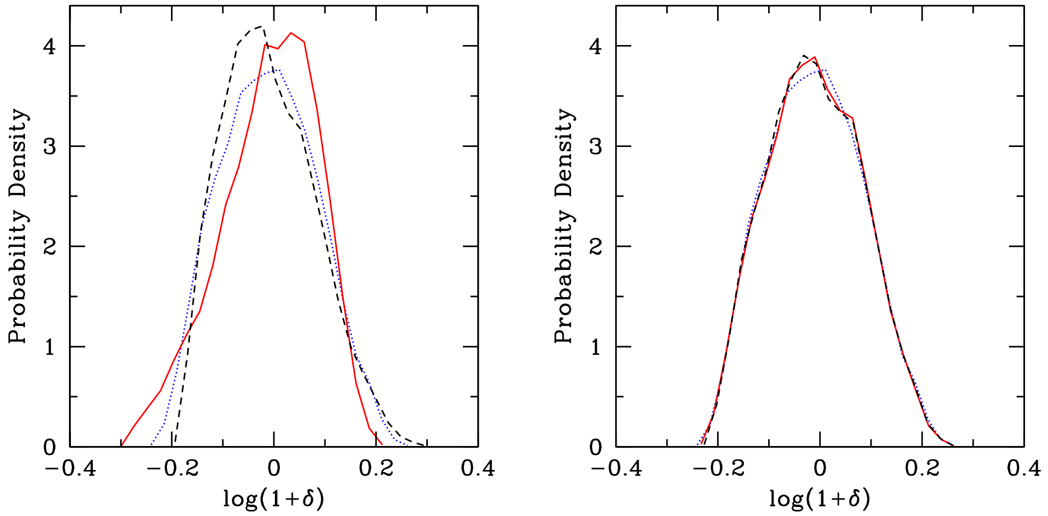

Figure 4.7: Left panel: Probability distribution function for $\delta_{\mathrm{h}}(\boldsymbol{x})$ (dotted blue) compared to $\delta_{\mathrm{h},f_1}(\boldsymbol{x})$ (black dashed) and $\delta^{\mathrm{SPT}}_{\mathrm{h},f_1}(\boldsymbol{x})$ (red solid). Right panel: Probability distribution function for $\delta_{\mathrm{h}}(\boldsymbol{x})$ (dotted blue) compared to $\delta_{\mathrm{h},f_3}(\boldsymbol{x})$ (black dashed) and $\delta^{\mathrm{SPT}}_{\mathrm{h},f_3}(\boldsymbol{x})$ (red solid). Both panels show mass bin I with $R = 12\,h^{-1}\,\mathrm{Mpc}$.

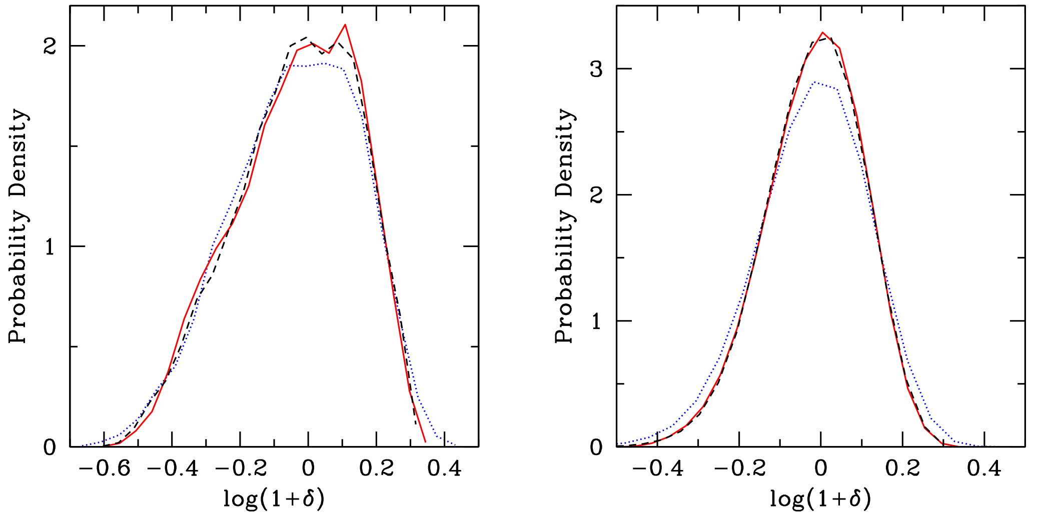

Figure 4.8: Probability distribution function for $\delta_{\mathrm{h}}(\boldsymbol{x})$ (dotted blue) compared to $\delta_{\mathrm{h},f_3}(\boldsymbol{x})$ (black dashed) and $\delta^{\mathrm{SPT}}_{\mathrm{h},f_3}(\boldsymbol{x})$ (red solid). Left panel: Medium mass bin III with $R = 12\,h^{-1}\,\mathrm{Mpc}$. Right panel: High mass bin V with $R = 28\,h^{-1}\,\mathrm{Mpc}$.

fitting $\delta^{\rm s}_{\rm h}(\boldsymbol{x})$ vs. $\delta^{\rm s}_{\rm SPT}(\boldsymbol{x})$. Similar to Equation (4.22), we define

$$\begin{aligned}
\delta^{\rm SPT}_{\rm h,f_1}(\boldsymbol{x}) &\equiv \bar{b}_1 \delta^{\rm s}_1(\boldsymbol{x}), \\
\delta^{\rm SPT}_{\rm h,f_2}(\boldsymbol{x}) &\equiv \bar{b}_1 \left[\delta^{\rm s}_1(\boldsymbol{x}) + \delta^{\rm s}_2(\boldsymbol{x})\right] + \frac{\bar{b}_2}{2} \left(\delta^{\rm s}_1(\boldsymbol{x})\right)^2 , \\
\delta^{\rm SPT}_{\rm h,f_3}(\boldsymbol{x}) &\equiv \bar{b}_1 \left[\delta^{\rm s}_1(\boldsymbol{x}) + \delta^{\rm s}_2(\boldsymbol{x}) + \delta^{\rm s}_3(\boldsymbol{x})\right] + \frac{\bar{b}_2}{2} \Big[\left(\delta^{\rm s}_1(\boldsymbol{x})\right)^2 \\
&\qquad + 2\delta^{\rm s}_1(\boldsymbol{x})\delta^{\rm s}_2(\boldsymbol{x})\Big] + \frac{\bar{b}_3}{6} \left(\delta^{\rm s}_1(\boldsymbol{x})\right)^3 . \qquad (4.24)
\end{aligned}$$

The bar is used to point out that b_i and $\bar{b}_i$ are generally different. This is not particularly surprising because while we showed before that $\delta^{\rm s}_{\rm m} \approx \delta^{\rm s}_{\rm SPT}$, the density contrast we fit to for the $\bar{b}_i$ is truncated at the order of the fit in δ_1. Considering the third-order fit, one therefore expects the largest deviation between the fit parameters to occur between b_3 and $\bar{b}_3$, which is indeed the case except for the high mass bin V (see Table 4.2). This truncation is consistent with previous works, and using the full SPT density contrast in the second- and third-order fits does not improve the agreement between b_i and $\bar{b}_i$. The errors on $\bar{b}_i$ are also jackknife errors from 8 subsamples, and again the third-order fit is always preferred by the AIC. Note that the bias parameters also depend on the smoothing scale R. Both effects can be seen in Table 4.2 for mass bins I, III and V. We do not show the values of the fitted bias parameters for all possible combinations of mass bins and smoothing scales, but we comment on their mass dependence for a specific smoothing scale in section section 4.5.1. The fact that the allowed range for b_3 does not include zero in most cases is also an indication that the third-order bias parameter is required for the fit.

Testing the Local Bias Assumption

Now we can compare the fitted halo density contrast $\delta_{{\rm h,f}_i}(\boldsymbol{x})$ to the true halo density contrast from the simulations $\delta^{\rm s}_{\rm h}(\boldsymbol{x})$. Figure 4.7 shows the probability distribution function of the fitted halo density contrast compared to the simulation. The left panel shows the results using a linear fit, i. e. only one bias parameter b_1, fitting to $\delta^{\rm s}_{\rm m}$ and $\delta^{\rm s}_1$ respectively. Neither fit gives a satisfactory halo density contrast, and the PDFs look quite different. The right panel shows the same but this time using the third-order fit with all three bias parameters. Here, both fits give PDFs which are almost indistinguishable, but the agreement with the simulated halo density contrast is still not perfect. This discrepancy can be seen for all mass bins: Figure 4.8 shows the PDFs of the third-order fits for mass bins III and V compared to the simulated density contrast. This indicates that the local bias model does not capture all the properties of the halo density field.

Figure 4.9 shows again two slices of the simulation along one axis, extending over the box length $L = 150\,h^{-1}\,$Mpc. Overplotted are the fits of different order to $\delta^{\rm s}_{\rm m}$ (left panel) and to $\delta^{\rm s}_{\rm SPT}$ (right panel). In the former case, the different fits behave very similarly. In the latter case, the third-order fit (black long-dashed)

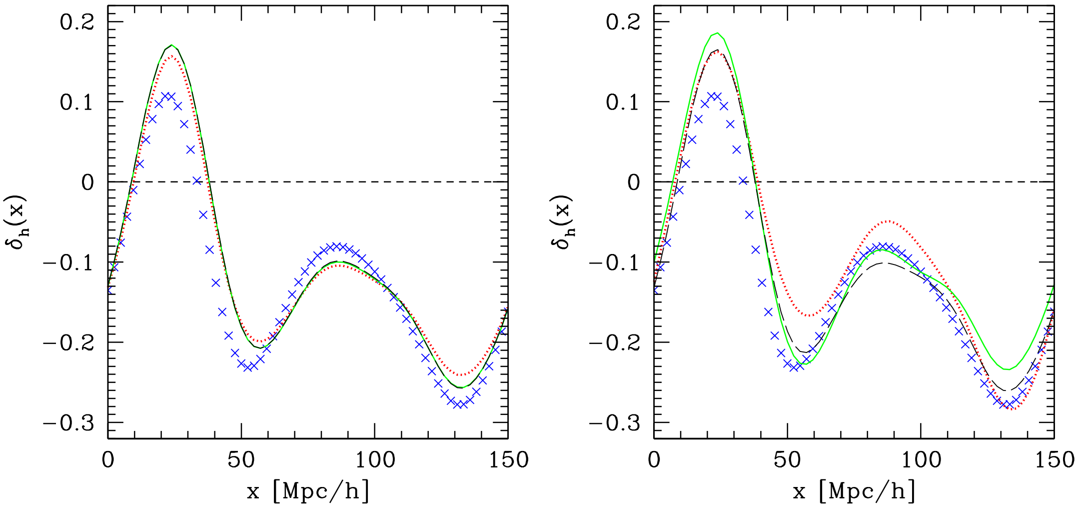

Figure 4.9: Left panel: Amplitude of the halo density contrast from the simulation (blue crosses) compared to $\delta_{\mathrm{h,f}_i}(\boldsymbol{x})$. Red dotted: linear fit, green solid: second-order fit, black long-dashed: third-order fit. Right panel: Same for $\delta^{\mathrm{SPT}}_{\mathrm{h,f}_i}(\boldsymbol{x})$. Both panels show mass bin I with $R = 12\,h^{-1}\,$Mpc. Note that this is the same region of the density field as in the left panel of Figure 4.1.

seems to be closest to the simulation. It is clear that in both cases, the local bias model does not resemble the simulated halo density field on small scales. We also show the results for a medium and a high mass bin (III and V) in Figure 4.10, but here we show only the fits to $\delta^{\mathrm{s}}_{\mathrm{m}}$. Also for higher masses and larger smoothing scales, the ELB model is not satisfactory on a point-by-point level, although the PDFs shown before were similar.

We also calculate the linear correlation coefficient c and the Spearman ranked correlation coefficient c_{Sp} which allows for a general, non-linear correlation between the simulated and fitted $\delta_{\mathrm{h}}(\boldsymbol{x})$ (Table 4.3). The values are very close to 1, which shows as well that overall, the local bias assumption is not such a bad model for out data. From now on we will use the third-order fit because it is the model preferred by the AIC, the correlations also support this choice and it is consistent with the order of the SPT calculation.

Mass Dependence

Figure 4.11 shows the mass dependence of the bias parameters (fitting $\delta^{\mathrm{s}}_{\mathrm{h}}$ vs. $\delta^{\mathrm{s}}_{\mathrm{m}}$) for the mass bins defined in Table 4.1 and $R = 12\,h^{-1}\,$Mpc. Different symbols distinguish the different parameters: red squares for b_1, green triangles for b_2 and blue crosses for b_3. The lines show a theoretical prediction for the bias parameters obtained as follows: The first step is to apply the peak-background-split model to the halo mass function in the simulation (from the fitting formula presented in Pillepich et al. 2010). The second step uses the spherical collapse model to relate the Lagrangian and Eulerian bias parameters (Giannantonio & Porciani, 2010). The two sets of bias parameters show the same trend with halo mass, but they are not in a perfect agreement given the jackknife errors bars (see also Manera et al.

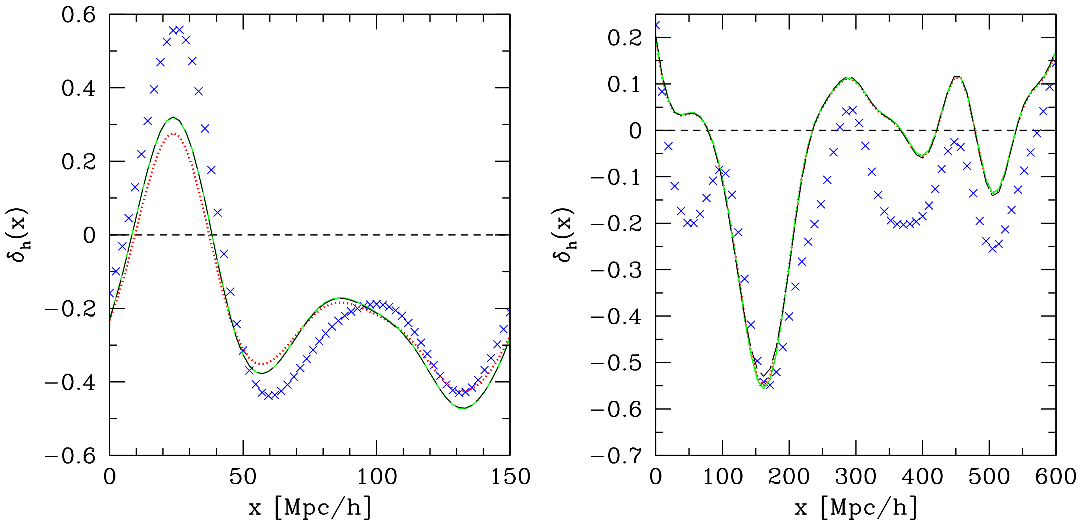

Figure 4.10: Amplitude of the halo density contrast from the simulation (blue crosses) compared to $\delta_{\mathrm{h,f}_i}(\boldsymbol{x})$. Red dotted: linear fit, green solid: second-order fit, black long-dashed: third-order fit. Left panel: Medium mass bin III with $R = 12\,h^{-1}$ Mpc. (Same region as in Figure 4.9). Right panel: High mass bin V with $R = 28\,h^{-1}$ Mpc.

Table 4.3: Correlation coefficients for the halo density contrast for a selection of mass bins and smoothing scales R $[\,h^{-1}\,\mathrm{Mpc}]$.

Mass bin	R	Pair	c	c_{Sp}
I	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_1}\rangle$	0.987	0.989
I	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_2}\rangle$	0.990	0.989
I	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_3}\rangle$	0.990	0.989
I	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_1}^{\mathrm{spt}}\rangle$	0.951	0.961
I	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_2}^{\mathrm{spt}}\rangle$	0.985	0.986
I	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_3}^{\mathrm{spt}}\rangle$	0.989	0.989
III	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_1}\rangle$	0.934	0.956
III	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_2}\rangle$	0.943	0.956
III	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_3}\rangle$	0.043	0.956
III	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_1}^{\mathrm{spt}}\rangle$	0.902	0.922
III	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_2}^{\mathrm{spt}}\rangle$	0.937	0.953
III	12	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_3}^{\mathrm{spt}}\rangle$	0.944	0.957
V	28	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_1}\rangle$	0.876	0.875
V	28	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_2}\rangle$	0.876	0.875
V	28	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_3}\rangle$	0.877	0.875
V	28	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_1}^{\mathrm{spt}}\rangle$	0.866	0.869
V	28	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_2}^{\mathrm{spt}}\rangle$	0.876	0.875
V	28	$\langle\delta_{\mathrm{h}}^{\mathrm{s}}\ \delta_{\mathrm{h,f}_3}^{\mathrm{spt}}\rangle$	0.877	0.875

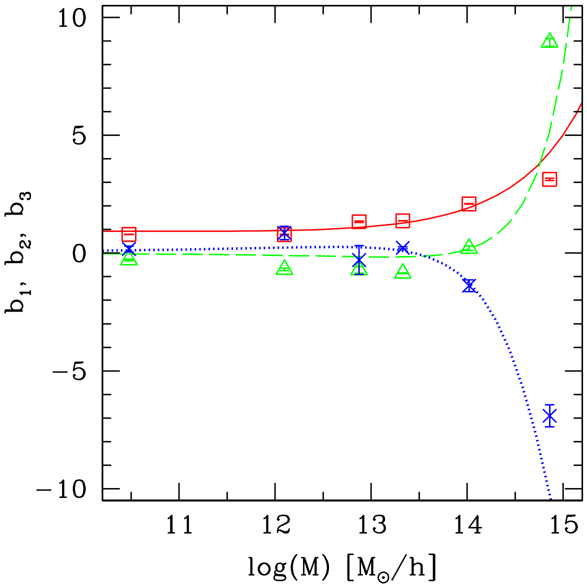

Figure 4.11: Mass dependence of the fitted bias parameters for $R = 12\,h^{-1}$ Mpc. Symbols: fit parameters, lines: predictions for the simulated halo mass function. b_1 (red squares and solid line), b_2 (green triangles and dashed line) and b_3 (blue crosses and dotted line).

2010). The error in b_3 is strongly influenced by shot noise, because there are fewer points in the very high and low density regions which determine the shape of the polynomial.

4.5.2 Halo power spectra

In this section we test the accuracy of SPT and ELB for the halo power spectrum. In order to calculate the SPT halo power spectra, we have to re-arrange $\delta^{\mathrm{SPT}}_{\mathrm{h},f_i}(\boldsymbol{x})$ into terms of the same order in $\delta^{\mathrm{s}}_1(\boldsymbol{x})$, e. g.

$$\delta^{(2)}_{\mathrm{h}}(\boldsymbol{k}) \equiv b_1 \delta^{\mathrm{s}}_2(\boldsymbol{x}) + \frac{b_2}{2}(\delta^{\mathrm{s}}_1(\boldsymbol{x}))^2. \tag{4.25}$$

The biased, third-order SPT halo power spectrum then consists of three terms:

$$P^{\mathrm{SPT}}_{\mathrm{f}_3}(k) \equiv P_{\mathrm{h},11}(k) + P_{\mathrm{h},22}(k) + 2\,P_{\mathrm{h},13}(k), \tag{4.26}$$

with $P_{\mathrm{h},mn}(k) \propto \langle \tilde{\delta}^{(m)}_{\mathrm{h}}(\boldsymbol{k})\, \tilde{\delta}^{(n)}_{\mathrm{h}}(\boldsymbol{k'})\rangle$, as in Equation (4.6). In this way we can make sure that also the halo power spectrum does not contain terms of order higher than δ^4_1.

Figure 4.12 shows the ratio of the reconstructed halo power spectrum to the one from the simulation (left panel: large box, right panel: small box), which has been corrected for (Poissonian) shot noise in the following way:

$$P_{\mathrm{h}}(k) \equiv P^{\mathrm{sim}}_{\mathrm{h}}(k) - \frac{L^3}{\overline{N}}, \tag{4.27}$$

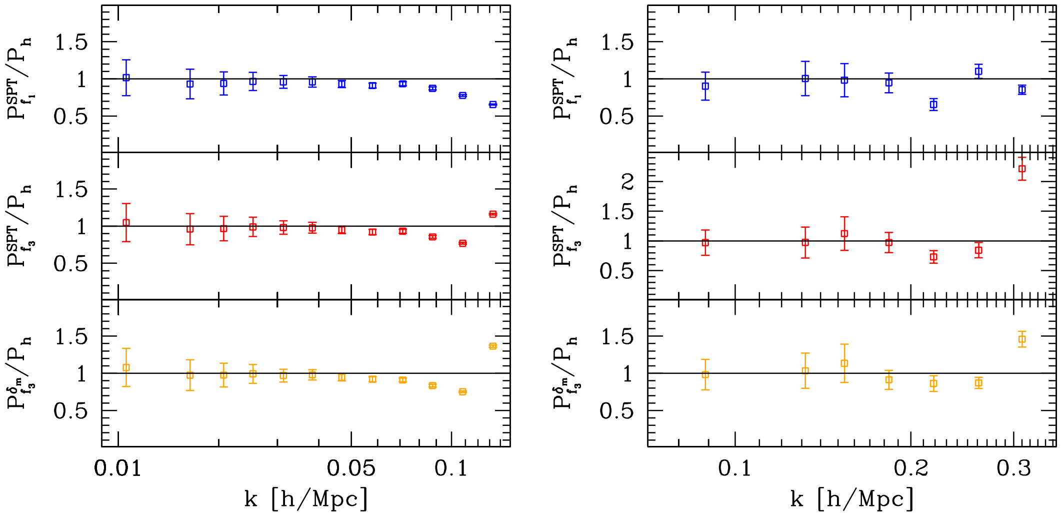

Figure 4.12: Ratios of *halo power spectra* from our fits to the simulation. Left panel: big simulation volume (mass bin IV, $R = 28\,h^{-1}\,\mathrm{Mpc}$), right panel: small simulation volume (mass bin I, $R = 12\,h^{-1}\,\mathrm{Mpc}$). From top to bottom: linear fit to the linear matter density; third-order fit to full SPT density; third-order fit to the non-linear matter density field.

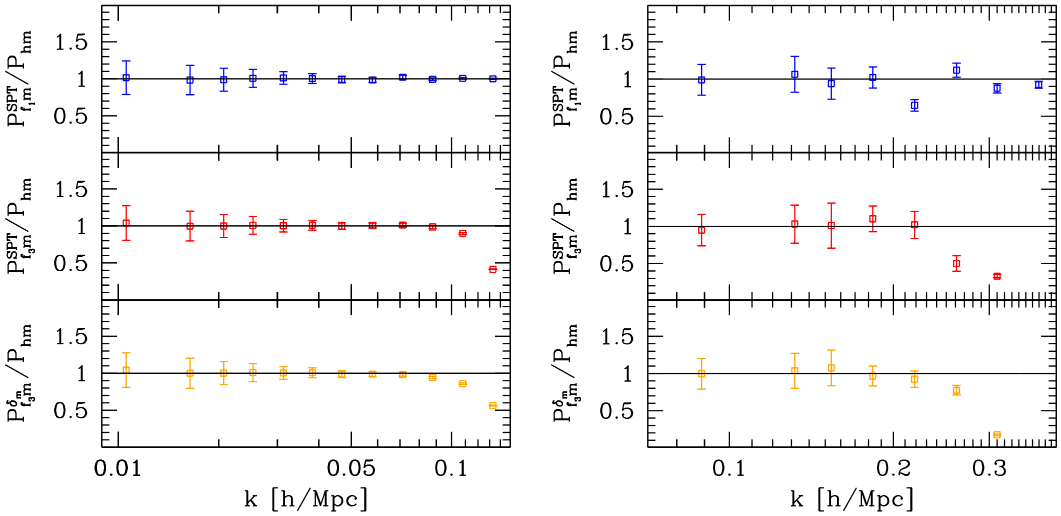

Figure 4.13: Ratios of *halo-matter cross spectra* from our fits to the simulation. Left panel: big simulation volume (mass bin IV, $R = 28\,h^{-1}\,\mathrm{Mpc}$), right panel: small simulation volume (mass bin I, $R = 12\,h^{-1}\,\mathrm{Mpc}$). From top to bottom: linear fit to the linear matter density; third-order fit to full SPT density; third-order fit to the non-linear matter density field.

where $\overline{N}$ is the number of halos in each bin and L is the box size. The top panel on each side corresponds to the power spectrum of $\delta^{\mathrm{SPT}}_{\mathrm{h,f_1}}$, while the middle panel corresponds to $\delta^{\mathrm{SPT}}_{\mathrm{h,f_3}}$. In the lowest panel we consider $\delta_{\mathrm{h,f_3}}$. Two effects can be noticed: In general, both linear and third-order SPT underestimates the power apart from the largest scales in both boxes. However, when $kR \approx \pi$, the unsmoothing effect sets in, eventually leading to a diverging ratio. The lowest panel shows that this behaviour is not caused by SPT, because $\delta^{\mathrm{SPT}}_{\mathrm{h,f_3}}$ and $\delta_{\mathrm{h,f_3}}$ have very similar power spectra.

We also tried to reverse the order of smoothing and fitting in the following way: Determining the bias parameters from a scatter plot of the unsmoothed[4] halo and matter densities, then multiplying the (unsmoothed) matter density with these new bias parameters and applying Equation (4.8) to the resulting biased halo field. The halo power spectra from this fit do not show the unsmoothing effect by construction, but are also incompatible with the simulation halo power spectra on any scale. In fact, not even the linear bias parameter from this new method agrees with the previous estimates, so that the halo power spectrum is different from the simulation even on large scales where the smoothing should have no effect (because $P_{\mathrm{h}} \approx b_1^2 P_{\mathrm{m}}$). We therefore conclude that (not surprisingly) the order of smoothing and fitting can not be reversed, and one has to live with the limitations of the non-linear local bias model that re-introduces some small-scale fluctuations.

We show the ratio of the halo-matter cross spectra in Figure 4.13, the order of the panels is the same as in the previous figure. The cross spectrum is not affected by shot noise, and ELB spectra agree with the simulation over a larger range of scales for the big box. Contrary to the halo power spectra, the unsmoothing effect leads to an additional loss in power for the cross spectra on the very small scales, particularly evident for the small box. Note that the best agreement between simulation and fits is actually achieved using the linear bias parameter and the linear matter power spectrum (top panel). For the large box, this is not surprising because the scales in consideration are large enough for linear theory to be a good approximation. For the small box, linear theory is underestimating the power for the small scales, as seen before in Figs. Figure 4.5 and Figure 4.12, but only on scales where the cross spectra from the higher-order fits are also too low. This means that the unsmoothing effect (which does not affect the top panel) has an even stronger influence on the halo power spectra from SPT+ELB than the deviation from linear theory on small scales, at least within the accuracy of our simulations. In order to investigate the systematic effects seen in Figs. Figure 4.12 and Figure 4.13, we now consider the influence of the fit residuals on auto and cross spectra, which include both the intrinsic scatter as well as the unsmoothing effect from the local bias scheme. We define the residual $\Delta(\boldsymbol{x})$ for the third-order fit by

$$\Delta(\boldsymbol{x}) \equiv \delta^{\mathrm{s}}_{\mathrm{h}}(\boldsymbol{x}) - \delta_{\mathrm{h,f_3}}(\boldsymbol{x}), \tag{4.28}$$

---

[4] That is without applying (Equation (4.8)). Of course, the fields are always smoothed at least on the scale of the grid used in the CIC algorithm.

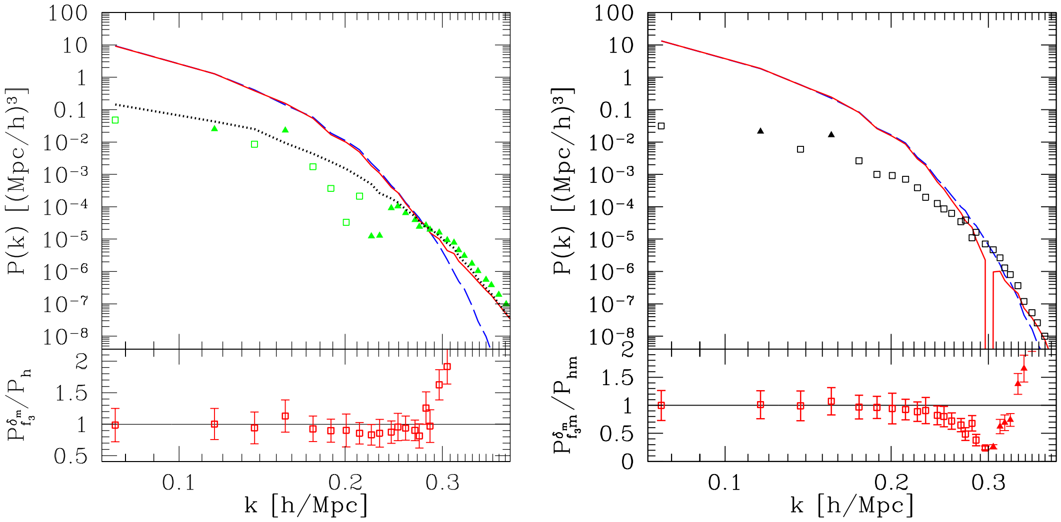

Figure 4.14: Left panel: Halo power spectrum and residuals for $\delta_{\mathrm{h,f_3}}(\boldsymbol{x})$ (upper panel, mass bin I, $R = 12\,h^{-1}\,\mathrm{Mpc}$). Dashed blue line: $P_{\mathrm{h}}(k)$, solid red line: $P_{\mathrm{h,f_3}}(k)$, dotted black line: $P_{\Delta\Delta}(k)$, points: the cross-term $2\,P_{\Delta\,\delta_{\mathrm{h,f_3}}}(k)$. Open squares show positive values, solid triangles show negative values of the cross-term. Ratio between $P_{\mathrm{h,f_3}}(k)$ and $P_{\mathrm{h}}(k)$ (lower panel). Right panel: Same for the halo matter cross spectrum. Dashed blue line: $P_{\mathrm{hm}}(k)$, points: $P_{\Delta\,\delta_{\mathrm{m}}}(k)$, solid red line: $P_{\delta_{\mathrm{h,f_3}}\delta_{\mathrm{m}}}(k)$. Note the change in sign at $k \approx 0.3\,h\,\mathrm{Mpc}^{-1}$.

from which

$$\begin{aligned} P_{\mathrm{h}}(k) &= P_{\mathrm{h,f_3}}(k) + 2\,P_{\Delta\,\delta_{\mathrm{h,f_3}}}(k) + P_{\Delta\Delta}(k), \\ P_{\mathrm{hm}}(k) &= P_{\delta_{\mathrm{h,f_3}}\delta_{\mathrm{m}}}(k) + P_{\Delta\,\delta_{\mathrm{m}}}(k), \end{aligned} \tag{4.29}$$

follows for the halo power spectrum and the halo matter cross spectrum. The left panel of Figure 4.14 shows the different terms for $P_{\mathrm{h}}(k)$: $P_{\mathrm{h,f_3}}(k)$ (solid red) is very close to $P_{\mathrm{h}}(k)$ (dashed blue) until $k \approx 0.3\,h\,\mathrm{Mpc}^{-1}$. The cross-term $2\,P_{\Delta\,\delta_{\mathrm{h,f_3}}}(k)$ (green points) can have both positive and negative values. This is indicated in the figure by the different symbols: open squares for positive values, solid triangles for negative values. Finally, $P_{\Delta\Delta}(k)$ is shown with a black dotted line. Generally, $2\,P_{\Delta\,\delta_{\mathrm{h,f}_i}}(k)$ and $P_{\Delta\Delta}(k)$ have a small amplitude, but their contribution becomes important on smaller scales. The lower panel shows the ratio of the fit and the simulated halo power spectrum for comparison. The region where $P_{\mathrm{h,f_3}}(k)$ is systematically lower than $P_{\mathrm{h}}(k)$ roughly corresponds to the region where the cross-term is positive.

The right panel of Figure 4.14 shows the terms contributing to $P_{\mathrm{hm}}(k)$ (blue dashed): the cross spectrum of the fit and the dark matter (red solid) and the cross spectrum of the residuals and the dark matter (black points, squares positive values, triangles negative values). In contrast to the fitted halo power spectrum which rises on small scales, the fitted cross spectrum becomes negative at around the same scales, indicated by the "hole" in the red solid line (for smaller scales we then show $-P_{\delta_{\mathrm{h,f_3}}\delta_{\mathrm{m}}}$). The lower panel shows the ratio of the fit and the simulation cross spectrum. As before, squares correspond to positive values of this ratio, and

Table 4.4: Comparing b_1, $\bar{b}_1$, b_{eff}, $b^{\mathrm{H}}_{\mathrm{eff}}$ and $b^{\mathrm{spt}}_{\mathrm{eff}}$ for different mass bins and smoothing scales R. $\langle\cdot\rangle$ here denotes the average over a specific k-range (see text). R is given in h^{-1} Mpc. Errors are jackknife errors using 8 subsamples.

Mass bin	R	b_1	$\bar{b}_1$	$\langle b_{\mathrm{eff}}\rangle$	$\langle b^{\mathrm{H}}_{\mathrm{eff}}\rangle$	$\langle b^{\mathrm{spt}}_{\mathrm{eff}}\rangle$
I	6	0.773 ± 0.005	0.573 ± 0.011	0.684 ± 0.051	0.606 ± 0.040	0.623 ± 0.028
II	6	0.928 ± 0.011	0.680 ± 0.013	0.756 ± 0.057	0.484 ± 0.067	0.800 ± 0.032
III	6	1.476 ± 0.030	1.108 ± 0.043	1.198 ± 0.091	0.722 ± 0.112	1.100 ± 0.061
I	12	0.719 ± 0.009	0.694 ± 0.001	0.684 ± 0.051	0.545 ± 0.031	0.703 ± 0.044
IV	28	1.279 ± 0.008	1.278 ± 0.001	1.253 ± 0.076	1.056 ± 0.042	1.243 ± 0.059
IV	50	1.244 ± 0.028	1.241 ± 0.028	1.253 ± 0.076	1.185 ± 0.060	1.237 ± 0.054
V	28	2.040 ± 0.023	2.037 ± 0.022	1.958 ± 0.119	1.553 ± 0.295	1.947 ± 0.095
V	50	2.029 ± 0.042	2.016 ± 0.039	1.958 ± 0.119	1.697 ± 0.111	1.972 ± 0.089
VI	28	3.704 ± 0.054	3.615 ± 0.049	3.520 ± 0.222	4.739 ± 0.134	3.662 ± 0.159
VI	50	3.761 ± 0.088	3.739 ± 0.087	3.520 ± 0.222	4.461 ± 0.296	3.877 ± 0.155

triangles show where it is negative.

We conclude that the halo auto and cross spectra computed with the ELB deviate from the simulation on scales where $kR \approx \pi$. This affects especially the auto spectra, and is not related to using SPT instead of the non-linear matter field.

Biased SPT on Large Scales

Heavens et al. (1998) discuss two effects on the halo power spectrum that are caused by using SPT. The basis for their analysis is as follows: as for the case of the one-loop matter power spectra, one can also express the SPT halo power spectra as integrals over products of the linear matter power spectrum and the bias parameters (Jain & Bertschinger, 1994). The $P_{\mathrm{h},22}$ and $P_{\mathrm{h},13}$ terms defined before can then be written as

$$\begin{aligned} P_{\mathrm{h},22}(k) = {} & 2\int \frac{d^3q}{(2\pi)^3} P_{11}(q) P_{11}(|\boldsymbol{k}-\boldsymbol{q}|) \times \\ & \left[b_1 F_2^{(s)}(\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}) + \frac{b_2}{2} \right]^2 , \\ P_{\mathrm{h},13}(k) = {} & 6\, b_1 P_{11}(k) \int \frac{d^3q}{(2\pi)^3} P_{11}(q) \left[b_1 F_3^{(s)}(\boldsymbol{q},-\boldsymbol{q},\boldsymbol{k}) \right. \\ & \left. + \frac{b_3}{6} + b_2 F_2^{(s)}(-\boldsymbol{q},\boldsymbol{k}) \right] , \end{aligned} \tag{4.30}$$

and $P_{\mathrm{h},11}(k) = b_1^2 P_{11}(k)$. By taking the limit $k \to 0$ of the foregoing equations, one can study the behaviour of the SPT halo power spectrum on large scales:

1. **Is the large-scale bias not b_1?** While the linear bias model should be valid on large scales, Heavens et al. (1998) predict that the large-scale bias is not b_1, but can be approximated by an effective bias

$$b^{\mathrm{H}}_{\mathrm{eff}} = \sqrt{b_1^2 + b_1 \left(\frac{68}{21} b_2 + b_3 \right) \sigma_R^2}, \tag{4.31}$$

where σ_R^2 is the variance of the smoothed linear density field at redshift 0:

$$\sigma_R^2 \equiv \int \frac{\mathrm{d}^3q}{(2\pi)^3} P_{11}(q)\, e^{-(qR)^2}. \tag{4.32}$$

Equation (4.31) tells us that the difference between $b_{\mathrm{eff}}^{\mathrm{H}}$ and b_1 depends on the choice of the smoothing scale through σ_R^2. However, b_1 is supposed to describe the large-scale behaviour of the power spectrum, where the smoothing scale should not have any effect. The value of σ_R^2 will be large for smaller R and b_2 is negative in many cases, so the term in parenthesis can even become negative, leading to $b_{\mathrm{eff}}^2 < 0$. (Although this never happens for the smoothing scales we investigate.) We can compare this with the effective bias directly from the simulation power spectra

$$b_{\mathrm{eff}} \equiv \sqrt{\frac{P_{\mathrm{h}}}{P_{\mathrm{m}}}} \tag{4.33}$$

on large scales ($k \leq 0.02\, h\,\mathrm{Mpc}^{-1}$ for the large box and $k \leq 0.15\, h\,\mathrm{Mpc}^{-1}$ for the small box), where this ratio is found to be constant within the errorbars. Note that P_{h} has been corrected for shot noise, and for this specific ratio, the dependence on the smoothing scale cancels out. Examples of the different bias parameters are shown in Table 4.4, for mass bins in both boxes and several smoothing scales. Here, b_1 and $\bar{b}_1$ are the linear bias parameters from the third-order fits, and $b_{\mathrm{eff}}^{\mathrm{H}}$ was also calculated from these values. We always find the large-scale bias to be very close to b_1 for reasonable[5] smoothing scales. Note that the value for b_{eff} for the small box (mass bins I-III) is only an approximation for the large-scale bias, because the box is so small. Comparing b_1, $\bar{b}_1$ and $\langle b_{\mathrm{eff}} \rangle$ for the low mass bins, it is clear that the different estimates do not agree with each other if the smoothing scale gets too small ($R = 6\, h^{-1}\,\mathrm{Mpc}$). Note also that exchanging P_{h} with one of the fitted halo power spectra in Equation (4.33) does not significantly affect the value of b_{eff} (exchanging $P_{\mathrm{m}} \to P_{\mathrm{SPT}}$ for the SPT fit as well).

Extending the comparison to $b_{\mathrm{eff}}^{\mathrm{H}}$ reveals a large discrepancy to the previous values over the whole mass range and different smoothing scales. Even though the errors on $b_{\mathrm{eff}}^{\mathrm{H}}$ are rather large (owing to the large errors in $\bar{b}_3$), the estimates are quite different from the fit parameters and b_{eff}.

In concordance with the derivation of Equation (4.31) in Heavens et al. (1998), we define an additional effective bias for the SPT fit:

$$b_{\mathrm{eff}}^{\mathrm{spt}} \equiv \sqrt{\frac{P_{\mathrm{f_3}}^{\mathrm{SPT}}}{P_{11}}}, \tag{4.34}$$

which is also averaged over the same scales as b_{eff}. The difference to Equa-

---

[5] Meaning that the SPT assumptions about the smallness of δ are still valid at $z = 0$, requiring $R \geq 8\, h^{-1}\,\mathrm{Mpc}$.

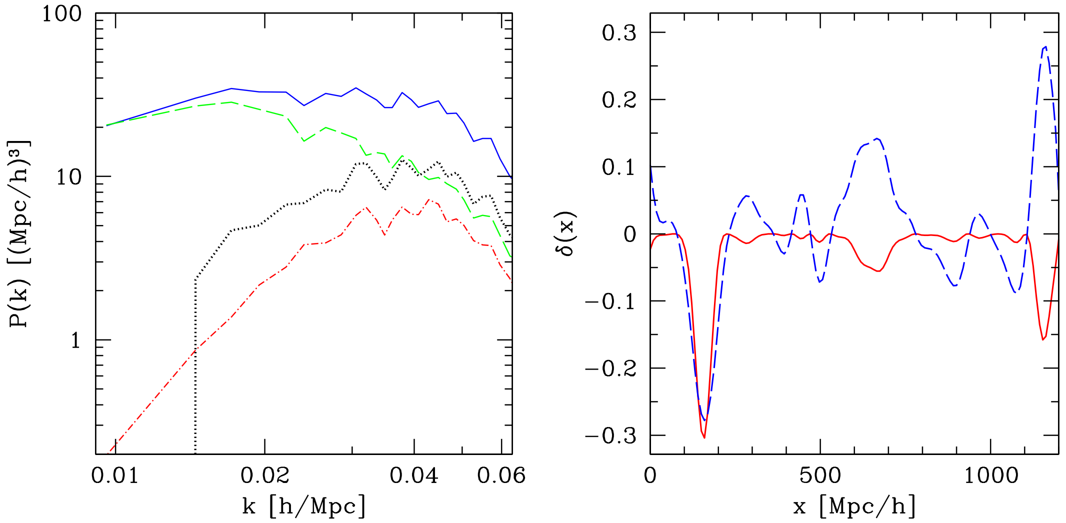

Figure 4.15: Left panel: Different contributions to $P_{\mathrm{h},22}(k)$ (see text) and their sum (blue solid) for mass bin IV with $R = 28\,h^{-1}\,\mathrm{Mpc}$. The constant power on large scales is dominated by $b_2^2\langle\tilde{\delta}_1^2(\boldsymbol{k})\tilde{\delta}_1^2(\boldsymbol{k'})\rangle/4$ (green line). Right panel: The contribution $A \cdot \delta_1^2(\boldsymbol{x})$ (solid red, with $A = 5\,b_2/2$ for enhanced visibility), which coincides with the peaks and troughs of the underlying SPT matter density field (dashed blue) as suggested by Heavens et al. (1998).

tion (4.33) comes because the denominator contains only the linear power spectrum. However, the values of $b_{\mathrm{eff}}^{\mathrm{spt}}$ do not differ much from b_{eff} because on large scales $P_{\mathrm{m}} \approx P_{\mathrm{SPT}} \approx P_{11}$.

We conclude that b_{eff} matches the fit parameters b_1 and $\bar{b}_1$ within the error-bars. Equation (4.31) was derived neglecting the contribution of $P_{\mathrm{h},22}(k)$, which seems not to be a valid approximation, as the values for $b_{\mathrm{eff}}^{\mathrm{H}}$ do not agree with b_{eff} and $b_{\mathrm{eff}}^{\mathrm{spt}}$. It is also important to note that the agreement of the different bias estimations gets worse when smaller smoothing scales are considered. This implies that using the bias parameters to infer halo masses (by choosing R to correspond to the Lagrangian radius of the halo) can lead to wrong results.

2. **Is the SPT halo power spectrum constant on very large scales?** From Equation (4.30) we see that the $P_{\mathrm{h},22}(k)$-term is not directly proportional to $P_{11}(k)$ which falls off as $k \to 0$. It follows that this term will eventually dominate, and it can be shown to lead to a constant halo power spectrum on very large scales. We can study this behaviour using the SPT halo density contrast from the third-order fit. Even with our large simulation volume, we can only see that $P_{\mathrm{h},22}(k)$ tends to a constant, but the scales where it actually dominates are out of reach. However, we can look in more detail at the $P_{\mathrm{h},22}(k)$-term, to find out why it becomes constant. From Equation (4.25), we can see that there a three terms that contribute, which are shown in the left panel of Figure 4.15: $b_1^2 P_{22}(k)$ (red dot-dashed), the cross term $b_1 b_2 \langle\tilde{\delta}_2^{\mathrm{s}}(\boldsymbol{k})\tilde{\delta}_1^2(\boldsymbol{k'})\rangle$ (black dotted) and $b_2^2\langle\tilde{\delta}_1^2(\boldsymbol{k})\tilde{\delta}_1^2(\boldsymbol{k'})\rangle/4$ (green dashed). The latter term clearly dominates $P_{\mathrm{h},22}(k)$ (blue solid line) on large scales.

Heavens et al. (1998) suggest that this constant power comes from the peaks and troughs of the underlying density field, which we can confirm: The right panel of Figure 4.15 shows the SPT matter density contrast $\delta^{\mathrm{s}}_{\mathrm{SPT}}(\boldsymbol{x})$ and $b_2\,\delta_1^2(\boldsymbol{x})/2$ along a line in the simulation volume. The latter term can have either sign depending on b_2, here we show the mass bin IV where $b_2 < 0$. Note that $b_2\,\delta_1^2/2$ has been multiplied by a factor of 5 in the Figure to make the effect more visible.

4.6 Conclusion

In this chapter we have followed a novel approach by evaluating the SPT expansion of the mass density and velocity fields up to third order, starting from the same realisations of the linear density field that has been used to generate the initial conditions of two N-body simulations. This allowed us to make a point-by-point comparison between the non-linear mass and halo overdensities (using the Eulerian local bias model), while past studies have only focussed on two- or three-point statistics. Our results can be summarized as follows:

- We found that SPT provides a good approximation to the density field up to redshift 0, for smoothing scales $R \geq 8\,h^{-1}\,\mathrm{Mpc}$. This ensures that the linear density contrast is typically less than one.

- We have compared the redshift evolution of the non-linear matter density field of the simulations and SPT with the prediction of linear theory and the spherical collapse model. We found that SPT is very close to the simulated density for all redshifts, while both linear theory and the spherical collapse model are in poor agreement with the simulations. The lognormal model by Kayo et al. (2001) gives a good PDF for the non-linear density contrast but fails when used for a point-by-point comparison.

- In Equation (4.13), we presented a simple fitting formula for the non-linear density contrast as a function of the linear density contrast, which is accurate at the per cent level over the full range of redshifts available to us ($0 \leq z \leq 10$).

- We compared the SPT matter power spectra with the linear and non-linear matter power spectra from the simulation. On large scales, linear theory provides a good approximation to the simulated matter power spectra, but SPT is superior on smaller scales, up to the maximum wavenumber we can probe, $k \approx 0.3\,h\,\mathrm{Mpc}^{-1}$ at redshift 0.

- Assuming a deterministic, Eulerian local bias model (ELB) with up to 3 free parameters, we obtained values for these parameters by fitting polynomials to a scatter plot of the smoothed matter and halo density contrast. We find that the third-order bias model is always preferred by the data over models with less parameters. The reconstructed halo density is similar to the simulation, but the ELB can not accurately reproduce the simulated field.

- We found that the bias parameters from fitting the halo distribution to the simulated matter density field and to the SPT density field generally differ, but the corresponding power spectra are very similar. The mass dependence of the bias parameters shows the same trend as theoretical predictions based on the peak-background-split approach.

- We compared both the halo-halo power spectrum and the halo-matter cross spectrum from the fits with the ones from the simulation. The cross spectrum is in better agreement with the simulation than the auto spectrum which deviates from the simulation on scales much larger than the smoothing radius. This is not related to using SPT instead of the non-linear matter field.

- We have investigated two effects on the halo power spectra, which have been predicted by analytic considerations of SPT and the local bias model. First, we estimated the large-scale bias b_{eff} using the halo and mass power spectra from both the simulation and the fits. This large-scale bias is compatible with the linear bias parameter obtained from the polynomial fit if we make sure that the SPT assumptions are not violated, choosing the smoothing radius R such that $\delta \ll 1$. This suggests that the effective bias does not require perturbative corrections, contrary to previous results based on SPT. Second, we determined the origin of the constant shot-noise term on very large scales, which is caused by $b_2^2\langle\tilde{\delta}_1^2\tilde{\delta}_1^2\rangle/4$ as predicted in Heavens et al. (1998).

In summary, our study shows that SPT is a suitable approximation for the matter field even at redshift 0, provided a large enough smoothing radius is adopted. However, the Eulerian local bias model can not fully describe the halo density field, which is most evident from our point-by-point comparison in Figure 4.9.

5 Primordial non-Gaussianity and the large-scale halo bias

This chapter represents a modified version of Roth & Porciani (2012) with small changes in notation, additional references, and a brief discussion of the Planck *results and recent developments in the last section.*

5.1 Introduction

As we have shown in the previous chapter, galaxy (or halo) bias restricts our ability to accurately describe cosmological structure formation and to determine cosmological parameters. However, in the case of primordial non-Gaussianity (PNG), the bias between galaxies and dark matter can actually provide *additional information.* It turns out that dark matter and galaxies react differently to the change in the underlying potential, and this leads to deviation of the bias from the case of Gaussian initial conditions predicted from standard single-field inflation.

Currently, the most stringent constraints on PNG come from the Cosmic Microwave Background (CMB) bispectrum (e. g. Bennett et al., 2012; Planck Collaboration et al., 2013b). However, having another channel for a *detection* of PNG is important, as any non-zero value for $f_{\rm NL}$ etc. would imply some deviation from the standard single-field inflationary model. Additionally, results from large-scale structure probe a different range in redshift and scales, and can thus constrain PNG models that predict e. g. $f_{\rm NL}(k)$.

Here we consider the local model of PNG, in which the Bardeen curvature potential during matter domination (at high-redshift) is given by

$$\Phi_{\rm B}(\boldsymbol{x}) = \phi(\boldsymbol{x}) + f_{\rm NL}\left[\phi(\boldsymbol{x})^2 - \langle\phi^2\rangle\right] + g_{\rm NL}\left[\phi(\boldsymbol{x})^3 - 3\langle\phi^2\rangle\phi(\boldsymbol{x})\right], \tag{5.1}$$

where $\phi(\boldsymbol{x})$ is a zero-mean Gaussian random field, $f_{\rm NL}$ and $g_{\rm NL}$ are real constants, and the symbol $\langle\cdot\rangle$ denotes expectation values.[1] On large scales, PNG of the local type introduces a scale-dependent bias between galaxies and the underlying matter distribution (Dalal et al. 2008, Matarrese & Verde 2008, Slosar et al. 2008, Giannantonio & Porciani 2010), which could provide a promising signal for the detection of non-zero PNG even at late times.

Several authors have derived observational constraints on $f_{\rm NL}$ and $g_{\rm NL}$ from the galaxy power spectrum or 2-point correlation function (Slosar et al. 2008,

---

[1] These terms are required in order to ensure that the resulting non-Gaussian density field still has a zero mean.

Desjacques & Seljak 2010b), and recently Xia et al. (2011) reported a possible detection of positive $f_{\rm NL}$ at the 2σ level. However, these studies all assume that the underlying model for PNG is either purely quadratic ($g_{\rm NL} = 0$) or purely cubic ($f_{\rm NL} = 0$). In this chapter, we show that the assumption of a one-parameter model significantly biases the estimation of the PNG parameters when both $f_{\rm NL}$ and $g_{\rm NL}$ do not vanish. We also illustrate that model-selection techniques like the Bayes factor are powerful tools to identify the model that best describes the available data. This is particularly interesting in light of the next generation of galaxy and cluster surveys like eROSITA (Predehl et al., 2010), Euclid (Laureijs et al., 2011) or LSST (Ivezić et al., 2008b) which are expected to significantly reduce the error bars of PNG parameters (e. g. Pillepich et al. 2012, Giannantonio et al. 2012).

5.2 Scale-dependent bias

The linear matter density contrast in Fourier space $\tilde{\delta}_{\rm m}(k)$ follows from the curvature perturbations by

$$\tilde{\delta}_{\rm m}(k, z) = \alpha(k, z)\tilde{\Phi}(k), \tag{5.2}$$

with

$$\alpha(k, z) = \frac{2k^2 D(z)T(k)}{3\Omega_{\rm m} H_0^2}, \tag{5.3}$$

where $\Omega_{\rm m}$ and H_0 are the current matter-density parameter and Hubble constant, $T(k)$ is the transfer function, and $D(z)$ is the growth factor normalised to $(1+z)^{-1}$ during the epoch of matter domination. (For more details see Chapter 2.)

Let us denote by $P_{\rm m}(k)$ the matter power spectrum and by $P_{\rm hm}(k)$ the cross spectrum between dark-matter halos (of a given mass) and mass. On large scales, the relation between $P_{\rm hm}(k)$ and $P_{\rm m}(k)$ is given by

$$\frac{P_{\rm hm}(k)}{P_{\rm m}(k)} = b_1(f_{\rm NL}, g_{\rm NL}) + \Delta b(k, f_{\rm NL}, g_{\rm NL}), \tag{5.4}$$

where b_1 is the linear bias parameter (a real-valued constant) and $\Delta b(k, f_{\rm NL}, g_{\rm NL})$ is a scale-dependent bias term introduced by PNG. Following the work of Smith et al. (2012) (hereafter SFL12), the latter term can be split up into different contributions (to first order in $f_{\rm NL}$ and $g_{\rm NL}$):

$$\Delta b(k, f_{\rm NL}, g_{\rm NL}) = \frac{\beta_f f_{\rm NL} + \beta_g g_{\rm NL}}{\alpha(k)}. \tag{5.5}$$

Its dependence on $f_{\rm NL}$ was shown to be (Dalal et al. 2008, Matarrese & Verde 2008, Slosar et al. 2008, Giannantonio & Porciani 2010):

$$\beta_f = 2\,(\tilde{\nu}^2 - 1), \tag{5.6}$$

with the peak height

$$\tilde{\nu} \equiv [\delta_c(b_1 - 1) + 1]^{1/2} \tag{5.7}$$

where δ_c is a parameter set to ~ 1.4 to fit the clustering properties of friends-of-friends halos in numerical simulations. Note that b_1 modulates the amplitude of the scale-dependent bias though $\tilde{\nu}$. [2] Similarly, the dependence on $g_{\rm NL}$ is (SFL12)

$$\beta_g = \kappa_3^{(1)}(M)\left[-0.7+1.4(\tilde{\nu}-1)^2+0.6(\tilde{\nu}-1)^3\right]-\frac{d\kappa_3^{(1)}(M)}{d\log\sigma^{-1}}\left(\frac{\tilde{\nu}-\tilde{\nu}^{-1}}{2}\right) \tag{5.8}$$

with $\kappa_3^{(1)}$ the skewness of the mass density field (smoothed on the mass scale M) in the case of $f_{\rm NL}=1$ and σ the linear root mean square (r. m. s.) mass-density fluctuation on the scale M. The skewness and its derivative can be expressed in terms of b_1 and z (see Equations 46 and 47 in SFL12).

In brief, PNG of the local type makes the bias of dark-matter halos scale dependent and proportional to k^{-2} for $k \to 0$. It is therefore conceivable to constrain PNG from measurements of the galaxy power spectrum, provided a model that links galaxies to halos. However, the $f_{\rm NL}$ and $g_{\rm NL}$ contributions to the scale-dependent bias in Equation (5.5) show the same scaling with the wavenumber and it is not possible to measure both parameters from studies of a single population of tracers of the large-scale structure. Given that β_f and β_g display different scaling with halo mass and redshift, this degeneracy can be broken by considering several populations of galaxies, possibly within widely separated redshift bins. Recent studies have accomplished this task by combining a number of observational datasets with simple models where either $f_{\rm NL}$ or $g_{\rm NL}$ are set to zero a priori. With the help of numerical simulations, in this chapter we study the effect of these simplifying assumptions on the determination of PNG.

5.3 Simulations

We have used GADGET-2 (Springel, 2005) to perform two sets of dark-matter-only simulations consisting of four realisations each. Each set corresponds to a different fiducial model of PNG assuming $f_{\rm NL}=50$ and $g_{\rm NL}=\pm 5\times 10^5$, in accordance with upper limits obtained from cosmic-microwave-background studies (Komatsu et al. 2011, Vielva & Sanz 2010). The other cosmological parameters are chosen in concordance with the WMAP5 analysis (Komatsu et al., 2009): $\Omega_{\rm m}=0.279$, $\Omega_\Lambda=0.721$, $\Omega_{\rm b}=0.0462$, $\sigma_8=0.817$, $h=0.7$, $n_s=0.96$. We follow 1024^3 particles within a cubic box of side $L=1200\,h^{-1}\,$Mpc, corresponding to a particle mass of $1.247\times 10^{11}\,h^{-1}\,{\rm M}_\odot$. The initial conditions are generated at redshift 50 using the Zel'dovich approximation to displace particles from a regular Cartesian grid.

Halos are selected using a friends-of-friends halo finder with a linking length of 0.2 times the mean interparticle separation. We consider only objects with at least

---

[2] Some authors write $\tilde{\nu}$ in term of the linear-bias factor obtained from Gaussian initial conditions, but Giannantonio & Porciani (2010) have demonstrated that the full b_1 should appear in Equation (5.7). Note, however, that the correction to b_1 due to PNG is generally small and, in most practical cases, the difference between the two formulations is negligible.

Table 5.1: The top section describes the galaxy samples used in the analysis of Xia et al. (2011): redshift, type, effective linear bias parameter b_{eff} and number of objects N_{gal}. The middle section illustrates the properties of specific halo subsamples from our simulations that have been selected to approximate the observational dataset described above: redshift, mass range (in $10^{13}\,h^{-1}\,\mathrm{M}_{\odot}$), linear bias parameter b_1, number of objects in each mass bin, N_{halo}, after combining the four realisations. We denote this as the "current sample". The bottom section refers to an hypothetical sample of more massive objects that should become available in the coming times. We label this the "future sample".

z	0	0.5	1	1.5	2	Name
Type	-	LRGs	NVSS	QSOs	-	Observations
b_{eff}	-	1.9	2	2.3	-	
N_{gal}	-	10×10^5	1.4×10^5	1.7×10^5	-	
Mass	-	1.6-2.5	1-1.25	1-1.25	-	Current sample
b_1	-	1.8	2	2.9	-	
N_{halo}	-	6×10^5	4×10^5	2.7×10^5	-	
Mass	6.25-9.5	4-6.25	4-6.25	4-6.25	2.5-4	Future sample
b_1	1.8	2.5	4.0	4.9	5.9	
N_{halo}	1.7×10^5	2.1×10^5	1.1×10^5	0.4×10^5	0.4×10^5	

80 particles, so the halo mass range is $10^{13}\,h^{-1}\,\mathrm{M}_{\odot} \leq M_{\mathrm{halo}} \leq 2 \times 10^{15}\,h^{-1}\,\mathrm{M}_{\odot}$ at $z = 0$. We use outputs at $z = \{0, 0.5, 1, 1.5, 2\}$. The scale-dependent bias parameter $b(k)$ for the halos is measured from the ratio of the (k-binned) halo-matter cross spectrum and the matter auto spectrum, averaged over the four realisations. We only consider the range $0.007\,h\,\mathrm{Mpc}^{-1} \leq k \leq 0.04\,h\,\mathrm{Mpc}^{-1}$ where Equation (5.8) holds as shown in SFL12. The errors on $b(k)$ are calculated using an estimator derived in the appendix of Smith & LoVerde (2011).

To make sure that Equations (5.5)-(5.8) (with $\delta_c = 1.42$) can also be applied to our simulations, we fix f_{NL} and g_{NL} to the input values and use a standard χ^2-fit to measure b_1 as a function of halo mass and redshift. We find good χ^2 values in all cases, and the fitted b_1 is compatible with the value expected for the case of Gaussian initial conditions (see Equation 14 from Pillepich et al. 2010). We can therefore safely assume that the fitting formulae given by SFL12 accurately describe the scale-dependent bias for our sets of simulations.

5.4 Parameter estimation

We employ the simulations to build mock observational data for $b(k)$ and use them to estimate the PNG parameters under different model assumptions. We consider two sets of halo mass and redshift bins as described in Table 5.1. The first set is chosen to closely resemble the observations used by Xia et al. (2011). The second set covers a wider range in redshift and linear bias parameter b_1 and exemplifies datasets (for groups and clusters of galaxies) that will be available in the near future (e. g. with the eROSITA mission plus spectroscopic follow up).

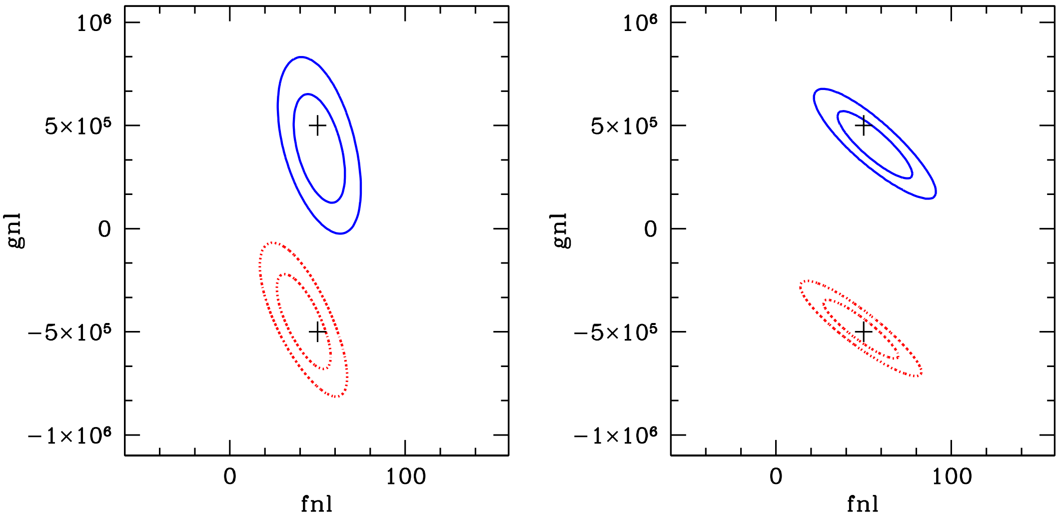

Figure 5.1: Credibility intervals (68.3 and 95.4 per cent) for $f_{\rm NL}$ and $g_{\rm NL}$. Solid: positive $g_{\rm NL}$, dotted: negative $g_{\rm NL}$. The black crosses indicate the input values for $f_{\rm NL}$ and $g_{\rm NL}$. Left panel: using the current sample, right panel: using the future sample.

5.4.1 Two-parameter model for PNG

We first consider a model (M_2) with three free parameters, $f_{\rm NL}$, $g_{\rm NL}$ and b_1, and assume Gaussian errors on $b(k)$ to build their likelihood function. We adopt a flat prior within a finite region of parameter space (considerably more extended than the likelihood function) and marginalize the posterior distribution over b_1. Figure 5.1 shows the marginal posterior probability density for $f_{\rm NL}$ and $g_{\rm NL}$ by marking the 68.3 and 95.4 per cent credibility intervals for the model parameters. The two panels refer to the different sets of mass and redshift bins defined in Table 5.1 (current data on the left, future data on the right). Different sets of curves in the same panel refer to the different fiducial models used to generate the initial conditions of the simulations. In all cases, the fiducial values for $f_{\rm NL}$ and $g_{\rm NL}$, indicated by the black crosses, lie within the 95.4 per cent credibility region region. Note that the current datasets are not expected to give tight constraints on the model parameters which are also degenerate. The credibility intervals shrink when more mass/redshift bins are used. The small rotation of the contour levels in the right panel is caused by including data from $z = 2$, which significantly improves the constraints on $g_{\rm NL}$, and only slightly weakens the constraints on $f_{\rm NL}$ (with respect to the current dataset).

5.4.2 One-parameter models for PNG

We now consider two simpler models by assuming that PNG is either purely quadratic ($g_{\rm NL} \equiv 0$, M_{1f}) or purely cubic ($f_{\rm NL} \equiv 0$, M_{1g}). All the rest is left unchanged with respect to M_2, so that M_2, M_{1f} and M_{1g} form a set of nested models.

The marginal posterior probability distribution for $f_{\rm NL}$ obtained combining the current dataset with model M_{1f} is presented in Figure 5.2. The left panel shows

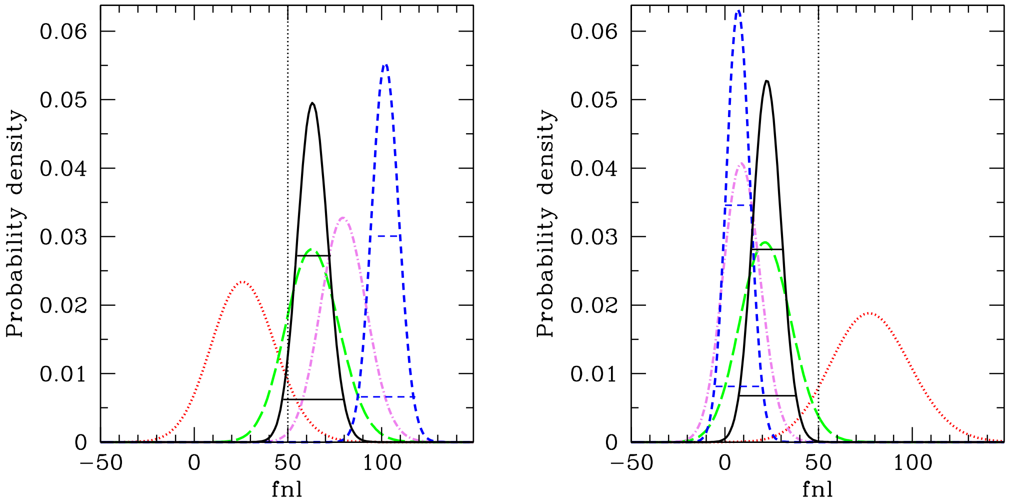

Figure 5.2: Constraints on $f_{\rm NL}$ for model M_{1f} (where $g_{\rm NL} \equiv 0$ is assumed a priori). Left: simulation with $g_{\rm NL} = 5 \times 10^5$, right: simulation with $g_{\rm NL} = -5 \times 10^5$. Constraints from the 3 redshift bins in the current sample are shown with different line styles: $z = 0.5$ dotted, $z = 1$ long-dashed, $z = 1.5$ dot-dashed. The posterior pdf obtained combining the 3 redshift bins is shown with a solid line. For comparison, we also show the combined posterior for the future dataset (short-dashed). The two horizontal lines indicate the 68.3 and 95.4 per cent credibility intervals for each of the combined posteriors. The vertical dotted line shows the input value $f_{\rm NL} = 50$.

the result for the simulation with positive $g_{\rm NL}$ and the right panel for negative $g_{\rm NL}$. Different line styles indicate the individual redshifts bins: dotted, long-dashed, dot-dashed for $z = \{0.5, 1, 1.5\}$, while the solid curve shows the combined probability for all redshifts. For the simulation with $g_{\rm NL} = 5 \times 10^5$, the resulting 95.4 per cent credibility region (lower horizontal solid line) marginally includes the input value $f_{\rm NL} = 50$, indicated by the vertical dotted line, but the maximum is shifted towards higher values. When $f_{\rm NL}$ and $g_{\rm NL}$ have opposite sign, the maximum a posteriori value is biased low, and the input value is excluded at high significance. In both cases, there is an apparent redshift dependence of $f_{\rm NL}$, and this trend depends on the sign of $g_{\rm NL}$. Thus, if a redshift dependence was ever detected from observations, it may just indicate that the wrong model for PNG is assumed. We note that the results in Xia et al. (2011) actually hint at a redshift dependence of $f_{\rm NL}$ although the error bars are quite large. If taken at face value, the ordering of the estimates would suggest that $g_{\rm NL} < 0$. Additionally, Figure 5.2 includes the combined probability for the 5 mass and redshift bins of the future dataset for comparison (short-dashed). The shift away from the input value is qualitatively the same, but *larger* for the future dataset. This can be understood by comparing the degeneracy between $f_{\rm NL}$ and $g_{\rm NL}$ in Figure 5.1.

The marginal posterior probability distribution for $g_{\rm NL}$ under model M_{1g} is shown in Figure 5.3. As in the previous Figure, the two panels show the results for the simulations with $g_{\rm NL} = \pm 5 \times 10^5$. In both cases, the maximum a posteriori estimate is biased high, and the input value lies well outside the 95.4 per cent credibility regions. The estimates from the different redshifts are almost

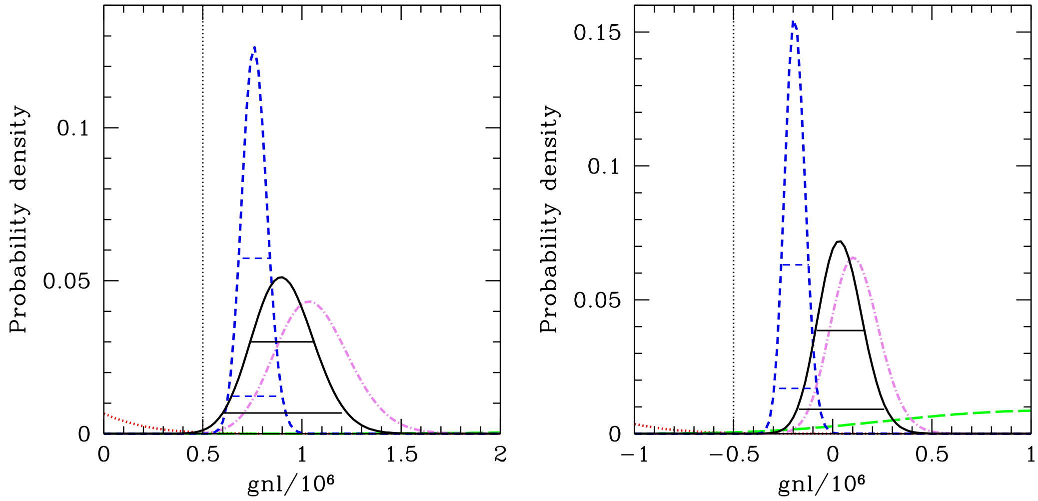

Figure 5.3: Constraints on g_{NL} for model M_{1g} (where $f_{\mathrm{NL}} \equiv 0$ is assumed a priori). Left panel: simulation with $g_{\mathrm{NL}} = 5 \times 10^5$, right panel: simulation with $g_{\mathrm{NL}} = -5 \times 10^5$. Line styles are the same as in 2. The vertical dotted line shows the input value $g_{\mathrm{NL}} = \pm 5 \times 10^5$. Note that the long-dashed curve ($z = 1$) seems to be missing in the left panel because the corresponding probability density is too low.

incompatible with each other, and there is no apparent trend with z as seen in model M_{1f}. For the case of negative g_{NL}, the estimation procedure favours a Gaussian model although both input PNG parameters are substantially different from 0. The combined probability of the future dataset is shown by the short-dashed curve. Here, the shift away from the input value for g_{NL} is *smaller*, and the constraints are tighter than for the current dataset, in accordance with the rotation of the contours in Figure 5.1.

5.4.3 Model selection

Our results show that the choice of the underlying model for PNG has a significant effect on the estimates of f_{NL} and g_{NL}. Only model M_2 returned estimates which were in agreement with the fiducial values, while setting a parameter to zero a priori (as often done in the literature) produced biased results. Here we investigate if the data themselves can be used to establish which parameterisation should be preferred.

Given a set of data D, the relative evidence in favour of a model with respect to a second one can be quantified in terms of the Bayes factor (e. g. Kass 1993, Trotta 2008)

$$B_{ab} = \frac{P(D|M_a)}{P(D|M_b)} = \frac{\int \mathcal{L}_a(D|M_a, \theta_a)\, \pi_a(\theta_a)\, \mathrm{d}\theta_{\mathrm{a}}}{\int \mathcal{L}_b(D|M_b, \theta_b)\, \pi_b(\theta_b)\, \mathrm{d}\theta_{\mathrm{b}}}. \tag{5.9}$$

Here $\mathcal{L}_i(D|M_i, \theta_i)$ gives the probability of getting the data D given the model M_i with parameters θ_i (i. e. $\mathcal{L}_i$ is proportional to the likelihood function for the model parameters), $\pi_i(\theta_i)$ is the prior probability density for the model parameters and $P(D|M_i)$ denotes the Bayesian evidence for model M_i (i. e. the probability of getting the data under model M_i after marginalizing over the values of the model parameters).

We find that M_2 is strongly favoured over M_{1f} and M_{1g}, with a Bayes factor > 30, except for one case[3]. Based on the Jeffreys scale (Jeffreys, 1961), this provides “very strong” evidence in favour of M_2. We thus recommend the use of Bayes factors in future determinations of PNG from observational data based on galaxy 2-point statistics.

5.5 Conclusion

Using two sets of N-body simulations starting from non-Gaussian initial conditions of the local type with $\{f_{\rm NL}, g_{\rm NL}\} = \{50, \pm 5 \times 10^5\}$, we have investigated the bias of dark-matter halos on large scales. Motivated by the widespread practice of considering only one-parameter models for PNG to fit observational data, we have studied the effect of model assumptions on the determination of model parameters based on 2-point statistics of the galaxy distribution. We have considered two sets of mass and redshift bins: the first one is chosen to emulate the current observational samples, while the second approximates data from future cluster surveys. Our main conclusions can be summarized as follows:

- Fitting the mock data with the input (two-parameter) PNG model gives unbiased results although estimates for $f_{\rm NL}$ and $g_{\rm NL}$ turn out to be degenerate. This degeneracy can be at least partially broken using higher-order statistics of galaxy clustering, like the bispectrum.

- If we assume a purely quadratic model ($g_{\rm NL} = 0$), the maximum a posteriori estimate for $f_{\rm NL}$ is biased high (low) for the simulations with $g_{\rm NL} = 5 \times 10^5$ ($g_{\rm NL} = -5 \times 10^5$), respectively. The input value of $f_{\rm NL} = 50$ is almost always excluded at very high confidence. If we consider each redshift bin separately, a spurious systematic shift of the $f_{\rm NL}$-estimate with z appears, which also depends on the sign of $g_{\rm NL}$. If seen in observations, this effect could actually indicate the necessity for a more complex model.

- If we assume a purely cubic model ($f_{\rm NL} = 0$), the maximum a posteriori estimates for $g_{\rm NL}$ are artificially shifted to larger values, which differ from the input value much more than the broadness of the posterior distribution. When the fiducial PNG parameters have opposite sign, the two contributions to the scale-dependent bias can cancel out, leading to a false null detection.

- Our analysis could be extended by considering even higher-order terms in the expansion of Equation (5.1). However, this phenomenological approach adds degrees of freedom which are not necessarily independent because they may be related to each other through the underlying physics. Alternatively, one could test specific inflationary models and put direct constraints on their properties (e. g. coupling constants) that determine the PNG parameters.

---

[3] Using the current sample, M_{1f} is only disfavoured by a factor of 9 (3) for the simulation with positive (negative) $g_{\rm NL}$.

- Given a dataset, statistical model selection techniques are capable to identify the optimal PNG parameterization and thus avoid estimation biases due to incorrect simplifying assumptions. The Bayes factor very strongly favours the two-parameter model over a purely quadratic or purely cubic model for our simulations. It is thus important to apply this model selection to future constraints on PNG from observational data.

Note on recent developments The recent first year data release of the *Planck* mission led to significant improvement in the accuracy of CMB constraints on the parameter $f_{\rm NL}$. Previous constraints from WMAP give for the local model $f_{\rm NL} = 37 \pm 20$, for equilateral-type NG $f_{\rm NL} = 51 \pm 136$, while for the orthogonal shape $f_{\rm NL} = -245 \pm 100$ (68% confidence level, Bennett et al., 2012). Due to its higher precision, *Planck* finds $f_{\rm NL}^{\rm loc} = 2.7 \pm 5.8$, $f_{\rm NL}^{\rm eq} = -42 \pm 75$ and $f_{\rm NL}^{\rm orth} = -25 \pm 39$ (68% confidence level, Planck Collaboration et al., 2013b). New constraints on $g_{\rm NL}$ from the CMB have not yet been published.

Giannantonio et al. (2014) used a combination of large-scale structure probes to measure the PNG parameters, while also considering the degeneracy for the scale-dependent bias. Their constraints are compatible with the null-detection from the CMB for both parameters with $|f_{\rm NL}| \lesssim 30$ and $g_{\rm NL} \lesssim 10^5$ at 95% confidence level, and the use of Bayesian model selection criteria showed no evidence for non-zero PNG. Nevertheless, the error bars on these parameters are still large compared to the CMB, so a definitive conclusion from LSS data will require at least the next generation of galaxy and cluster surveys (e. g. Pillepich et al., 2012).

6 Near-field Cosmology: How to find a Unicorn in a Haystack

6.1 Introduction

The previous chapters have focused on cosmological structure formation on large scales, using galaxy groups or clusters at intermediate and high redshifts to better understand structural evolution. However, observations of the local Universe – the region surrounding the Milky Way galaxy – can also be used to place important constraints on cosmological models. This is known as near-field cosmology, a recently coined term (Freeman & Bland-Hawthorn, 2002) with roots dating back to the 1960s (e. g. Eggen et al., 1962).

The Local Group consists of two massive galaxies, the Milky Way (MW) and Andromeda, the slightly smaller spiral galaxy M33, as well as a large number of "dwarf" galaxies orbiting as satellites. The halos of dwarf galaxies are the first objects to collapse that are capable of trapping baryons to fuel star formation, and they are therefore considered to be the basic building blocks of all subsequent galaxy formation. Their abundance at high redshift makes them suitable candidates for the sources of reionisation (Stark et al., 2007), and their low metal content allows a comparison of the primordial helium abundance to predictions of Big Bang Nucleosynthesis (Izotov et al., 2007). Understanding the origin of these objects, as well as how they impact the formation of more massive systems, is therefore of fundamental importance to unraveling the complex physics of galaxy formation.

Before the first data release of the Sloan Digital Sky Survey (SDSS, York et al. 2000), there were 11 known dwarf satellites of the Milky Way, now referred to as the "classical" dwarfs. Recent dedicated surveys of the local group, such as the Pan-Andromeda Archaeological Survey (McConnachie et al., 2009) and HST/ANGST[1] (Dalcanton et al., 2009), as well as wide field surveys, such as the Two Micron All-Sky Survey (Skrutskie et al., 2006) and further data releases of the SDSS have more than doubled the number of known Galactic dwarf galaxies, with many more likely to lie below current detection limits. Dwarf galaxies are characterized by low luminosities, $M_V > -17$ (Tammann, 1994), and small characteristic sizes, as inferred from their surface brightness profiles. There is, however, no obvious distinction between the physical properties of normal and dwarf galaxies; rather there is a continuous distribution of sizes and luminosities (Graham & Guzmán,

---

[1] Hubble Space Telesc. Advanced camera for surveys Nearby Galaxy Survey Treasury Program

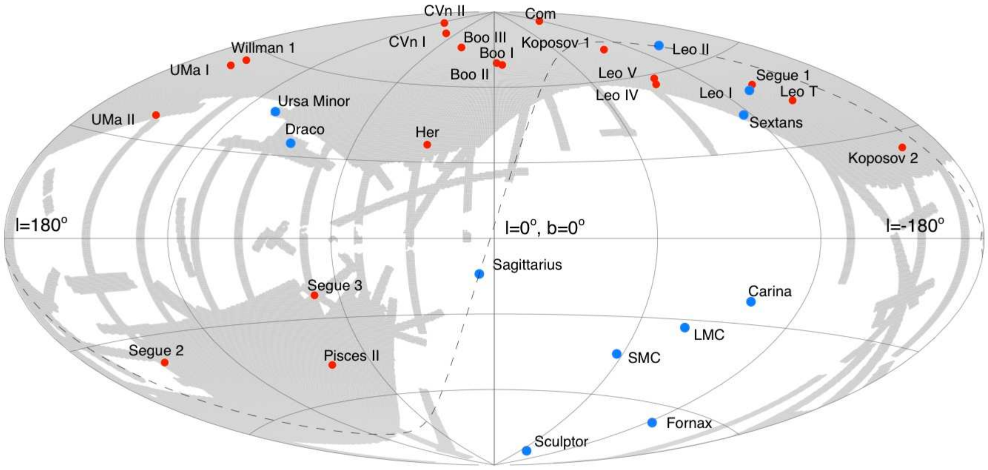

Figure 6.1: The locations of the Milky Way's dwarf satellite galaxies shown in Galactic coordinates (the Galactic center is located at $l = 0°$, $b = 0°$ and the Galactic plane lies along $b = 0°$). Blue circles show the classical dwarf galaxies; red circles show the ultra-faint satellites identified in the SDSS (including 3 globular clusters: Koposov 1 & 2 and Segue 3). The grey shaded region shows the SDSS footprint (data-release 8), and the dashed line marks the orbit of the Sagittarius dwarf galaxy as predicted by Niederste-Ostholt et al. (2010). [Figure taken from Belokurov (2013).]

2003, and references therein) extending down to mass scales normally associated with globular star clusters. At these low masses, dwarf galaxies differ from star clusters due to the presence of dark matter (DM), which is absent in the latter.

As with normal galaxies, dwarfs can also be subdivided into different morphological classes ranging from gas poor dwarf spheroidal and ellipticals, to gas rich dwarf irregulars (e. g. Kormendy et al., 2009). As for normal spirals and ellipticals, dwarf galaxies also follow a density-morphology relation: dwarf spheroidals preferentially live in the high-density regions surrounding more massive galaxies, whereas irregulars are predominantly found in low-density peripheral regions (e. g. Grebel, 2001). This suggests an environmental dependence to dwarf galaxy morphology which can be used to decipher the main mechanisms responsible for their evolution. Figure 6.1 shows the spatial distribution of the currently known Milky Way dwarf galaxies (taken from Belokurov 2013), split into the classical dwarfs (blue circles) and the more recently detected ultra faint dwarfs (red circles). The grey area shows the footprint of the SDSS (data-release 8), and the dashed line marks the orbit of the Sagittarius dwarf galaxy as proposed by Niederste-Ostholt et al. (2010) (see below).

Within hierarchical structure formation models, low mass dark matter halos are, on average, the first to collapse, with more massive systems forming through the merger and accretion of these smaller subunits. One consequence of this process is the presence of "substructure" within dark matter halos: these are the surviving self-bound cores of previous generations of accreted systems, and are thought to be the primary sites of dwarf galaxies within the Local Group. As a result, understanding both the global properties of the Local Group's dwarf galaxy system, as well as their internal properties, is essential for reconstructing exactly how the Milky Way and Andromeda galaxies formed.

One puzzle, for example, is the spatial distribution of dwarf galaxies around the Milky Way and Andromeda: a large fraction of the Milky Way's dwarf galaxies lie in a plane that is roughly perpendicular to the disc of the Milky Way (Lynden-Bell, 1976; Pawlowski et al., 2012), and recent observations have hinted at a similar ring-like structure around Andromeda (Ibata et al., 2013; Conn et al., 2013). Some authors consider this a serious challenge to ΛCDM (Kroupa et al., 2005; Pawlowski et al., 2012), since the observed planar structure is difficult to reconcile with the roughly isotropic distribution of dark matter substructures seen in numerical simulations of Milky Way-like dark matter halos. However, others have pointed out that preferential accretion of massive satellites along large-scale filaments can naturally explain the observed distribution within the context of the ΛCDM model (Libeskind et al., 2011; Lovell et al., 2011). Alternatively, the majority of the MW's dwarf galaxies may have formed directly from the tidal debris stripped from a previous merger. One of the most well-known examples is the Magellanic stream (Mathewson et al., 1974), created by the interaction of the Milky Way and the Magellanic clouds. The stripped material includes stars, but also atomic and molecular gas and dust, which can fuel the formation of tidal dwarf galaxies, which should be devoid of dark matter (e. g. Mirabel et al. 1992; Barnes & Hernquist 1992). Since tidal dwarf galaxies would be aligned with the tidal arms in which they form, they naturally explain the disk-like distribution of the Galactic satellites as originating from a previous interaction between the MW and a now-disrupted companion (Metz et al., 2008). This scenario seems unlikely, however, given the large mass-to-light ratios inferred for most Local Group dwarf galaxies, implying that they contain a significant amount of dark matter.

Similar to tidal streams, stellar streams are remnants of past or ongoing mergers, often found in the inner, dense regions of galaxies where tidal forces are strongest. They are typically devoid of gas, which has either been consumed in a star burst at infall or stripped by ram pressure. Stellar streams closely trace the orbit of their progenitor and can be stable for several gigayears (Springel & White, 1999); their structure and location within the host can be used to probe the host's potential, resulting in robust mass profile constraints which can then be compared with theoretical expectations. Stellar streams therefore offer powerful probes of galactic structure and evolution and can be used to constrain models of structure formation on small scales. Perhaps the most prominent stellar stream within the Milky Way is the remnant of the Sagittarius dwarf galaxy (Mateo et al., 1996), although numerous others have been detected in both the Milky Way and Andromeda, as well as in other nearby galaxies (e. g. Ibata et al., 2001a,b; Martínez-Delgado et al., 2008). The Sagittarius stream is extended and lies well outside of the Galactic plane, allowing it to be studied in detail. Niederste-Ostholt et al. (2010), for example, modeled Sagittarius' orbit (shown as dashed line in Figure 6.1) and estimated that the progenitor has lost up to 70% of its initial stellar mass, which is now mixed with the intrinsic stellar populations of the Milky Way's halo and disk.

Another peculiar stellar structure within the Milky Way is the Monoceros Overdensity (MO, Greek for “Unicorn”), which lays at a distance of $\sim$ 18 kpc from the Galactic center (Newberg et al., 2002) and has an estimated stellar mass of $2 \times 10^7 \leq M_* \leq 5 \times 10^8\,\mathrm{M}_\odot$ (Yanny et al., 2003). Similar to the Sagittarius stream, the Monoceros Overdensity (MO) is rather broad and extends over a large range in Galactic longitude ($\sim 180°$), but instead of lying perpendicular to the Galactic disk like Sagittarius, the MO forms a ring-like structure that roughly aligns with the Galactic plane (Ibata et al., 2003; Crane et al., 2003; Rocha-Pinto et al., 2003; Yanny et al., 2003; Martin et al., 2005, 2006; Conn et al., 2005a,b, 2007, 2008; Jurić et al., 2008; Ivezić et al., 2008a; de Jong et al., 2010; Sollima et al., 2011; Conn et al., 2012, and references therein). A possible progenitor for the MO (if indeed it is the remnant of a tidally disrupted MW satellite) is the Canis Major dwarf galaxy (CMa, Martin et al., 2004), which lies at $(l, b) = (-120°, -9°)$ and a distance of 7 kpc from the Sun. Alternatively, the alignment of the MO with the disk could be explained if the MO and CMa are only *apparent* stellar overdensities resulting from, for example, the excitation of the Galactic disk by a past merger event, or a strong warp in the disk (Moitinho et al., 2006; Momany et al., 2006; Kazantzidis et al., 2008). However, this is disfavored by measurements of the stellar metallicities of MO stars, which are considerably lower than those of disk stars with similar Galactic coordinates (e. g. Ivezić et al., 2008a).

The left panel of Figure 6.2 shows an ensemble of observations of MO and CMa, collected by Conn et al. (2007) (referred to as “this paper” in the plot legend). Each symbol marks a detection of stars that can be associated with those structures by their position in a color-magnitude diagram (CMD). Filled symbols lie above the galactic plane and open symbols lie below. The solid line shows an additional stellar stream observed by Grillmair (2006).[2] The $\sim$1 kpc error on the position measurements is indicated by the bar in the bottom right corner. The right panel shows an edge-on view of the stellar overdensity as a function of distance R to the galactic center and height Z above the Galactic plane determined using SDSS data (Jurić et al., 2008). A smooth model for the Milky Way’s stellar halo has been subtracted from the observed stellar distribution; the MO is the residual at $Z \sim 0$ and $R \sim 20\,\mathrm{kpc}$.

Considerable observational and theoretical effort has been devoted to understanding the origin of the MO, and it is now generally accepted that its kinematics are consistent with a stellar stream stripped from a merging dwarf galaxy. Under this assumption, the multitude of available observations is sufficient to allow a reconstruction of the progenitor’s orbit. Peñarrubia et al. (2005) performed a large number of simulations of satellite stripping in MW-like potential in an attempt to constrain the properties of the MO’s progenitor. They were able to numerically reproduce the geometry and kinematics of the stream by assuming that the infalling

---

[2] Curiously, this author concluded that the detected stream is not physically associated with the MO, but this is not mentioned in Conn et al. (2007). A more recent study by Carlin et al. (2010) is also unable to determine unambiguously whether this stream is part of the MO.

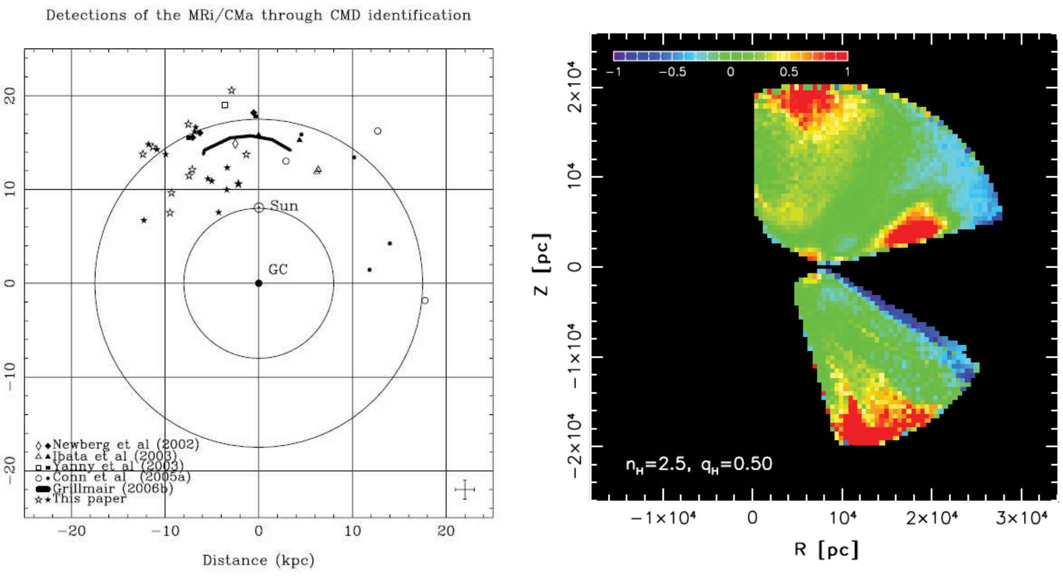

Figure 6.2: *Left panel*: A collection of observations of the Monoceros overdensity and the Canis Major dwarf galaxy. [Figure taken from Conn et al. (2007), see text for a detailed description.] *Right panel*: Edge-on view of the residual stellar overdensity (color scale) in SDSS data after subtracting a smooth model of the Milky Way's halo. The Monoceros overdensity is shown as a residual excess of stars close to the Galactic plane ($Z = 0$ and $R \sim 20\,\mathrm{kpc}$). The parameters n_H and q_H describe the halo model. [Figure taken from Jurić et al. (2008).]

satellite had a mass of $\sim 6\times10^8\,\mathrm{M}_\odot$ and followed a prograde, coplanar and nearly circular orbit before being disrupted. However, due to uncertainties in some of the model's parameters (mainly the axis ratios of the Milky Way's halo) both the mass and position of the progenitor were poorly constrained. Similar quality fits, for example, were obtained for progenitor masses in the range $\sim (3-9)\times10^8\,\mathrm{M}_\odot$, provided its current position is $\sim 15\,\mathrm{kpc}$ from the center of the galaxy, roughly twice as distant as the CMa overdensity.

Recent proper motion measurements of member stars have been used to constrain the orbit of CMa (Dinescu et al., 2005). The orbit has an apocentric distance of $14\pm0.2\,\mathrm{kpc}$, and pericenter of $10\pm0.9\,\mathrm{kpc}$, suggesting that its orbital circularity is roughly in line with what is expected for the MO progenitor. In fact, the best-fit N-body orbits obtained by Peñarrubia et al. (2005) are consistent with the proper motion-based orbits derived by Dinescu et al. (2005), supporting the hypothesis that the MO is a remnant of CMa. Whether or not this can be confirmed, it remains difficult to explain how a low-mass satellite came to be on such a circular orbit: the results of simulations of Milky Way-mass dark matter halos indicate that most satellites are accreted on highly eccentric orbits. The orbital decay rate due to dynamical friction is proportional to the mass of the satellite and the density of the host at its location (Chandrasekhar, 1943; Binney & Tremaine, 2008). Dynamical friction timescales for low-mass satellites are therefore very long, and this process alone is unlikely to be sufficient to fully circularize a the orbit of a typical satellite given the finite age of the Milky Way.

The physical mechanism responsible for circularizing progenitor's orbit is therefore unclear, though a number of possibilities exist. One of them, as discussed by Peñarrubia et al. (2005), is that the mass of the MO's progenitor is substantially larger than predicted by their best-fit model. A $1.8 \times 10^9\,\mathrm{M}_\odot$ progenitor, for example, will be dragged by dynamical friction to a typical distance of $\sim 7.2\,\mathrm{kpc}$, which agrees with the observed distance to CMa. Alternatively, the progenitor of the MO may have been "born" relatively close to the Galactic center and on a circular orbit. This would allow the system to survive to the present day by evading the strong tidal effects of the Galactic center. Or, as proposed by Michel-Dansac et al. (2011), three-body interactions between the Milky Way and two (or more) infalling satellites could have led to rapid changes in the orbital energies and angular momenta of the infalling systems and, under the right circumstances, result in the circularization of one of their orbits.

In this chapter, we use high resolution N-body simulations of the formation of MW-like dark matter halos to study which of these two possibilities best explains the origin of the MOs progenitor. To do so, we trace the full dynamical history of all subhalos that were ever accreted by their host galaxy, as well as their association with other dark matter halos before being accreted onto the main halo. In doing so, we are able to identify subhalos which are *today* on circular orbits and identify the physical mechanisms responsible for their circularization.

6.2 Numerical methods

In this section we provide a brief description of our main numerical and analysis methods. This includes a description of the *Aquarius* simulations (Springel et al., 2008), our halo finding and orbit tracking algorithms, as well as other important analysis techniques. We discuss a variety of possible subhalo orbits and define selection criteria to create idiosyncratic subsamples based on their dynamical properties.

6.2.1 The *Aquarius* simulations

The *Aquarius* project comprises a suite of six dark matter-only simulations of Milky Way-like halos. These were first identified in a larger volume "parent" simulation (the Millennium-II simulation, Boylan-Kolchin et al. 2009), have virial masses on the order of $M_{200} \sim 10^{12}\,h^{-1}\,\mathrm{M}_\odot$ and have no neighbors within $1\,h^{-1}\,\mathrm{Mpc}$ exceeding half of their present-day mass. Each halo was then resimulated at higher resolution, with increasing numbers of particles in order to identify limitations due to numerical resolution (Navarro et al., 2010). For any given run, a minimum of 100 snapshots were saved between $z = 20$ and 0, which is more than sufficient for accurately characterizing the full dynamical evolution of the dark matter halos and subhalos.

All simulations adopted the same set of cosmological parameters: $\Omega_\mathrm{m} = 0.25$, $\Omega_\Lambda = 0.75$, $h = 0.73$, $n_\mathrm{s} = 1$, and $\sigma_8 = 0.9$. (For a definition of these parameters see Chapter 1). Although these values are not strictly consistent with the most

Table 6.1: Various properties of the *Aquarius* simulations used in this work (see Springel et al. (2008) for a detailed description). Halo names consist of a letter (A to F) for each halo, and a number encoding the resolution level (2-4, with 2 being the highest resolution used in this work). $m_{\rm part}$ is the dark matter particle mass in the high-resolution region, and $M_{\rm sub}^{\rm min} = 20\, m_{\rm part}$ is the minimum subhalo mass adopted for our group and halo finders. The properties of the host halo at $z = 0$ are M_{200}, r_{200} and v_{200}, which denote the mass, radius and circular velocity within a spherical region with a mean density of 200 times the critical density for closure. The last column indicates the method of subhalo tracking: using merger trees (MT), or our own method when particle data (PD) is available; both described in section 6.2.3.

Halo	$m_{\rm part}\,[h^{-1}\,{\rm M}_\odot]$	$M_{\rm sub}^{\rm min}\,[h^{-1}\,{\rm M}_\odot]$	$M_{200}\,[h^{-1}\,{\rm M}_\odot]$	$r_{200}\,[h^{-1}\,{\rm kpc}]$	$v_{200}\,[{\rm km\,s^{-1}}]$	Comments
Aq-A-2	1.000×10^4	2.000×10^5	1.345×10^{12}	179.41	179.55	MT
Aq-A-3	3.585×10^4	7.170×10^5	1.341×10^{12}	179.31	179.36	MT
Aq-A-4	2.868×10^5	5.736×10^6	1.342×10^{12}	179.36	179.37	PD
Aq-B-2	4.706×10^3	9.412×10^4	5.982×10^{11}	137.02	137.06	MT
Aq-B-4	1.637×10^5	3.274×10^6	6.092×10^{11}	137.86	137.86	PD
Aq-C-2	1.021×10^4	2.042×10^5	1.295×10^{12}	177.26	177.31	MT
Aq-D-2	1.020×10^4	2.040×10^5	1.295×10^{12}	177.28	177.33	MT
Aq-D-4	1.954×10^5	3.908×10^6	1.307×10^{12}	177.82	177.83	PD
Aq-E-2	7.002×10^3	1.400×10^5	8.652×10^{11}	154.96	155.01	MT
Aq-F-2	4.946×10^3	9.892×10^4	8.282×10^{11}	152.72	152.76	MT
Aq-F-4	1.335×10^5	2.670×10^6	8.024×10^{11}	151.11	151.12	PD

recent estimates from *Planck* and WMAP (see section 2.5), the differences are expected to mainly affect halo collapse times and are largely irrelevant for the evolutionary and dynamical properties of halo and subhalo orbits that we discuss here. The analysis in this chapter uses both high resolution (level-2) and lower resolution (level-4) versions of each *Aquarius* halo. In Table 6.1 we provide a naming convention for the simulations and a list of several important properties of each *Aquarius* halo.

The high-resolution level-2 runs contain a large number of low-mass structures; this is important for statistical analysis, since circular subhalo orbits are expected to be rare in cosmological halos. The level-4 data is used to study the effects of environment, mass resolution, and to visualize the unbound filamentary structure along which the majority of infalling halos are accreted (Libeskind et al., 2011; Lovell et al., 2011). This is computationally less cumbersome than using the level-2 particle data directly, but numerically accurate given that the relevant properties of the halos converge at the percent level between the two resolutions (Springel et al., 2008; Navarro et al., 2010).

6.2.2 Halo and subhalo catalogs

Dark matter halos are identified in each simulation output using a friends-of-friends (FOF) halo finder with a linking-length of $0.2\,b$, where b is the mean inter-particle distance. Each FOF halo is then dissected into self-bound substructure halos using the SUBFIND algorithm (Springel et al., 2001, summarized in section 3.2.3) , which separates the main dark matter halo from its smaller system of satellite halos. For clarity, we refer to the largest "substructure" halo within any FOF halo as the *main* or host halo; any sub-dominant halos we refer to as substructure halos, or simply as subhalos, for short. SUBFIND ensures that each halo or subhalo is gravitationally self-bound, which rids the halo catalogs of loose groupings of particles that may (temporarily) pollute the FOF lists. In addition, it records the position, velocity, mass and various other useful characteristics of halos that greatly simplifies later analysis. For our runs we have adopted a minimum of 20 particles per halo or subhalo.

6.2.3 Orbit tracking

In order to study the orbits of subhalos in our simulations we require a method of tracking them across different SUBFIND outputs. We consider *all associated* subhalos, defined as those that have been within the time-dependent virial radius of the main halo's progenitor at some point in the past (see also Ludlow et al., 2009). We therefore consider both bound and unbound objects, as well as those that have been disrupted by tidal interactions with the host halo since $z = 2$.

For the level-2 simulations our orbit tracking was based on halo merger trees, which were constructed in the following way (Boylan-Kolchin et al., 2009). Starting at the initial redshift, the particles within each halo are identified in two subsequent time steps. If all particles within a particular halo are found within another halo at subsequent time-step then this is the descendant of that halo. If, instead, they are spread over several halos, then the particles are weighted by their binding energy, and the halo that contains most of the highly bound particles from the previous snapshot defines the descendant. Most of the time this will be the halo which contains most particles, but the weighting allows for a distinction in the case where several halos contain roughly the same number of tracked particles. With this methodology, a halo may have multiple progenitors but only one descendant, reflecting hierarchical structure formation. By linking halos across cosmic time in this way we can characterize the full dynamical history of a given dark matter halo, including the evolution of its structural properties, as well as its its dynamical association with any other structures. Note that it is also possible for a halo to “skip” one time step if it, for example, enters a high density environment and the SUBFIND algorithm fails to detect it, or if it temporarily falls below the 20 particle limit. We have carefully identified and excised such instances in our runs so that they do not impact our results and interpretation.

For the level-4 simulations we follow a different approach and generate our own subhalo tracking algorithm that uses both the SUBFIND outputs as well as the dark matter particle data. The method is similar to the merger tree method described above, although it differs in several details. The biggest difference is that we first identify halos at low redshifts and trace their evolution backward in time, as opposed to the forward-constructed merger trees. Nevertheless, we consider all subhalos that have been within the virial radius of the main halo, and were not fully disrupted prior to $z \sim 2$. To identify the progenitor of a particular subhalo we first identify all subhalos in the previous snapshot that contain some of its particles. If any of these halos contain more than 70% of the subhalo’s total mass then we define this as the *progenitor* subhalo. Otherwise, we weight particles by their binding energy and identify which of the progenitor candidates contributes most of its mass to the most-bound parts of the descendants mass distribution. Matching particles across several snapshots is computationally time consuming and so it is only employed in cases where a large fraction of the subhalo’s mass is distributed among many possible progenitors. In addition, we enforce that the mass of a subhalo can not increase by more than a factor of 3 between two snapshots. This prevents mismatching of subhalos with larger objects with which they may have had a close encounter, and also accounts for subhalos that may have been temporarily lost while orbiting within the dense background of the main halo. In this case we simply skip one (or more) snapshot and repeat the process again at the next (earlier) time. This allows the subhalo to re-emerge from its interaction and, in most instances, we successfully recover its orbit even if *several* outputs have been skipped (an example orbit with “holes” near its pericentric passage is shown in Figure 6.3 using solid red lines).

The majority of the results presented in section 6.3 use only the orbits calculated from the merger trees, so the second method just described serves mainly as a consistency check. As we will show below, both methods lead to very similar statistical distributions of subhalo properties. This means that we have created a reliable tracking method that can be used in future projects for which merger trees are unavailable.

Orbits tracked in this way can be used to identify subhalos occupying roughly circular orbits, and to study their dynamical and structural evolution. Figure 6.3 shows a selection of approximately circular subhalo orbits (top panel) and their mass evolution (bottom panel) as a function of look-back time in the Aq-F-4 run. The connected black circles show the virial radius r_{200} (top) and mass M_{200} (bottom) of the host halo. These orbits exhibit considerable complexity and subhalos that are on circular orbits *today* may have been accreted on highly eccentric orbits in the recent past. Consider, for example, the three subhalos that infall at at look-back time of $t_{\rm lb} = 4.5$ Gyr (shown as black dotted, purple dashed and green dot-dashed lines in Figure 6.3). These halos were accreted on highly radial orbits but rapidly circularized upon reaching their first pericentric passage. The "wiggles" in their obits prior to infall betray the fact that these systems are in fact part of a larger infalling group; the interaction between the merging group and the potential of the main halo resulted in rapid changes in their orbital energy and angular momentum and, in these cases, to the complete circularization of their orbits. Such 3-body interactions have previously been discussed in the context of unbound orbits (Ludlow et al., 2009), and may also result in systems gaining orbital energy (cyan triple-dot-dashed line) and being fully ejected from the system (solid blue line). In contrast, the three subhalos that fall in at early times, roughly 11 Gyrs ago (red solid, orange dashed and cyan triple-dot-dashed), have maintained roughly the same orbital eccentricity over the past 10 Gyr. These objects have avoided complete disruption because their approximately circular orbits have kept them away from the strong disruptive tides associated with the central density cusp of their host halo.

These results indicate that there may be several mechanism behind the generation of circular orbits in cosmological dark matter halos. We will return to a discussion of these possibilities in the following section, but first provide some basic characteristics of subhalo orbits in the next section.

6.2.4 Orbit classification

The halo tracking methods described above provide a list of subhalo properties as a function of time that can be used to categorize subhalo orbits. Here we provide a brief summary of some important characteristics of orbits that will be useful when interpreting the results presented in the following section.

- The infall redshift, $z_{\rm in}$, is defined as the time when a subhalo's orbit first crosses the virial radius of the host halo. Any quantity measured at infall will carry the subscript "in". Note that we will often refer to the *time since infall*, which is simple the look-back time, $t_{\rm lb}$, from $z = 0$ to $z_{\rm in}$.

- Associated subhalos initially recede with the Hubble flow, reaching a maximum distance from the center of their host before $z_{\rm in}$. We will use this

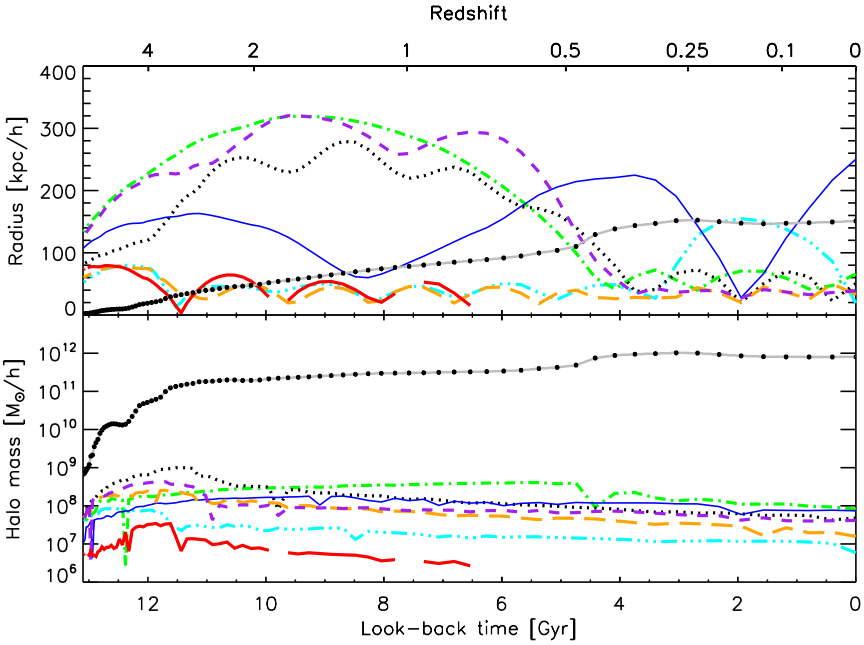

Figure 6.3: The evolution of subhalo properties in Aq-F-4 as a function of time. Shown are the radial separation from the main halo's center (*top panel*) and the bound mass as identified by the SUBFIND algorithm (*bottom panel*) as a function of look-back time from $z = 0$. The solid grey line with black points shows the evolution of the virial radius and mass of the main halo, respectively; each dot marks a simulation snapshot. See the text for a more complete description of the different subhalo orbits.

"turn-around" radius, r_{turn}, to characterize the orbital energy of a subhalo and to determine whether subsequent interactions with other halos have led to net losses or gains in orbital energy. Although this is technically the first apocenter of a given orbit, we will refer to it as the turn around radius so that apocenters are only defined for associated systems.

- Not every subhalo survives until $z = 0$. They may, for example, completely merge with another structure (not necessarily the main halo), be destroyed by interactions, or drop below the detection threshold of the SUBFIND algorithm. In this case we also record the minimum redshift z_{min} up to which a subhalo can be identified. We find that the majority of subhalos survive until the end of the simulation. However, this depends slightly resolution, with more subhalos disappearing in the low-res runs due to the decrease in their characteristic internal densities.

- Finally, the peri- and apocenter of a subhalo's orbit are defined as the minimum and maximum separation from the main halo center after infall, respectively. These extrema can be estimated directly from the trajectory of the subhalo over cosmic time, or from the Radial Energy Equation (REE) using its instantaneous orbital energy and angular momentum within the potential of the main halo (Binney & Tremaine, 2008). We will discuss both possibilities below.

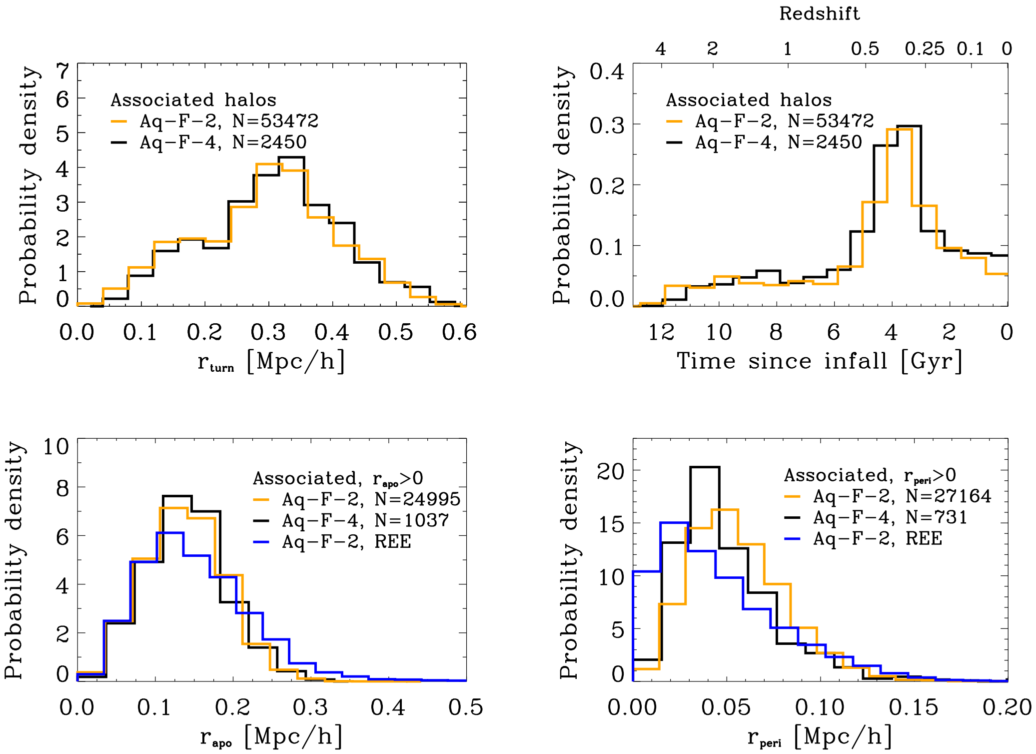

Figure 6.4: Probability distribution of subhalo properties with both tracking methods for halo Aq-F. The number of associated subhalos for which each quantity can be calculated is indicated in the legend. For example, each associated halo must have a turn-around radius and infall time, but not for all of them can we calculate the apo- and pericenter from the orbits. As expected, there are fewer halos identified in the level-4 run, but their properties are very similar. In addition, we show the distribution of apo- and pericenter as calculated from the radial energy equation (REE, blue), for the level-2 data. This shows that both tracking methods lead to the same results, but already indicates systematic differences between the two estimates for r_{apo} and r_{peri} (mostly for the latter).

The "shape" of a subhalo's orbit (whether radial or circular) can be characterized by its eccentricity, ϵ, defined

$$\epsilon = \frac{r_{\mathrm{apo}} - r_{\mathrm{peri}}}{r_{\mathrm{apo}} + r_{\mathrm{peri}}}. \tag{6.1}$$

Here r_{apo} and r_{peri} are the apocentric and pericentric distances between the object and the center of the main halo. Purely radial orbits will have $\epsilon = 1$, whereas $\epsilon = 0$ defines a perfectly circular orbit. Estimating ϵ requires knowledge of both r_{apo} and r_{peri}. The first estimate (ϵ_{rec}) comes directly from the subhalo's tracked orbit, by choosing the most recent local maximum and minimum as values of the apo- and pericenter, making sure that genuine extrema are found (so that $\dot{r} = 0$ at both r_{apo} and r_{peri}). The second estimate (ϵ_{pot}) uses the radial energy equation,

$$r^{-2} + \frac{2\left[\Phi(r) - E\right]}{|\boldsymbol{\ell}|^2} = 0, \tag{6.2}$$

at redshift 0. Here, $\Phi(r)$ is the gravitational potential of the host at radius r; $E = E_{\mathrm{kin}} + E_{\mathrm{pot}}$ and $\boldsymbol{\ell}$ are the total orbital energy and angular momentum of the satellite. The two solutions to this equation are r_{apo} and r_{peri}.

There are several reasons to prefer the subhalo orbits directly in the calculation of ϵ: 1) the radial energy equation only applies to bound orbits and so does not necessarily apply to all substructures; 2) the input potential for the radial energy equation must be calculated directly from the particle data assuming spherical symmetry, which can be problematic during mergers; 3) the direct method allows us to select halos that have completed at least one full orbit within the main halo (since we are searching for progenitors of a stellar stream which requires time to be stripped, this is a reasonable choice).

Figure 6.4 shows a comparison between both tracking methods and different ways to estimate the apo- and pericenter. In the top row, the distribution of subhalo properties (r_{turn} and t_{inf}) is shown for both resolution levels of Aq-F. The number of associated halos is given in the legend. Even though there are fewer subhalos in the level-4 data because of the lower mass resolution, the overall distribution are almost identical. In the lower panel, we show the distribution of r_{apo} and r_{peri}. The number of subhalos is lower here because these values are not defined for every associated subhalo: One drawback of the direct method is that r_{apo} is only well-defined for systems that have completed *at least* one full orbit within the body of the main halo. This is not satisfied, for example, in cases where the subhalo is on its approach to first apocenter; in these cases, and if necessary, we adopt the apocenter estimated from the radial energy equation[3]. Similarly, the r_{peri} may not be defined if the minimum radius is not a true extremum. Additionally subhalos in the low-res run are more likely to "disappear" close to their pericenter, which explains the rather low number with defined pericenters in Aq-F-4.

Circular subhalos We define circular orbits as those that satisfy the following three conditions: 1) the subhalo is inside r_{200} at z_{min}; 2) $\epsilon < 0.2$ (so that $r_{\mathrm{apo}} < 1.5\, r_{\mathrm{peri}}$), and 3) the subhalo's most recent apocenter was inside r_{200}. The first and second condition are straightforward; the third restricts the sample to subhalos whose entire orbits lie within r_{200}, and are therefore not simply "grazing" the virial boundary of the main halo.

How likely is it for a subhalo to be on a circular orbit and how does this probability differ between the six *Aquarius* halos? In Figure 6.5 we compare the eccentricity distribution for associated subhalos in each simulation. The red histogram corresponds to ϵ_{rec} and the blue hatched histogram to ϵ_{pot}. Note that we only show those halos for which both eccentricity estimates are defined (i. e. both apo- and pericenter are calculable for each method). One can see that the distribution of ϵ_{pot} is shifted slightly towards more radial orbits; for $\epsilon < 0.2$, however, the distributions are quite similar. This can be seen in the plot's legend, where the fraction of subhalos with $\epsilon < 0.2$ is provided for each estimate of ϵ. (Note that this fraction

---

[3] For systems that have completed at least one full orbit within their host, we have explicitly verified that both methods provide accurate estimates of the apo- and pericentric distances.

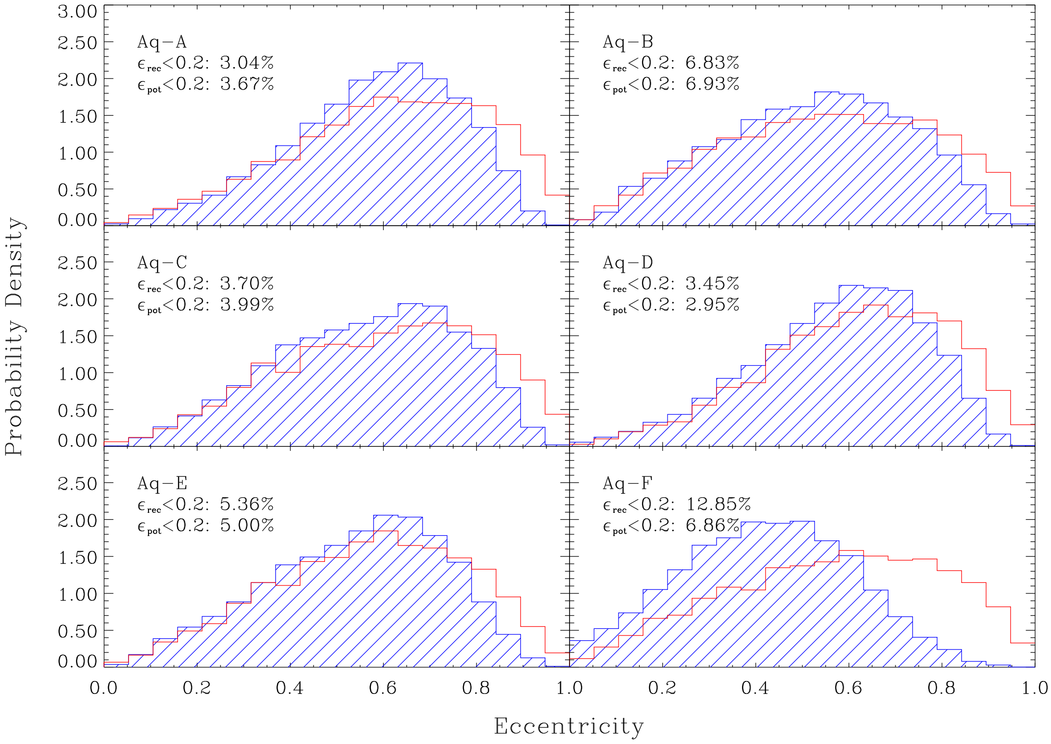

Figure 6.5: Distribution of eccentricities for all associated subhalos in the six level-2 *Aquarius* halos. Blue hatched regions corresponds to $\epsilon_{\rm rec}$, for which halo apocenters and pericenters are computed directly from their orbital tracks; red histograms show $\epsilon_{\rm pot}$, for which $r_{\rm apo}$ and $r_{\rm peri}$ are derived from the radial energy equation. In general, the distributions are quite similar. One exception is halo Aq-F, which has a major merger at low redshift that disturbs the symmetry of the halo potential and, consequently, biases estimates of $\epsilon_{\rm pot}$. (Note that perturbations from spherical symmetry in the halo potential induced by this merger affect mainly $r_{\rm peri}$; estimates of the apocentric distance are robustly estimated by both methods.) The legend gives the fraction of subhalos with $\epsilon < 0.2$ with respect to the total number of subhalos for which both estimates can be calculated, which is about half of the total number of associated subhalos.

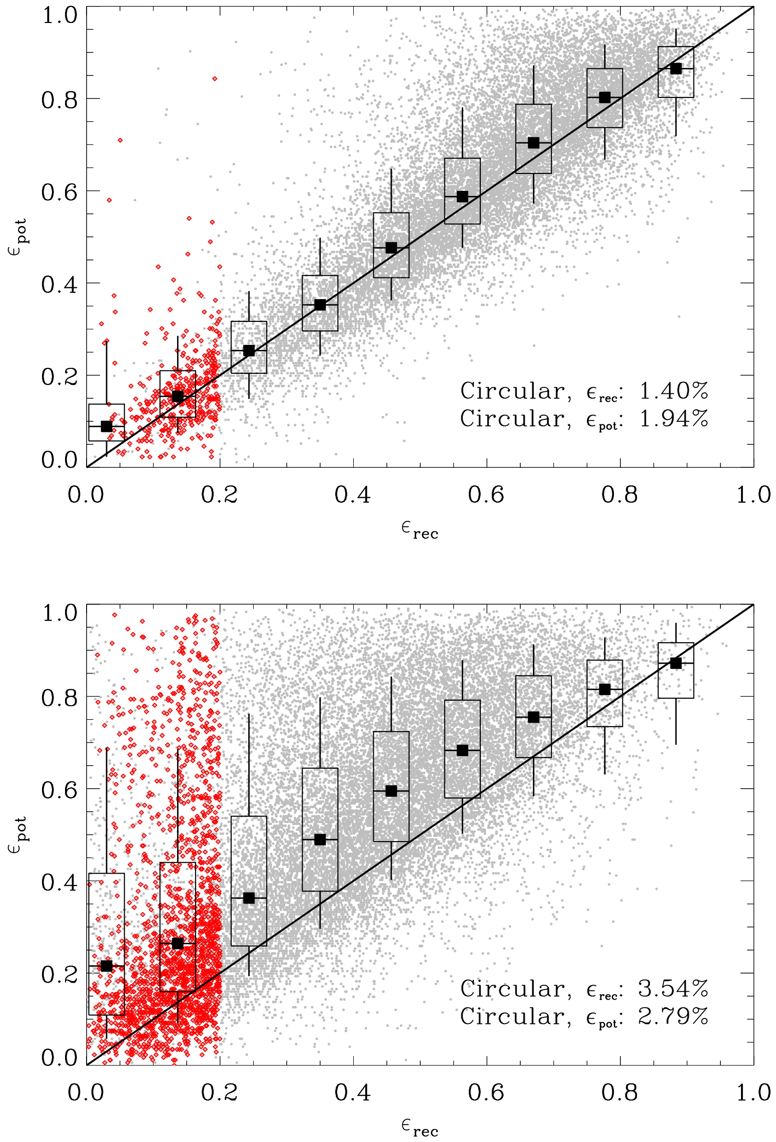

Figure 6.6: One-to-one comparison of the different definitions of eccentricity for Aq-A-2 (*left panel*) and Aq-F-2 (*right panel*). All subhalos for which both eccentricities are well defined are shown as grey circles, and those that fulfill all three criteria for circularity as red diamonds (see section 6.2.4 for details). The solid squares trace the median; boxes and whiskers indicate the 10^{th}, 25^{th}, 50^{th} and 90^{th} percentiles of the scatter. The fraction of all associated subhalos on circular subhalos is given in the bottom right corner of each panel. For relaxed halos, like Aq-A, there is a good correspondence between both estimates of the orbital eccentricity. For Aq-F-2, which has experienced a late-time merger, the correlation is somewhat weakened; we have traced this to inaccurate estimates of $r_{\rm peri}$ obtained from the radial energy equation.

is relative to the number of subhalos for which both estimates are defined, which is about 50% of all associated subhalos.)

A more direct comparison of $\epsilon_{\rm pot}$ and $\epsilon_{\rm rec}$ is shown in Figure 6.6 for halos Aq-A (left panel) and Aq-F (right panel). For a quiescent halo like Aq-A there is a tight correlation between $\epsilon_{\rm pot}$ and $\epsilon_{\rm rec}$, with only few outliers. These occur when a subhalo has not spent enough time inside the main halo to reach a stable orbit, in which case its most recent apo- and pericenter may not coincide with the estimates from Equation (6.2). The agreement between the two ϵ estimates for Aq-F is considerably worse. We have traced the origin of this discrepancy to a late time merger (occurring at $z \approx 0.45$), in which roughly half of the halo's present day mass is accreted over a very short time interval (see Ludlow et al. 2012). This accretion event perturbs the potential of the main halo leading to strong departures from spherical symmetry and, therefore, slightly biases estimates of $r_{\rm peri}$ obtained from the radial energy equation. Note that we have explicitly verified that the discrepancy between $\epsilon_{\rm pot}$ and $\epsilon_{\rm rec}$ is largely a result of differences in pericenter estimates, with apocenters largely unaffected. We will discuss Aq-F in more detail in the following sections.

The fraction of associated subhalos that are on circular orbits is provided in the bottom right corner of each panel. This value varies between ∼1% and ∼3.6% for the six *Aquarius* halos. Note that this fraction gives the number of circular halos with respect to the total number of associated halos, and so can not be directly compared to the fraction of subhalos with $\epsilon < 0.2$ quoted in Figure 6.5. As expected, the fraction of subhalos on circular orbits is quite low, which is why it is important to have simulations with large numbers of satellites in order to increase the sample statistics. In the following, we will use $\epsilon \equiv \epsilon_{\rm rec}$ when characterizing orbital circularity, unless explicitly stated otherwise.

Ejected subhalos Not every infalling subhalo will lose energy and angular momentum and end up on a tightly bound orbit. Some will *gain* energy from multi-body interactions and may be ejected from the system entirely. We consider a subhalo to have been ejected, when its position at $z_{\rm min}$ is $> r_{200}(z_{\rm min})$ and $> r_{\rm turn}$, such that its halo-centric distance is beyond both the host's virial radius as well as its original turn-around radius. A less strict limit would be to choose $r_{\rm apo} > r_{\rm turn}$, but this would include a large number of subhalos that are falling back towards the main halo at low redshift. Since the number of ejected orbits is large even with the strict selection criterion, we adopt $r(z_{\rm min}) > r_{\rm turn}$ because this defines the subsample of those that have *already* been ejected.

6.3 Results

6.3.1 Satellite accretion and the origin of circular and ejected subhalos

After defining samples of circular and ejected halos we can then investigate their dynamical history. Since the dynamical properties of a subhalo will be strongly affected by interactions with the main halo, the infall time will play an important role: The longer a subhalo has been orbiting within the body of the main halo the more time has been available for processes such as dynamical friction and tidal stripping to affect its orbit. In Figure 6.7 we show the radial velocity vs. the distance to the main halo for all ($z = 0$) associated subhalos in Aq-A-2. The color-coding indicates the time since infall in gigayears. There is a clear separation in $v_\mathrm{r}(r)$ space between subhalos that fall in early ($t_\mathrm{in} > 10$ Gyrs) and those that fall in at later times. Systems that are accreted early are more likely to end up at small radii, while most that fall in at later times occupy the outer regions of the host halo. that are accreted late. These are systems that are just now located at their first pericenter, and are rare because of the large characteristic velocities of subhalos near the center of the host. The radial of infall time is consistent with previous work (Rocha et al., 2012) based on the Via Lactea II simulation.

Interestingly, the most recently accreted systems have typically large halo-centric radii. In fact, roughly 45% of subhalos accreted in the past Gyr are currently beyond the virial radius of the host, indicated by the vertical line in Figure 6.7. This is actually *expected* from simple self-similar, secondary infall models (SSIM, e. g. Fillmore & Goldreich 1984, Bertschinger 1985). SSIMs envision three distinct regions surrounding a growing dark matter halo: 1) an envelope of material, beyond the current turn-around radius, in which mass shells still expand with the Hubble flow; 2) an intermediate “infall” region in which material has decoupled from the Hubble flow but have not yet experienced shell-crossing (halos in this region are indicated by grey dots in Figure 6.7) and, 3) an inner region in which mass shells have experienced repeated shell crossing and virialization has occurred.n Within SSIMs, accreted mass shells have their orbital energy gradually degraded until they reach a stable configuration. In Bertschinger’s SSIM, for example, the first apocenter (after z_turn) reaches a distance of roughly $0.9\,r_\mathrm{turn}$, and a stable orbital configuration with $r_\mathrm{apo} \sim 0.8\,r_\mathrm{turn}$. Therefore, orbits with apocenters as large as $\sim 1.6\,r_{200}$ are entirely consistent with expectations from SSIMs. Because $r_\mathrm{turn}(z)$ increases steadily with time, subhalos with higher accretion redshifts, z_in, are expected to dominate the inner regions of a halo. This is consistent with the radial segregation of subhalos by accretion time shown in Figure 6.7.

What is *not* consistent with the expectations of simple SSIMs are the presence of subhalos well beyond $\sim 1.6\,r_{200}$, roughly 288 h^{-1} kpc in the case of Aq-A. Nevertheless, $\sim$ 17% of all associated subhalos occupy such orbits, with several having apocentric radii exceeding $4\,r_{200}$! As previously discussed by Ludlow et al. (2009), the dynamics of these system are a result of 3-body interactions which can propel low-mass objects onto unorthodox orbits that would otherwise be unattainable

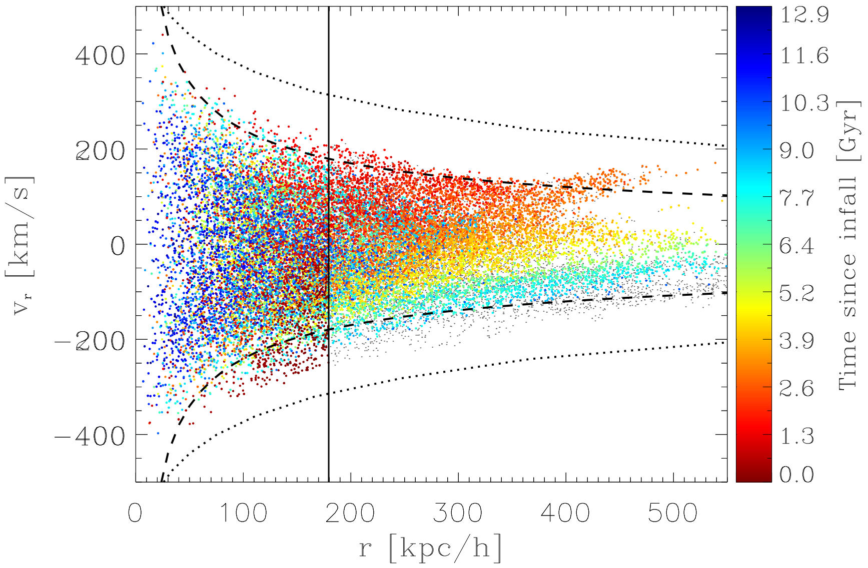

Figure 6.7: Radial velocity as a function of halo-centric distance for all ($z = 0$) halos and subhalos in Aq-A-2. Colored points show the distribution of *all associated* subhalos, defined as those that have at some point in the past been within the (time-dependent) virial radius of the main progenitor halo. All points have been colored by their accretion time onto the main halo; grey dots are used for halos that are on their first approach but have not yet crossed the halo's virial radius. Clearly subhalos that fall in early occupy tightly bound orbits in the central region of the main halo, with only moderate contamination by those that fall in at late times, as expected from simple self-similar infall models. The dashed (dotted) lines show the circular (escape) velocity of the main halo at a given radius. The virial radius is indicated using a solid vertical line. Note that many associated subhalos are today well beyond the virial radius of the main halo, with several receding at speeds that exceed the system's nominal escape velocity at their location.

(see also Sales et al. 2007). The sample of "ejected" subhalos, defined in section 6.2.4, contains such objects.

Can a similar dynamical process result in the rapid circularization of subhalo orbits, or is it a gradual process brought on by, for example, the orbital decay associated with dynamical friction? Clues can be gleaned from Figure 6.8, where we show the distribution of accretion times for associated subhalos in all six level-2 *Aquarius* runs. In each panel the grey histograms correspond to the distribution of $z_{\rm in}$ for *all* associated subhalos; red and blue histograms show the subsamples for objects whose orbits satisfy the criteria for circular or ejected orbits, respectively. For comparison, we show the mass accretion history[4] of the main halo as black points, which also illustrates the time resolution of the simulations (one point for each timestep). In most cases, the total number of infalling subhalos is a rather smooth function of time for intermediate and low redshifts, which reflects the

---

[4] We define the mass accretion history as the time evolution of the virial mass, M_{200}, of each halo's main progenitor.

smooth mass accretion histories. One exception is Aq-F which has a major merger between 3 and 5 Gyr ago, and a corresponding increase in the total number of accreted satellites. Note, however, that the $z_{\rm in}$ distributions for the circular and ejected subhalos are much more discontinuous. The majority of subhalos which are on circular orbits today fall in at early times, with a distribution that tapers off towards lower redshift. In fact, averaged over all simulations, roughly 50% of all circular subhalos are accreted in the first 4 Gyr after the Big Bang. This is roughly of the same order as the characteristic half-mass assembly times of each main progenitor halo (except for Aq-F, where this takes $\gtrsim$ 8 Gyrs), and is coincident with the rapid assembly phase of MW-mass DM halos (Bullock & Johnston, 2005). This preference for early accretion extends as well to the ejected subhalo population, but their distribution is shifted towards slightly later infall times relative to circular subhalos. The exception is Aq-F, in which roughly 50% of circular, and 70% of ejected subhalos are clearly associated with the late-time major merger, suggesting that three-body interactions are *capable* of leading to orbital circularization just as they are to ejection.

The early accretion times of subhalos on circular orbits, however, may indicate that their circularization is largely a result of dynamical friction decaying their orbits over long periods of time. In Figure 6.9 we plot the halo-centric radius and bound mass for several circular subhalos in Aq-F. Note that the circular subhalos with early accretion times (shown as black curves in this plot) *do not* show the characteristic orbital decay normally associated with dynamical friction. Rather, their orbits change very little since $z_{\rm in}$; they appear to have *always* been on nearly circular orbits. The evolution of their bound mass agrees with this interpretation: for the most part, the mass is smoothly stripped from the subhalo; it does not exhibit the periodic mass loss associated with the tidal decay induced by dynamical friction. The red curves in this show the orbits of several subhalos whose accretion times coincide with the late-time merger. These orbits clearly evolve differently than those associated with early accretion. These systems reach large turn-around radii ($r_{\rm turn} \approx 2 - 4\, r_{200}$) before infalling onto the main halo. After their first pericenter, the subhalos *lose* a considerable fraction of their infall energy leading to the immediate reduction of their apocentric radii. Take as an example the solid red curve in Figure 6.9. This subhalo had a turn-around radius of $\sim 3\, r_{200}$. Based on Bertschinger's SSIM model this object would therefore be expected to reach a first apocenter of $r_{\rm apo} \approx 0.9\, r_{\rm turn} \approx 410\ h^{-1}\,$kpc, rather than the observed apocenter of $\sim 100\ h^{-1}\,$kpc.

These results indicate that major mergers can lead both the the ejection of low-mass halos from the system, but also to the rapid circularization of others orbits. In the next section we focus on the late time merger in Aq-F in order to better understand how the merger dynamics can lead to such extreme events.

6.3.2 The major merger of Aq-F

Halo Aq-F is an interesting case because it contains the largest fraction of circular subhalos, and most of them accrete at late times. In addition, the infall time

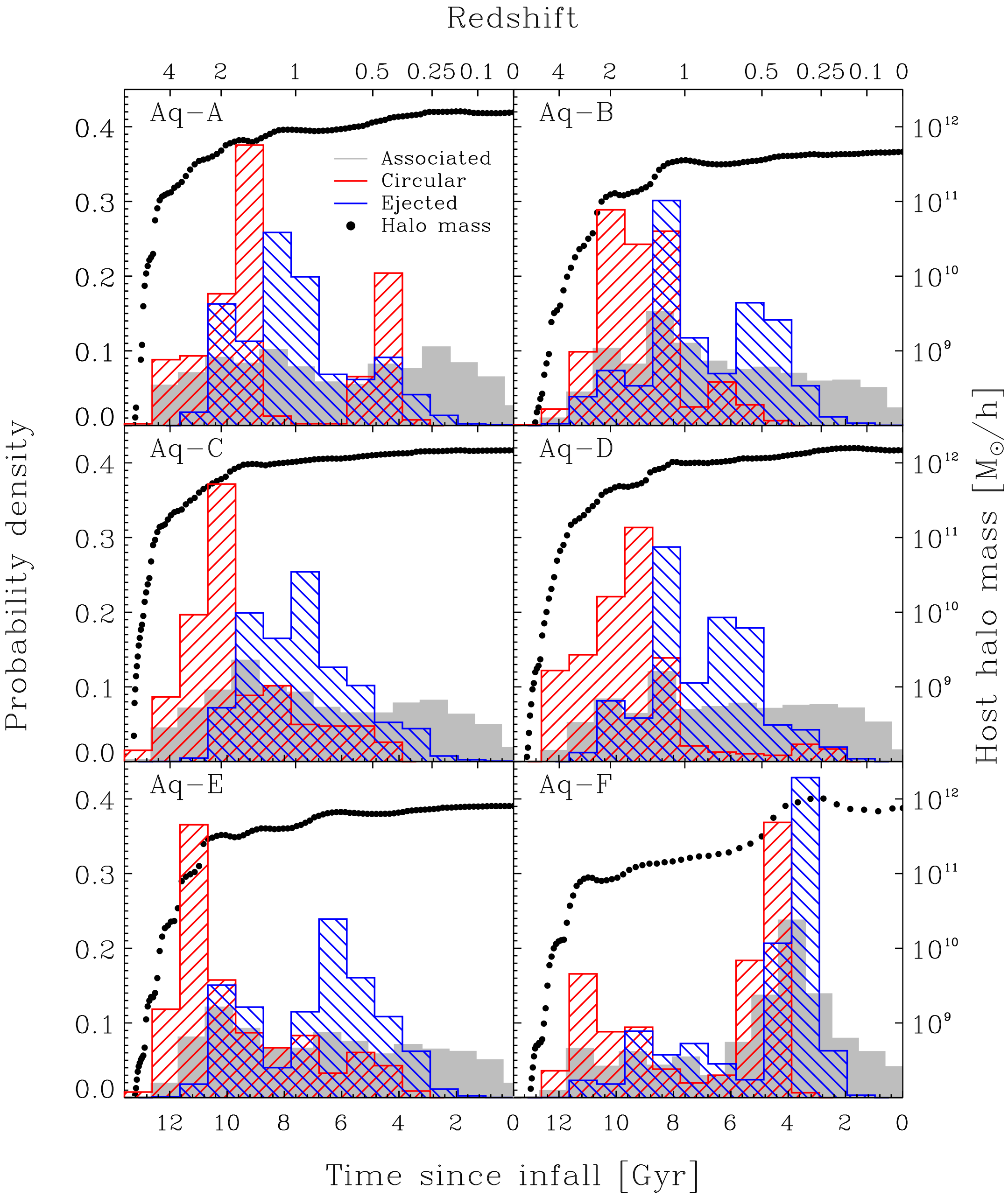

Figure 6.8: Accretion histories of dark matter (black circles) and subhalos (histograms) for all six *Aquarius* halos. Black circles trace the evolution of the virial mass, M_{200}, of the each halo's main progenitor. Grey histograms show the distribution of accretion times for all $z = 0$ associated subhalos; the red and blue histograms correspond to the subsamples of circular and ejected subhalos, respectively. The majority of circular subhalos appear to have been accreted at early times, when the halo's main progenitor was in a phase of rapid merging and assembly. One exception to this is halo Aq-F, in which ~ 45% of all circular subhalos are accreted during a late-time major merger. This suggests that rapid fluctuations in the potential driven by mergers can lead to the complete circularization of otherwise eccentric orbits.

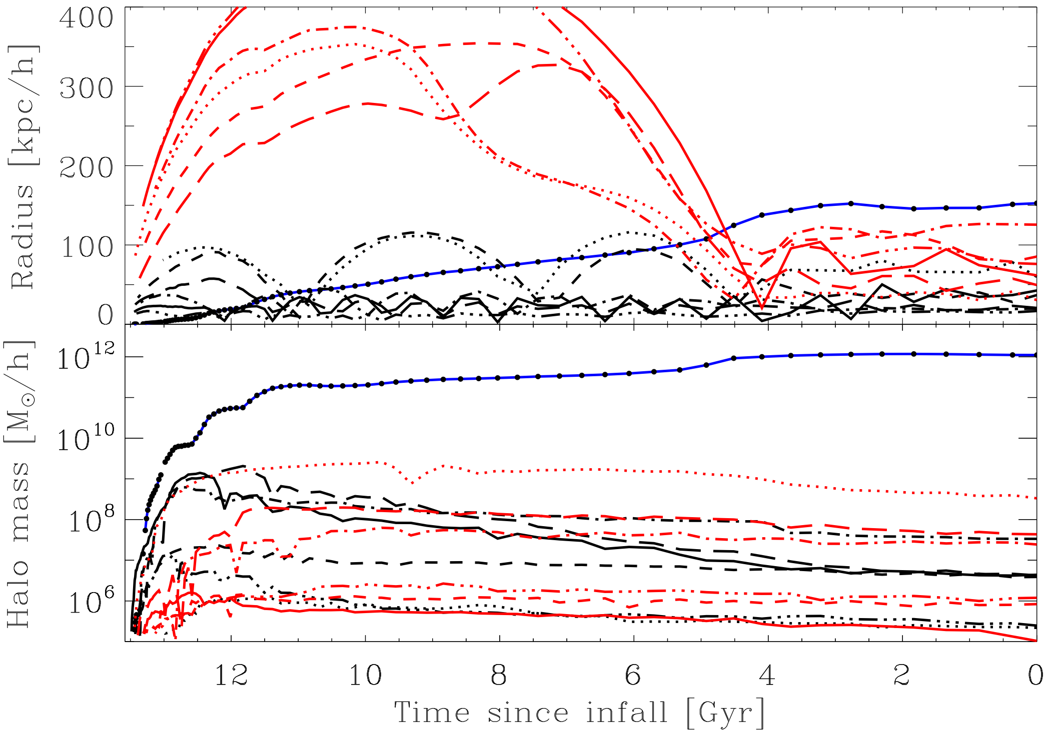

Figure 6.9: Halo-centric distance (top panel) and bound mass (bottom panel) as a function of look-back time for several circular subhalos identified in Aq-F. The virial radius and mass of the main halo is shown as blue lines with black dots representing each simulation snapshot. Black lines correspond to halos with early infall times, between 10 and 12 Gyrs ago, while red lines show several subhalos that are associated with the late-time major merger event (see Figure 6.8). Linestyles for a given subhalo are the same in both panels. Subhalos that fall in early are characterized by a small turn-around radius with very little change in subsequent apocentric radii. These objects are gradually stripped of a significant fraction of their original mass. This is contrary to the *episodic* reduction of apocentric radius and bound mass at pericenter expected from dynamical friction. Late infalling subhalos have large turn-around radii and circularize abruptly at first pericenter, which we interpret as being a result of the rapid loss of orbital energy induced by a three-body interaction.

distribution in Figure 6.8 also shows a peak for ejected subhalos at a similar, but slightly later time than the one for circular subhalos. In Figure 6.10 we plot the joint distribution of r_{turn} and r_{apo} for associated subhalos in Aq-F, each expressed in units of the halo virial radius. The full distribution of associated subhalos are shown as light grey points; the colored points show both the circular subhalos (open circles) as well as those that are ejected (crosses). The color coding of points indicates the infall time of each individual subhalo, with blue highlighting early and red late accretion times as in previous Figures.

The majority of subhalos (64%) have $r_{\mathrm{apo}} < r_{\mathrm{turn}}$, as expected from SSIMs. Note also that the circular subhalos (open circles, mostly above the 1:1 line) have a bimodal distribution of accretion times. This is also indicated by the red histogram in the right-hand panel, which gives the distribution of r_{turn} for circular subhalos. There is a clear peak at roughly $r_{\mathrm{turn}}/r_{200} \approx 0.8$ and a second at ≈ 2.2, which is consistent with bimodal accretion time distribution of the same subhalos shown in Figure 6.8, and the characteristic orbits shown in Figure 6.9. The “cloud” of orange points with $r_{\mathrm{turn}}/r_{200} \approx 2.2$ correspond to the subhalos whose orbits are rapidly circularized by the late-time merger. Ejected subhalos (crosses, mostly below the 1:1 line) have a narrower distribution of turn-around radii, with a median value of $r_{\mathrm{turn}}/r_{200} \approx 1.6$. They also have typical accretion times which are distinct from those of the circular subhalos, as can be seen from the slightly darker shade of red for the majority of ejected systems. The distribution of r_{apo} is flat and closely resembles that for all associated subhalos (grey histogram).

These differences in accretion times and turn-around radii can, in fact, be traced to the unique characteristics of the merger. Figure 6.11 shows two orthogonal projections of the spatial distribution of FOF dark matter halos at $z \approx 0.75$, roughly 1.5 Gyr before the start of the merger event. For clarity, individual FOF halos are shown as thin black circles (provided that their mass exceeds $3 \times 10^9\,h^{-1}\,\mathrm{M}_\odot$) with radii proportional to their mass: $r = (3\,M/800\,\pi\,\rho_{\mathrm{crit}})^{1/3}$. All halos are centered with respect to the location of the main progenitor of Aq-F at this redshift (thick black circle at the origin). The two most massive halos (apart from the main halo) are two merging partners with approximately the same accretion times, marked with M_1 and M_2 in the right hand panel. These objects have mass ratios of $M_1/M_{\mathrm{host}} \sim 0.36$ and $M_2/M_{\mathrm{host}} \sim 0.1$.

Also plotted in Figure 6.11 are the locations of circular (red) and ejected (blue) *subhalos* that have accretion times that approximately span the duration of the merger. This reveals an astonishing disparity: $\sim 90\%$ of the circular subhalos are associated with M_1 prior to infall, and $\sim 75\%$ of ejected ones with M_2. Note that basically all of these subhalos are associated with one of the merging partners prior to infall. This suggests that complex merger geometries can readily generate extreme subhalo orbits, from highly circular to radial.

6.3.3 Orbital circularization

How does a merger event give rise to a population of circular or ejected subhalos? Insight can be gathered from Figure 6.12, where we plot the radial velocity-distance

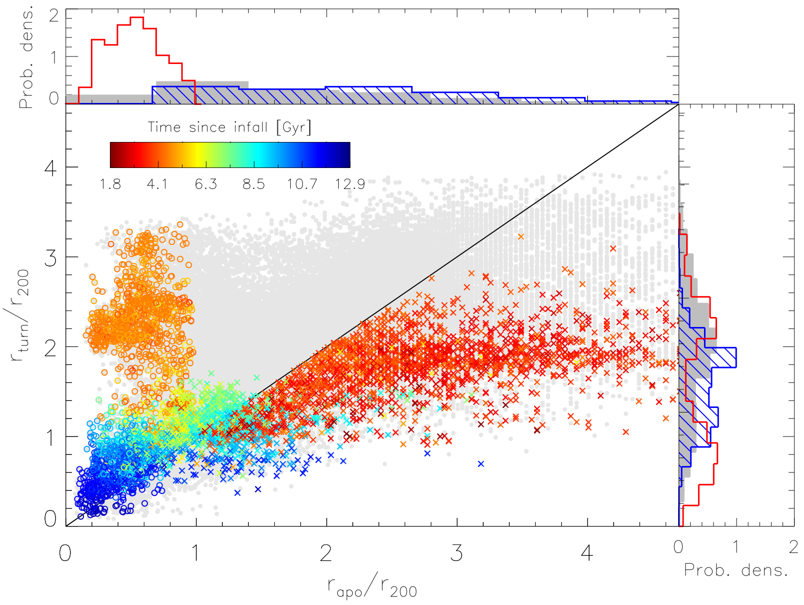

Figure 6.10: Turn-around radius vs. orbital apocenter for associated subhalos in Aq-F-2, normalized to $r_{200}(z = 0)$. Circular and ejected subhalos are marked by open circles and crosses, respectively, and are colour-coded by their infall time. Grey points show the remaining associated subhalos. The histograms in the upper and right-hand panels show the distribution of r_{apo} and r_{turn} for each of the 3 subsamples (red for circular subhalos, blue for ejected ones, and grey for the full associated sample). The bimodal distribution of accretion times for circular subhalos is clearly seen in their turn-around radii. Because of their late accretion times, many subhalos associated with this merger have not yet completed one full orbit within the main halo. In these cases, apocenters have been estimated from the radial energy equation.

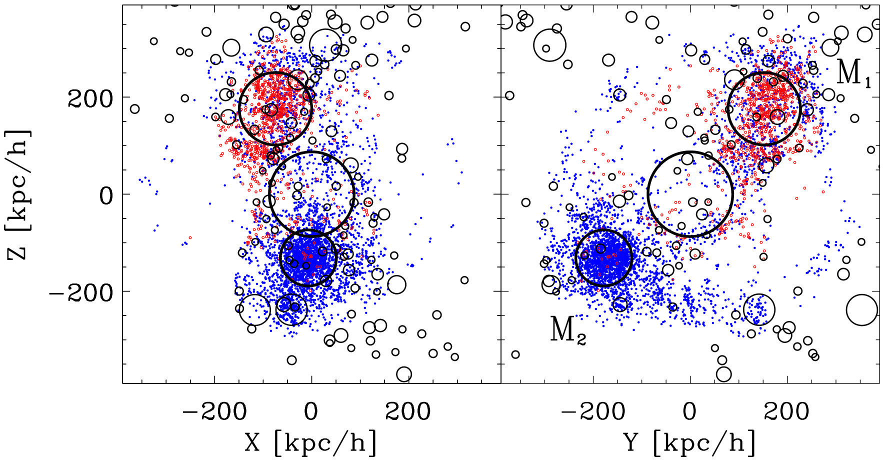

Figure 6.11: Z-X and Z-Y projections of halo positions in a cube with sidelength 800 h^{-1} kpc for Aq-F-2 at $z \sim 0.75$, shortly before the major merger, centered on the main halo (central thick black circle). Thin black circles correspond to isolated friends-of-friends halos with $M_{\text{FOF}} > 3 \times 10^{9}\, h^{-1}\, \mathrm{M}_{\odot}$; the radius of the circle is calculated from the FOF mass: $r = (3\, M/800\, \pi\, \rho_{\text{crit}})^{1/3}$. Colored points indicate subhalos that, at $z = 0$, are on either circular orbits (red) or ejected ones (blue), and whose accretion times coincide with the time of the major merger. The circular and ejected subhalos are clearly associated with two distinct mergers that occur nearly simultaneously (thick black circles marked with M_1 and M_2 in the right panel).

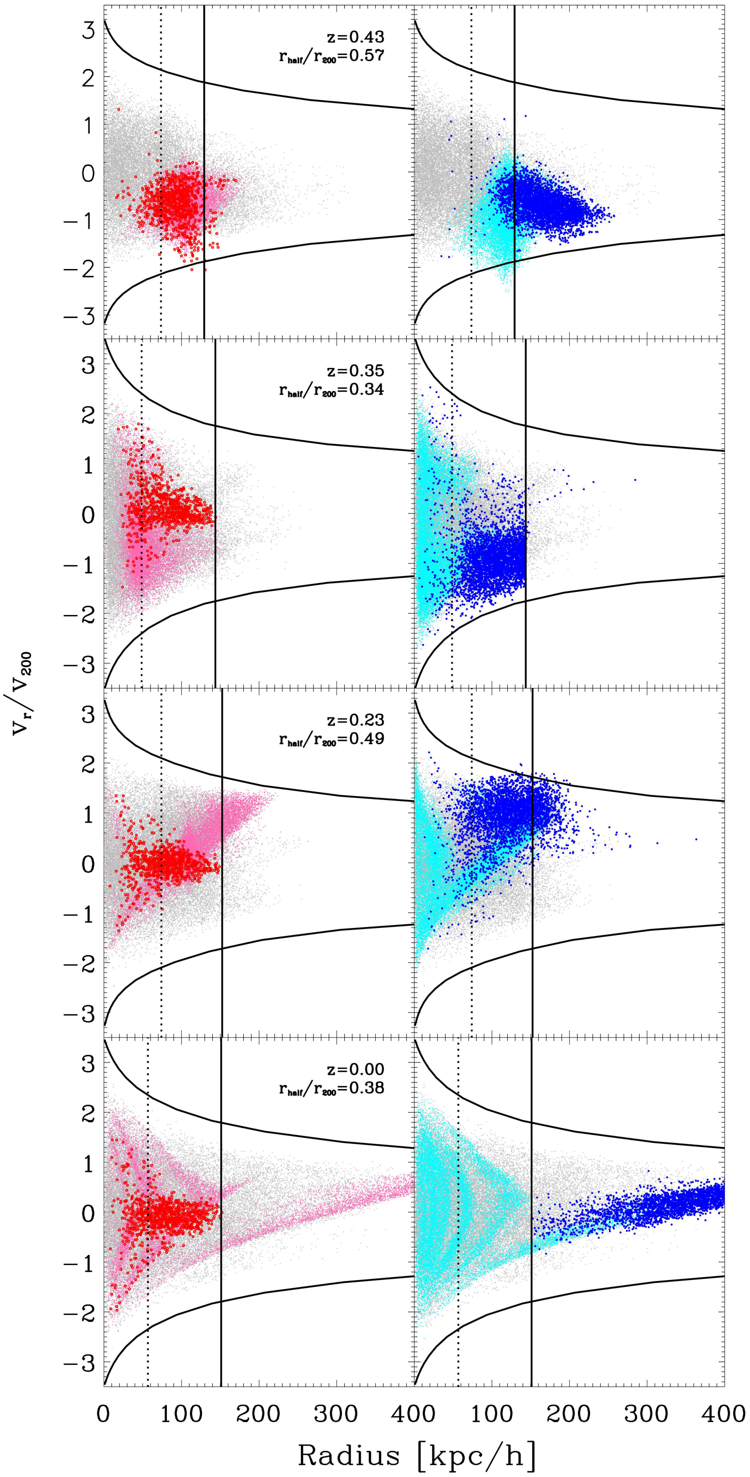

Figure 6.12: Radial velocity as a function of distance from the main halo for subhalos and dark matter particles associated with the late-time merger in Aq-F. In each panel, grey dots show a random subsample of dark matter particles (from the level-4 data) belonging to the $z = 0$ main halo or, at $z > 0$, to its most massive progenitor. Colored circles correspond to circular subhalos (red), or ejected ones (blue); pink or cyan dots show a random sample of DM particles associated with each merging object. From top to bottom, different panels correspond to different characteristic times during the merger: the first infall of M_2 (top panels), M_2's first pericentric passage (second panels), one crossing time since $z_{\rm in,1}$ (third panels), and finally at $z = 0$. The lines show the circular velocity profile of the host halo at each snapshot. Vertical solid and dotted lines mark the host's virial and half-mass radius; the "concentration", $r_{\rm half}/r_{200}$, is given in each panel.

relation for dark matter and subhalos at four characteristic times during the merger event. We adopt the same color-coding as in Figure 6.11: circular subhalos are shown as red open circles, and ejected subhalos as blue filled circles. The pink and cyan dots correspond to *dark matter particles* associated with M_1 and M_2, respectively, and the grey dots show a random sample of DM particles drawn from the main halo (or, at $z > 0$, from its most massive progenitor). From top to bottom, the redshifts correspond to $z_{\rm in,2}$, $z_{\rm peri,2}$, $z_{\rm char,1}$ and $z = 0$, respectively. Here, $z_{\rm char,1}$ corresponds to the time $t_{\rm in,1} + 2\, r_{200}(t_{\rm in,1})/V_{200}(z_{\rm in,1})$, which is roughly one crossing time after $z_{\rm in,1}$ ($\sim$ 2 Gyrs after infall). The legend also provides a measure of the main halo concentration, $r_{\rm half}/r_{200}$, where $r_{\rm half}$ is the radius inside which half of the mass is located (dashed vertical line); a low value of this parameter therefore indicates that the halo is more compact.

It is clear from the upper panels of Figure 6.12 that $z_{\rm in,1} > z_{\rm in,2}$; M_1 has already crossed the virial radius of the main halo (indicated by the solid vertical line) at the point that M_2 is first infalling at $z_{\rm in,2} \approx 0.43$. At $z = 0.35$ M_1 has gone through is first pericentric passage and is now moving towards its apocenter, but the mass added by M_2 creates a deeper potential well that "traps" subhalos around $v_{\rm r} \sim 0\,{\rm km/s}$. This interpretation is clearly supported by the red points in Figure 6.12, which remain approximately stationary in the $v_{\rm r} - r$ plane since $z_{\rm peri,2} \approx 0.35$, caused by a decrease in orbital energy. Such an event is not conducive to orbital energy gain, however, which explains why M_1 contains virtually no ejected subhalos. The ejection of subhalos can nevertheless be understood in terms of the same merger dynamics. The right hand panels of Figure 6.12 shows that ejected satellites accrete onto the main halo slightly later than their host's center of mass. These subhalos are themselves approaching pericenter within the infalling group at the same time that the group itself is approaching pericenter within the main halo. The deepened gravitational potential associated with the merger of M_2 results in an excess gravitational pull on these subhalos, leading to uncharacteristically high pericentric velocities, in contrast to those associated with M_1 where $v_{\rm r} \sim 0\,{\rm km/s}$. The dense merging core then oscillates into a state of maximal expansion (during the virialization process) and the subhalos can escape the shallow core with excess energy.

This process results in distinct differences in the eccentricity and orbital energy distributions of subhalos associated with either M_1 or M_2. In Figure 6.13 we plot the distribution of ϵ^5 and the orbital energy $E_{\rm orb}$ normalized to $V_{200}^2(z = 0)$ of the host for subhalos that are associated with either M_1 (red histograms) or M_2 (blue histograms) prior to infall. The grey histograms show for comparison the same distributions for *all* subhalos that are accreted over the narrow redshift range associated with the merger. The overall redistribution of energy discussed above is clearly seen here. For example, the median eccentricity for subhalos infalling with M_1 is $\widetilde{\epsilon}_1 \sim 0.51$, whereas $\widetilde{\epsilon}_2 \sim 0.67$. Similar trends can be seen in the subhalo's present-day orbital energies: those infalling with M_1 are *today* on more tightly bound orbits, with median values $\widetilde{E}_{\rm orb,1} = -2.5\, V_{200}^2$ and $\widetilde{E}_{\rm orb,2} = -1.8\, V_{200}^2$.

---

[5] To construct this plot we use ϵ_{pot} only for subhalos for which $\epsilon_{rec}$is undefined.

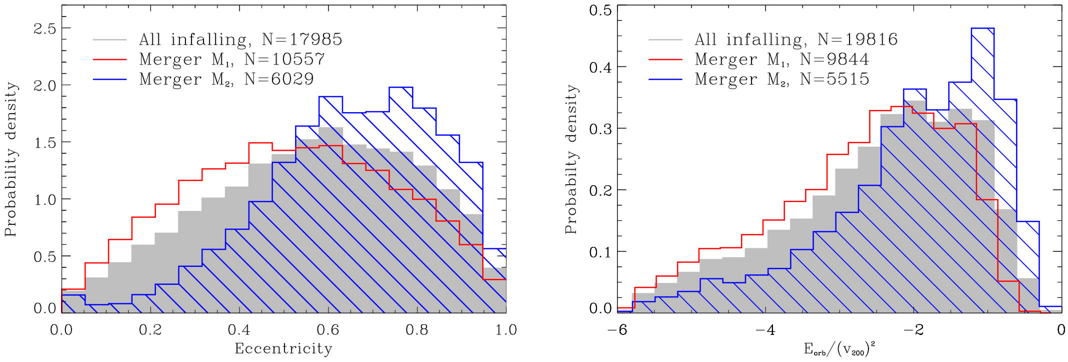

Figure 6.13: Probability distribution of orbital eccentricities (left panel) and orbital energy (right panel) for subhalos that are associated with either M_1 (red) or M_2 (blue, hatched), as well as *all* subhalos that are accreted in the same narrow redshift interval (grey, filled). The total number of halos in each subsample is provided in the legend, and differs between the panels because for some halos these values can not be calculated.

Net energy gains and loses such as these are a direct consequence of a violent form of relaxation driven by global fluctuations in the potential of the main halo during major mergers. The statistics of this process were first laid out in Lynden-Bell (1967) and invoked to explain the origin of the smooth and symmetric luminosity profiles of elliptical galaxies. A key point is that the timescale associated with violent relaxation is on the order of a crossing time, and so results in very rapid redistribution of the energies and angular momenta. With this in mind we provide a final view of the merger process in Figure 6.14. Here, the connected blue circles trace the evolution of the virial mass, M_{200}, of the main progenitor of Aq-F, and the green squares show the corresponding concentration of the halo, $r_{\rm half}/r_{200}$. The arrows indicate the time of infall, and of pericentric passage, where $t_{{\rm peri},i} = t_{{\rm in},i} + r_{200}(t_{{\rm in},i})/V_{200}(t_{{\rm in},i})$, one half of a crossing time after infall for halo i.

The start of the merger coincides with a sharp decrease in halo compactness, represented by an increase in the concentration parameter. At $t_{\rm in,2}$, for example, $r_{\rm half}/r_{200} \sim 0.57$. This is a factor of ~ 1.7 higher than the *minimum* concentration of ~ 0.34 occurring shortly after, at roughly $t_{\rm peri,2}$. This temporarily renders the system in a state of maximal compression leading to a sudden drop in the central gravitational potential, followed by an increases as the system oscillates to a new state of equilibrium. These oscillations in the dynamical state of the halo affect subhalo orbits differently depending on a complex interplay between their initial locations and energies. Imagine, for example, a test particle falling into the gravitational potential of a dark matter halo. If the potential of the halo is static no relaxation occurs; the total energy and angular momentum of the particle are integrals of motion and hence remain constant in time. If the potential rapidly deepens after the particle crosses pericenter the particle *loses* energy, since it sees a deeper potential than when it first fell in. If instead the potential rapidly expands then a net energy *gain* results from the weaker gravitational attraction of the host halo.

This concludes our description of the double merger of Aq-F. We have shown

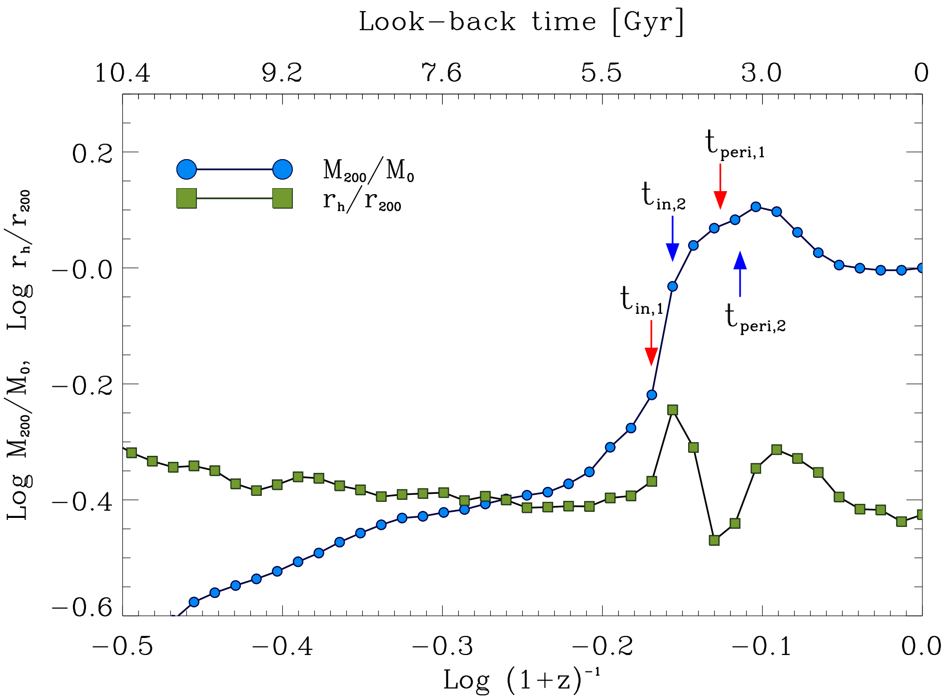

Figure 6.14: Evolution of the virial mass, M_{200} (blue circles), and concentration (green squares) for the main progenitor halo of Aq-F. The halo concentration is defined here as the ratio between the half-mass and virial radii at any given time, $r_h(z)/r_{200}(z)$. Arrows labeled $t_{in,1}$ and $t_{in,2}$ indicate the infall times of M_1 and M_2, respectively. The subsequent arrows, labeled $t_{cr,i}$, correspond to one half of a crossing time later, i.e. to $t_{in,i} + r_{200}(t_{in,i})/V_{200}(t_{in.i})$, which roughly coincide with the pericenter of halo i. Note that the start of the merger phase coincides with a rapid decrease in the concentration of the main halo; a maximum concentration is reached at approximately the same time as the first pericenters of the merging pair.

that a specific merging event can create a large number of both circular and ejected subhalos. In the next section, we attempt to connect these simulation-based results to the observed dwarf galaxies of the Local Group.

6.3.4 Monoceros candidates

In the previous sections we discussed how mergers can result in the rapid circularization of infalling subhalo orbits. A low eccentricity orbit, however, is only one prerequisite for the Monoceros stream. Other aspects to consider are its location within the halo, its mass, and whether or not the system could have plausibly trapped gas in order to fuel star formation. In this section we discuss how these constrains affect the frequency of Monoceros-like systems in the *Aquarius* simulation. We emphasize however that large uncertainties associated with modeling the MO and identifying *luminous* galaxies in collisionless N-body simulations precludes a detailed modeling of MO-like systems.

Observations suggest that the separation of the MO from the Galactic center is at most ~ 20 kpc, and a lower limit on its mass comes from the observed stellar mass of the stream: $M_* \geq 1.5 \times 10^7\, h^{-1}\, \mathrm{M}_\odot$ (converted to our value of $h = 0.73$,

Table 6.2: Number of subhalos that fulfill the selection criteria discussed in this section, for the level-2 simulations as well as Aq-A-3 for comparison. The first column shows the number of subhalos that exist at $z_{\mathrm{ion}} = 9$, out of which we select those with $v_{\max} > v_{\mathrm{crit}} = 10\,\mathrm{km/s}$, which significantly reduces the sample. The third column shows the number of "luminous" halos with radial separation $r < 0.1\,r_{200}$ which is the upper observational bound for the MO. This already emphasises the rarity of such objects, even without any constraints on their orbital properties. The fourth column shows the total number of circular subhalos found at redshift 0. When the star formation criterion is applied to this sample, we obtain our final "MO-like" sample (no radius constraint is used here).

Halo	N_{sub} at z_{ion}	$v_{\max} > v_{\mathrm{crit}}$	$r < 0.1\,r_{200}$	$\epsilon < 0.2$	MO-like
Aq-A-2	19058	697	98	409	14
Aq-A-3	5502	598	21	121	7
Aq-B-2	19919	371	116	1117	19
Aq-C-2	16328	707	80	477	21
Aq-D-2	19772	829	161	490	22
Aq-E-2	20078	605	100	818	36
Aq-F-2	29194	521	171	1891	36

Yanny et al. 2003). Note that this is a conservative lower limit since the mass-to-light ratios of Galactic dwarfs are $M/L \geq 5\,\mathrm{M}_\odot/L_\odot$, with significantly larger values for lower luminosities (e. g. McConnachie 2012). An estimate of the progenitor's *initial* mass can be obtained from the best-fitting model of Peñarrubia et al. (2005), which gives $M_{\mathrm{in}} = 0.73 \times 10^9\,h^{-1}\,\mathrm{M}_\odot$. Nevertheless, as discussed by Peñarrubia et al. (2005), estimates of the progenitors mass are poorly constrained and can vary significantly for different assumptions regarding the shape of the Milky Way's dark matter halo. For this reason, we do not attempt to restrict our sample of circular subhalos to Monoceros-like systems based on (poorly known) mass constraints.

Stellar streams require progenitors that are capable of forming stars. Star formation is generally thought to occur in dark matter halos massive enough to trap and cool gas prior to its heating by reionization. Because the *Aquarius* simulations do not contain any gas or star formation recipe we must resort to an empirical model and the properties of the dark matter halos to judge whether they could have hosted star formation. Hydrodynamic simulations of galaxy formation have shown that halos with $v_{\max} \gtrsim 10\,\mathrm{km/s}$ at or before reionization are sufficiently massive to trap and cool gas, which leads to significant star formation (Okamoto & Frenk 2009, Geen et al. 2013, Sawala et al. 2013). This offers a simple and plausible means of associating dark matter halos with luminous objects. In fact, these same arguments have been used to reconcile the disparity in the mass function of dark matter substructures found in N-body simulations with the paucity of observed satellites around the Milky Way (citations needed). However, reionization is not instantaneous and the redshift range over which it occurs is not particularly well known. Most observations indicate that it lies somewhere between $z = 12$ and $z = 6$. For simplicity, we assume $z_{\mathrm{ion}} = 9$ and use $\max\left[v_{\max}(z > 9)\right]$ to flag for subhalos that may today be visible through their starlight.

We therefore adopt constraints on orbital eccentricity ($\epsilon < 0.2$) and peak circu-

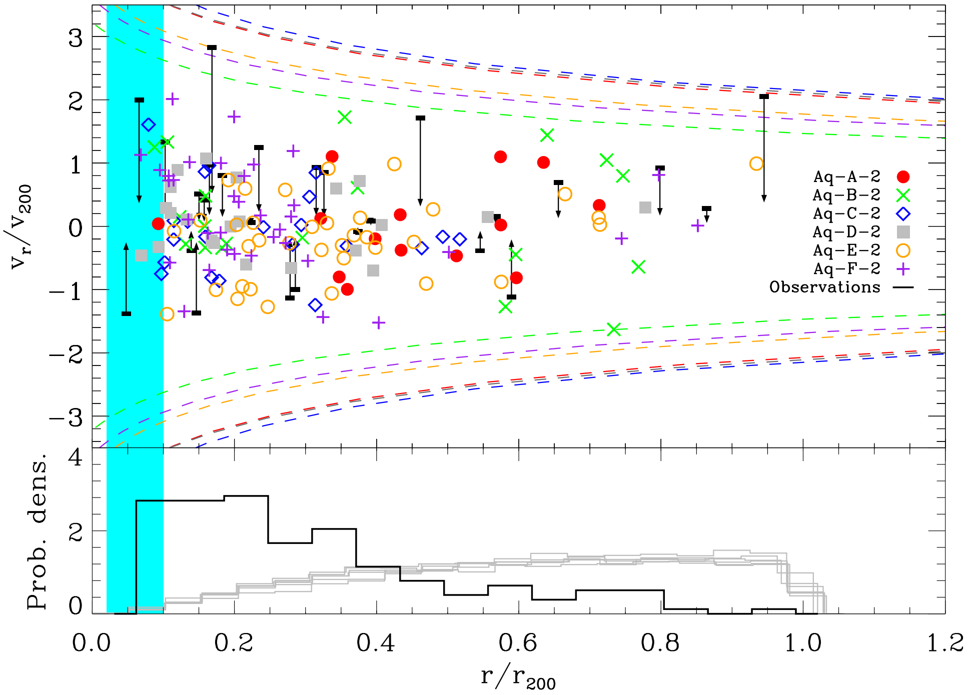

Figure 6.15: *Top panel:* Radial velocity as a function of distance for MO-like subhalos in all six level-2 *Aquarius* halos (different symbols), scaled with r_{200} and v_{200} of their respective halo. The black arrows show the properties of the Milky Way *dwarf galaxies* from McConnachie (2012). The length of the arrow indicates the uncertainty from the unknown tangential velocity of the dwarf galaxies (for details see text). The dashed contours show the escape speed for the individual halos. The cyan region marks the broad radial band where the Monoceros stream has been observed, which includes the Canis Major ($v_r < 0$) and Sagittarius ($v_r > 0$) dwarf galaxies. *Bottom panel:* Radial distribution of MO-like subhalos, combined for all six simulations (thick black line), and for all associated subhalos in each of the simulations (thin grey lines) for comparison.

lar speed at $z_{\rm ion}$ in order to estimate the rate of occurrence of MO-like systems. In general, the Monoceros overdensity (and Canis Major dwarf galaxy) are expected to be extremely rare in galaxies with quiescent assembly histories similar to the *Aquarius* halos. This is reflected in the total numbers of subhalos that pass the different cuts, given in Table 6.2. The constraint on $v_{\rm max}(z > 9)$ reduces significantly the number of subhalos which could be possible progenitors for Monoceros: only $3 - 10\%$ of associated subhalos are able to form stars based on this criterion, and of these systems only $\sim 1.2\%$ to 7% find themselves on circular orbits today. Since this leaves us with very few candidates, hereafter we combine the results from all six *Aquarius* halos by scaling their properties with those of their respective hosts.

The top panel of Figure 6.15 shows the radial velocity of all MO-like subhalos, scaled with r_{200} and v_{200} of their respective host. For comparison, we overplot the properties of the known Galactic dwarfs (from McConnachie 2012) as black arrows, assuming a virial radius for the MW of $r_{\rm MW} \sim 198\ h^{-1}\,$kpc[6]. Note that

---

[6] The observational constraints on the MW's virial radius are rather weak, ranging from $\sim 146\ h^{-1}\,$kpc to $\sim 250\ h^{-1}\,$kpc, with errors of the order of $50\ h^{-1}\,$kpc (e. g. Li & White 2008, Wang et al. 2012, Boylan-Kolchin et al. 2013). We therefore adopt the median value between

the velocities of the dwarf galaxies have been measured as radial velocities in the Solar rest frame, and the unknown tangential component translates into an systematic uncertainty in the Galactic rest frame, which is much larger than the actual measurement error (McConnachie, 2012). The tip of the arrow assumes the tangential velocity in the Solar rest frame to be zero, whereas the base of the arrow assumes a tangential component of the same magnitude as the measured radial component, therefore providing an upper limit. The dashed contours in Figure 6.15 show the escape speed for the different *Aquarius* halos. The cyan region marks the broad radial band between 7 and 15 h^{-1} kpc where the Monoceros stream has been observed, which includes the Canis Major ($v_\mathrm{r} < 0$) and Sagittarius ($v_\mathrm{r} > 0$) dwarf galaxies. Inside this narrow radial band, only a handful of subhalos remain: only 9 out of ~ 75000 *Aquarius* subhalos across all six simulations have properties similar to those predicted for a Monoceros progenitor.

6.4 Conclusion

We have used a set of six high-resolution dark matter halos from the *Aquarius* simulations to investigate the occurrence and origin of subhalos on circular orbits within the inner regions of a halo. By tracing each subhalo through its dynamical evolution, we were able to determine the orbital properties of *associated* subhalos, i. e. those that have been inside r_{200} of the host at any point in their evolution. We compared our own tracking method with the results from a merger tree analysis and found that neither the specific method of subhalo tracking nor the resolution level of the simulation has a significant influence on our results. This is in agreement with the excellent numerical convergence of the *Aquarius* simulations found in previous works.

In order to define a subsample of circular orbits, we investigated two ways of defining the eccentricity: through their orbital energy, Equation (6.2), and by directly determining the apo- and pericenters from their orbits. In general, we found that the accuracy of the orbital energy equation depends strongly on the degree of relaxation of the host halo. In particular, this affects the calculation of the pericentric radius, because the inner part of the potential is more susceptible to disruptions by mergers. Using the eccentricity determined from the orbits leads to a “cleaner” sample of circular orbits, but also reduces the sample to only those subhalos that have completed at least one full orbit within the main halo. In accordance with previous studies, we found that the infall time strongly influences the dynamical properties of a subhalo: those that fall in early will reach smaller radial separations and lose a larger fraction of their mass by tidal stripping.

We find clear indications that circular subhalos can either come from early times, when they fall in with already low eccentricity, or from late times in association with mergers. Not only do the infall times correlate with the mergers, but also the subhalo’s association with another substructure plays an important role. This is

---

these extremes to scale the observed dwarf galaxy distances and velocities. Note however that the results do not depend very strongly on this choice.

particularly evident in the case of Aq-F, which has a double merger at low redshift. While a significant fraction of subhalos associated with the first merger ends up on circular orbits, the second merging halo basically contributes ejected subhalos only. Investigating this double merger in both level-2 and level-4 simulations, and find that the deepening of the potential well by the second infall traps many subhalos belonging to the first merger and allows them to circularise more efficiently.

In the final section, we compare the properties of the subhalos in our circular subsample with those of the peculiar Monoceros stellar stream observed in the Milky Way, which may be a remnant of an infalling satellite galaxy. In addition to having a circular orbit at small radial separation, the subhalos must also be able to form stars prior to infall. We enforce this criterion by applying a cut in the maximum circular velocity, v_{max}, prior to the epoch of reionisation, which is commonly used to separate "luminous" from "dark" subhalos in simulations without star formation. We compare the properties of the circular, luminous subhalos with the distribution of dwarf galaxies around the Milky Way and the orbit of the Monoceros stream. We find that most subhalos have larger orbital radii than Monoceros, re-emphasizing the rarity of this object. Across all six simulations, only a handful of star-forming subhalos have properties similar to Monoceros.

We conclude that substructures on circular orbits are extremely rare in simulated cold dark matter halos. Nevertheless, major mergers may provide the physical conditions needed to rapidly circularize orbits of infalling systems and thereby give rise to stellar streams such as the Monoceros overdensity. Whether these findings apply generally to major mergers, or are the result of idiosyncrasies within the *Aquarius* halo sample, will require further investigation.

7 Summary & Outlook

In this thesis we have investigated several problems related to cosmological structure formation and numerical simulations. The first three chapters provided an introduction to the ΛCDM model, the concept of cosmological structure formation, and the principles of N-body simulations. We will now summarize our findings and provide a short outlook of future work for each topic.

Perturbation theory and biasing In Chapter 4 we tested third-order Standard Perturbation Theory (SPT) as an approximation to non-linear cosmological structure formation. A novel approach was used to numerically calculate the three-dimensional dark matter density field using SPT from the initial conditions of two high-resolution cosmological simulations. The calculated density field was compared to the non-linear dark matter field of the simulations both point-by-point and statistically. For smoothing scales above $8\,h^{-1}\,\mathrm{Mpc}$ it showed a good agreement up to redshift 0. We presented a simple fitting formula to relate the linear and non-linear density contrast that accurately recovers the non-linear time evolution for $0 \leq z \leq 10$ at the per cent level.

To address the problem of biasing between the matter field and the halos identified in the simulation, we employed the Eulerian local bias model (ELB), including non-linear bias up to the third order. The bias parameters were obtained by fitting a scatter plot of halo and matter density (both from the simulation and from SPT). Using these bias parameters, we were able to reconstruct the halo density field. We found that this reconstruction is not able to capture all the details of the halo distribution. In addition, we investigated if the large scale bias can be described by a constant and if it corresponds to the linear bias parameter b_1 of the local bias model. We also discussed how well the halo-halo power spectrum and the halo-mass cross spectrum from the reconstructed halo density field agree with the corresponding statistics from the simulation. The results show that while SPT is an excellent approximation for the matter field for suitably large smoothing scales even at redshift 0, the ELB model can only account for some of the properties of the halo density field.

This work is easily extendable because our SPT calculation produces the dark matter velocity field as a by-product of the density field. The velocity field is crucial for connecting observations and theoretical models of large-scale structure formation, because of *redshift-space distortions* (e. g. Kaiser, 1987; Hamilton, 1998; Hawkins et al., 2003). Since distances in cosmology are often measured by redshifts, the peculiar velocities of galaxies with respect to the homogeneous Hubble flow result in distortions of galaxy correlation functions on the smallest scales. For example, a galaxy falling into the potential well of a large cluster will appear further away (closer to us) when it is between us and the cluster (behind the cluster) because its peculiar velocity creates an additional positive (negative) contribution to its measured redshift. Corrections for these effects rely on accurately estimating the peculiar velocity field, often employing linear theory or perturbative extensions (e. g. Heavens et al., 1998; Scoccimarro, 2004; Taruya et al., 2010). Our calculations will enable us to test these assumptions and/or provide more accurate models for the mildly non-linear regime.

Primordial non-Gaussianity from large-scale structure In Chapter 5 we considered the effect of primordial non-Gaussianity (PNG) on the galaxy bias, which can be used to detect and constrain deviations from standard single-field Inflation. The *local* model of primordial non-Gaussianity makes strong predictions for the galaxy bias on large scales, introducing a scale-dependence which is absent in the Gaussian case. Current observational constraints from galaxy clustering on f_{NL} (skewness) and g_{NL} (kurtosis) assume that the other PNG parameters are vanishing. Using two sets of cosmological N-body simulations where both f_{NL} and g_{NL} are non-zero, we showed that this strong assumption generally leads to biased estimates and spurious redshift dependencies of the parameters. Additionally, if the signs of f_{NL} and g_{NL} are opposite, the amplitude of the scale-dependent bias is reduced, possibly leading to a false null detection. Finally, we showed that model selection techniques like the Bayesian evidence can (and should) be used to determine if more than one PNG parameter is required by the data. Future LSS surveys are expected to yield PNG parameters with an accuracy compatible with those from the CMB, so it is important to understand and avoid this possible degeneracy in order to confirm the constraints coming from the early Universe.

A natural extension of this work would be to explore the constraints from the galaxy bispectrum and trispectrum. In the local PNG model, the matter bispectrum is directly proportional to f_{NL}, and the trispectrum is proportional to g_{NL}. This can be used to break the degeneracy between the two parameters that appears for the scale-dependent bias in the power spectrum. Theoretical modeling for the bispectrum or trispectrum is even more complicated than for the power spectrum, but numerical simulations could be used to study the expected signals in detail. In order to beat down statistical uncertainties, these simulations have to cover large volumes, while keeping the particle resolution high enough to explore a large range in halo masses. Precise predictions of the effects of PNG on large-scale structure provide insights on the generation of perturbations in the very early universe, and a confirmation of a deviation from Gaussianity would constitute a major discovery in cosmology and fundamental physics.

In addition, further constraints from large-scale structure probes would require the improved understanding of systematics in the survey data. This could be achieved for example by cross-correlating the spatial clustering of different galaxy populations as shown for example in Giannantonio et al. (2014). This type of study not only yields robust results, but can also lead to the discovery of previously unknown effects caused by the observation strategy, the instruments, or the data processing pipeline.

The Monoceros stream and subhalo orbits In Chapter 6 we have focused on near-field cosmology and properties of the Local Group environment. The formation history of the Milky Way and local environment – especially its dwarf galaxy population – provide and important testbed for both cosmological models and galaxy formation. We have used a set of six high-resolution dark matter halos from the *Aquarius* simulations to investigate the occurrence and origin of subhalos on circular orbits within the inner regions of a halo. By tracing each subhalo through its dynamical evolution, we were able to determine their orbital properties like eccentricity and infall time. We found clear indications that circular subhalos either fall in with already low eccentricity during the formation of the main halo (*early mode*), or later on only in association with mergers of the host halo (*late mode*). In addition, we also defined a second sample of subhalos: those that were accreted onto the main halo but are not able to lose enough energy, and are subsequently ejected again.

Since the *Aquarius* halos were selected specifically to resemble the quiescent formation of the our Milky Way, these late-time mergers are rare, and only happen in one of the six halos. This double merger event at $z \sim 0.5$ lead to an uncharacteristic increase in circular subhalo orbits via the *late mode*, and was therefore studied in more detail. The two primary classes of subhalos orbits, circular and ejected, are clearly associated with the first and second (by infall time) merging partner, respectively. This can be explained by the change in potential well of the main halo during the double merger: the second merger happens when the first merging partner has reached its pericenter, and thus its subhalos are captured in the deepening potential well caused by the second merger. When the main halo potential rebounds, the subhalos associated with the second merger have not had have enough time to be captured, and are ejected very efficiently. This increase in the late accretion mode is of particular interest for a possible MO progenitor, because an early infall would have led to a much more pronounced effect of tidal stripping and is difficult to reconcile with the observed coherent stellar remnant.

In order to generate a stellar stream like the MO at redshift 0, the infalling subhalo must also be able to form stars prior to its accretion by the host. We enforced this criterion by applying a cut in the maximum circular velocity prior to the epoch of reionisation to separate “luminous” from “dark” subhalos in our dark-matter simulations. This allowed us to compare the properties of luminous subhalos on circular orbits with the observed distribution of dwarf galaxies around the Milky Way. The star-formation criterion significantly reduces the number of subhalos that can be compared to MW dwarfs, even before applying the orbital circularity condition. Finally, the MO’s observed orbital radius of ~ 18 kpc is found to be another strong selection criterion, further reducing the number of subhalos which can be considered as possible progenitors. We therefore conclude that subhalos on circular orbits (and in particular, with the properties of the

Monoceros stream) are very uncommon objects within the context of dark matter simulations of MW-type halos, even when their number is enhanced by rare merger events.

There are at least two general directions for future work on this topic.[1] Our finding that late-time (double) mergers significantly enhance the abundance of circular orbits is only based on one out of six *Aquarius* halos. A larger sample of MW-like halos of similar or higher mass resolution would therefore be required to validate this hypothesis. Higher mass and force resolution would also provide a more complete sample of subhalos with properties compatible to the MO, and allow us to probe the innermost regions of the halo more accurately.

Secondly, instead of using the star-formation cut on the maximum circular velocity, one could include an actual stellar component by "post-processing" of the simulations. Cooper et al. (2010) developed a method to "tag" a fraction of the dark matter particles with stellar properties based on the semi-analytical galaxy formation code GALFORM (Cole et al. 1994, 2000; Bower et al. 2006). This process is repeated for all snapshots, so a star formation history can be defined for each tagged particle. Their model is able to reproduce a satellite luminosity function that is in agreement with observations, without running a full hydrodynamical simulation. In addition to determining the origin of circular streams in a Milky Way-like halo, we would also be able to predict the dynamical properties, ages and metallicities of the stars (within the accuracy of the semi-analytical model). This method has been very recently applied by Gómez et al. (2013), who study the phase-space distribution of stellar streams in the *Aquarius* simulations. These numerical predictions can then be compared with the properties of stellar streams in the Milky Way, including those that will be newly detected by the recently launched *Gaia* satellite (Perryman et al., 2001). *Gaia* will measure the positions and peculiar velocities of roughly one billion stars in the Milky Way with high accuracy, leading to a significant improvement in the identification of merger remnants (Gómez et al., 2010).

---

[1] The most desirable study, a combination of both statistical power and accuracy (potentially using full hydrodynamical simulations) would require a significant computational effort.

A Hierarchical SPT calculation

For the following calculations, it is convenient to rewrite the equations of motion from section 2.1 in terms of the conformal time τ, defined by

$$\mathrm{d}\tau = \frac{\mathrm{d}t}{a(t)}, \tag{A.1}$$

introducing the conformal Hubble parameter $\mathcal{H}(\tau) \equiv \mathrm{d}\ln a/\mathrm{d}\tau = Ha$. The non-linear Continuity and Euler Equation are then given by

$$\frac{\partial\delta(\boldsymbol{x},\tau)}{\partial\tau} + \theta(\boldsymbol{x},\tau) = -\nabla\cdot\left[\delta(\boldsymbol{x},\tau)\boldsymbol{u}(\boldsymbol{x},\tau)\right], \tag{A.2}$$

and

$$\frac{\partial\theta(\boldsymbol{x},\tau)}{\partial\tau} + \mathcal{H}(\tau)\theta(\boldsymbol{x},\tau) + \frac{3}{2}\,\Omega_{\mathrm{m}}\mathcal{H}^2(\tau)\delta(\boldsymbol{x},\tau) = -\nabla\cdot\Big\{\left[\boldsymbol{u}(\boldsymbol{x},\tau)\cdot\nabla\right]\boldsymbol{u}(\boldsymbol{x},\tau)\Big\}. \tag{A.3}$$

Here, $\theta(\boldsymbol{x},\tau)$ is the velocity divergence, which in the case of vanishing vorticity[1] reads

$$\theta(\boldsymbol{x},\tau) = \nabla\cdot\boldsymbol{u}(\boldsymbol{x},\tau), \tag{A.4}$$

and in Fourier space

$$\tilde{\boldsymbol{u}}(\boldsymbol{k},\tau) = \frac{\mathrm{i}\boldsymbol{k}}{k^2}\,\tilde{\theta}(\boldsymbol{k},\tau). \tag{A.5}$$

We will sketch the derivation of the mode-coupling kernel for the Continuity Equation, the Euler Equation can be treated completely analogously. Transforming (A.2) into Fourier space yields

$$\begin{aligned}
\frac{\partial\tilde{\delta}(\boldsymbol{k},\tau)}{\partial\tau} + \tilde{\theta}(\boldsymbol{k},\tau) &= -\int d^3x\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{x}}\,\nabla\cdot\left[\delta(\boldsymbol{x},\tau)\boldsymbol{u}(\boldsymbol{x},\tau)\right] \qquad \text{(A.6)}\\
&= -\int \mathrm{d}^3x \int\frac{\mathrm{d}^3k_1}{(2\pi)^3}\int\frac{\mathrm{d}^3k_2}{(2\pi)^3}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{x}}\sum_{j=1}^{3}\nabla_j\left[\tilde{u}_j(\boldsymbol{k}_1)\mathrm{e}^{-\mathrm{i}\boldsymbol{k}_1\boldsymbol{x}}\tilde{\delta}(\boldsymbol{k}_2)\mathrm{e}^{-\mathrm{i}\boldsymbol{k}_2\boldsymbol{x}}\right]\\
&= -\int \mathrm{d}^3x \int\frac{\mathrm{d}^3k_1}{(2\pi)^3}\int\frac{\mathrm{d}^3k_2}{(2\pi)^3}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{x}}\,\tilde{\delta}(\boldsymbol{k}_2)\sum_{j=1}^{3}\tilde{u}_j(\boldsymbol{k}_1)\nabla_j\left[\mathrm{e}^{-\mathrm{i}\boldsymbol{x}(\boldsymbol{k}_1+\boldsymbol{k}_2)}\right]\\
&= \int \mathrm{d}^3x \int\frac{\mathrm{d}^3k_1}{(2\pi)^3}\int\frac{\mathrm{d}^3k_2}{(2\pi)^3}\,\mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)\boldsymbol{x}}\,\tilde{\delta}(\boldsymbol{k}_2)\tilde{\boldsymbol{u}}(\boldsymbol{k}_1)\left[\mathrm{i}(\boldsymbol{k}_1+\boldsymbol{k}_2)\right]\\
&= \int\frac{\mathrm{d}^3k_1}{(2\pi)^3}\int \mathrm{d}^3k_2\,\delta_{\mathrm{D}}(\boldsymbol{k}-\boldsymbol{k}_1-\boldsymbol{k}_2)\,\tilde{\delta}(\boldsymbol{k}_2)\,\tilde{\boldsymbol{u}}(\boldsymbol{k}_1)\left[\mathrm{i}(\boldsymbol{k}_1+\boldsymbol{k}_2)\right]
\end{aligned}$$

---

[1] This assumption is justified by the single-stream approximation (section 2.1).

where we have temporarily dropped the τ-dependence in the intermediate steps, used a subscript j to denote the components of $\boldsymbol{u}$ and ∇, and inserted the definition of the Dirac delta function

$$(2\pi)^3\delta_{\mathrm{D}}(\boldsymbol{k}) = \int \mathrm{d}^3x\, \mathrm{e}^{\mathrm{i}\boldsymbol{kx}}, \tag{A.7}$$

in the final step. Replacing $\tilde{\boldsymbol{u}}$ by $\tilde{\theta}$ using Equation (A.5) and defining

$$\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{\boldsymbol{k}_1 \cdot (\boldsymbol{k}_1 + \boldsymbol{k}_2)}{k_1^2}, \tag{A.8}$$

we finally obtain

$$\begin{aligned} &\frac{\partial\tilde{\delta}(\boldsymbol{k},\tau)}{\partial\tau} + \tilde{\theta}(\boldsymbol{k},\tau) \\ &\quad = -\int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \mathrm{d}^3k_2\, \delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{k}_1 - \boldsymbol{k}_2)\, \tilde{\theta}(\boldsymbol{k}_1,\tau)\, \tilde{\delta}(\boldsymbol{k}_2,\tau)\, \alpha(\boldsymbol{k}_1,\boldsymbol{k}_2). \end{aligned} \tag{A.9}$$

The Euler Equation can be treated in the same way, resulting in

$$\begin{aligned} &\frac{\partial\tilde{\theta}(\boldsymbol{k},\tau)}{\partial\tau} + \mathcal{H}(\tau)\, \tilde{\theta}(\boldsymbol{k},\tau) + \frac{3}{2}\Omega_{\mathrm{m}}\mathcal{H}^2(\tau)\, \tilde{\delta}(\boldsymbol{k},\tau) \\ &\quad = -\int \frac{\mathrm{d}^3k_1}{(2\pi)^3} \int \mathrm{d}^3k_2\, \delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{k}_1 - \boldsymbol{k}_2)\, \tilde{\theta}(\boldsymbol{k}_1,\tau)\, \tilde{\theta}(\boldsymbol{k}_2,\tau)\, \beta(\boldsymbol{k}_1,\boldsymbol{k}_2), \end{aligned} \tag{A.10}$$

where the second mode-coupling kernel is given by

$$\beta(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{(\boldsymbol{k}_1 + \boldsymbol{k}_2)^2(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)}{2k_1^2k_2^2}. \tag{A.11}$$

A.1 Solution for an EdS universe

For the following derivation we will assume an EdS universe ($\Omega_{\mathrm{m}} = 1$, $\Omega_\Lambda = 0$); the extension to ΛCDM will be discussed below. We make the Ansatz that the density and velocity divergence can be written as separable functions:

$$\tilde{\delta}(\boldsymbol{k},\tau) = \sum_{n=1}^{\infty} a^n(\tau)\, \tilde{\delta}_n(\boldsymbol{k}), \tag{A.12}$$

$$\tilde{\theta}(\boldsymbol{k},\tau) = -\mathcal{H}(\tau) \sum_{n=1}^{\infty} a^n(\tau)\, \tilde{\theta}_n(\boldsymbol{k}), \tag{A.13}$$

For the following derivation, it is also convenient to introduce the short-hand notation $\delta_{\mathrm{D}}(\boldsymbol{k}, \boldsymbol{q}_n) \equiv \delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{q}_1 - \cdots - \boldsymbol{q}_n)$.

Inserting these expansions into the previously derived non-linear equations of motion will lead us to a recursion relation for $\tilde{\delta}_n$ and $\tilde{\theta}_n$. But first we will calculate a few time derivatives that will be needed later. In the EdS case, $a = \tau^2/\tau_0^2$ and

$\mathcal{H} = 2/\tau$, so

$$\frac{\partial a}{\partial \tau} = \frac{2\tau}{\tau_0^2} = \mathcal{H} a \qquad \text{and} \qquad \frac{\partial \mathcal{H}}{\partial \tau} = -\frac{1}{2}\mathcal{H}^2. \tag{A.14}$$

The time derivative of $\tilde{\delta}(\boldsymbol{k}, \tau)$ then reads

$$\frac{\partial \tilde{\delta}(\boldsymbol{k}, \tau)}{\partial \tau} = \sum_{n=1}^{\infty} n a^{n-1} \frac{\partial a}{\partial \tau} \tilde{\delta}_n(\boldsymbol{k}) = \mathcal{H} \sum_{n=1}^{\infty} n a^n \tilde{\delta}_n(\boldsymbol{k}). \tag{A.15}$$

The linearised Continuity equation – Equation (A.6) with the right-hand side set to 0 – then gives the linear solution $\tilde{\delta}_1(\boldsymbol{k}) = \tilde{\theta}_1(\boldsymbol{k})$.

Inserting the expansions (A.12), (A.13), and the time derivative of $\tilde{\delta}(\boldsymbol{k}, \tau)$ into the Continuity equation (A.9), and dividing by a common factor $\mathcal{H}$, we obtain

$$\begin{aligned}
&\sum_{n=1}^{\infty} n a^n \tilde{\delta}_n(\boldsymbol{k}) - \sum_{n=1}^{\infty} a^n \tilde{\theta}_n(\boldsymbol{k}) \\
&\quad = \int \frac{\mathrm{d}^3 k_1}{(2\pi)^3} \int \mathrm{d}^3 k_2 \, \delta_{\mathrm{D}}(\boldsymbol{k}, \boldsymbol{k}_2) \, \alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) \sum_{n=1}^{\infty} a^n \tilde{\theta}_n(\boldsymbol{k}_1) \sum_{m=1}^{\infty} a^m \tilde{\delta}_m(\boldsymbol{k}_2) \\
&\quad = \int \frac{\mathrm{d}^3 k_1}{(2\pi)^3} \int \mathrm{d}^3 k_2 \, \delta_{\mathrm{D}}(\boldsymbol{k}, \boldsymbol{k}_2) \, \alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) \sum_{n=2}^{\infty} \sum_{m=1}^{n-1} a^n \tilde{\theta}_{n-m}(\boldsymbol{k}_1) \, \tilde{\delta}_m(\boldsymbol{k}_2),
\end{aligned} \tag{A.16}$$

where we note that the second equality is only applicable for $n > 1$, but we already know the linear solution. Comparing the left and right-hand sides we obtain the first recursion relation

$$n \tilde{\delta}_n(\boldsymbol{k}) - \tilde{\theta}_n(\boldsymbol{k}) = \int \frac{\mathrm{d}^3 k_1}{(2\pi)^3} \int \mathrm{d}^3 k_2 \, \delta_{\mathrm{D}}(\boldsymbol{k}, \boldsymbol{k}_2) \, \alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) \sum_{m=1}^{n-1} \tilde{\theta}_{n-m}(\boldsymbol{k}_1) \, \tilde{\delta}_m(\boldsymbol{k}_2) \tag{A.17}$$

For the Euler equation, we need to calculate the time derivative of $\tilde{\theta}(\boldsymbol{k}, \tau)$,

$$\begin{aligned}
\frac{\partial \tilde{\theta}(\boldsymbol{k}, \tau)}{\partial \tau} &= -\frac{\partial \mathcal{H}(\tau)}{\partial \tau} \sum_{n=1}^{\infty} a^n \tilde{\theta}_n(\boldsymbol{k}) - \mathcal{H}(\tau) \sum_{n=1}^{\infty} n a^{n-1} \frac{\partial a}{\partial \tau} \tilde{\theta}_n(\boldsymbol{k}) \\
&= \frac{1}{2} \mathcal{H}^2 \sum_{n=1}^{\infty} a^n \tilde{\theta}_n(\boldsymbol{k}) - \mathcal{H}^2 a \sum_{n=1}^{\infty} n a^{n-1} \tilde{\theta}_n(\boldsymbol{k}) \\
&= -\frac{1}{2} \mathcal{H}^2 \sum_{n=1}^{\infty} (2n - 1) \, a^n \tilde{\theta}_n(\boldsymbol{k}).
\end{aligned} \tag{A.18}$$

Inserting this into the Euler equation, and dividing by a common factor of $-\mathcal{H}^2$, we arrive at:

$$\begin{aligned}
&(2n + 1) \, \tilde{\theta}_n(\boldsymbol{k}) - 3 \, \tilde{\delta}_n(\boldsymbol{k}) \\
&\quad = 2 \int \frac{\mathrm{d}^3 k_1}{(2\pi)^3} \int \mathrm{d}^3 k_2 \, \delta_{\mathrm{D}}(\boldsymbol{k}, \boldsymbol{k}_2) \, \beta(\boldsymbol{k}_1, \boldsymbol{k}_2) \sum_{m=1}^{n-1} \tilde{\theta}_{n-m}(\boldsymbol{k}_1) \, \tilde{\theta}_m(\boldsymbol{k}_2)
\end{aligned} \tag{A.19}$$

Calling the right-hand side of Equations (A.17) and (A.19) $R_1^{(n)}$ and $R_2^{(n)}$, respectively, we can solve for $\tilde{\delta}_n$ and $\tilde{\theta}_n$

$$\tilde{\delta}_n(\boldsymbol{k}) = \frac{1}{(2n+3)(n-1)} \left[(2n+1) R_1^{(n)}(\boldsymbol{k}, \boldsymbol{k}_1, \boldsymbol{k}_2) + R_2^{(n)}(\boldsymbol{k}, \boldsymbol{k}_1, \boldsymbol{k}_2)\right] \tag{A.20}$$

$$\tilde{\theta}_n(\boldsymbol{k}) = \frac{1}{(2n+3)(n-1)} \left[3 R_1^{(n)}(\boldsymbol{k}, \boldsymbol{k}_1, \boldsymbol{k}_2) + n R_2^{(n)}(\boldsymbol{k}, \boldsymbol{k}_1, \boldsymbol{k}_2)\right] \tag{A.21}$$

One then makes the Ansatz (e. g. Goroff et al., 1986)

$$\tilde{\delta}_n(\boldsymbol{k}) = \int \mathrm{d}^3 q_1 \cdots \int \mathrm{d}^3 q_n \, \delta_{\mathrm{D}}(\boldsymbol{k}, \boldsymbol{q}_n) \, F_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) \, \tilde{\delta}_1(\boldsymbol{q}_1) \cdots \tilde{\delta}_1(\boldsymbol{q}_n) \tag{A.22}$$

$$\tilde{\theta}_n(\boldsymbol{k}) = \int \mathrm{d}^3 q_1 \cdots \int \mathrm{d}^3 q_n \, \delta_{\mathrm{D}}(\boldsymbol{k}, \boldsymbol{q}_n) \, G_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) \, \tilde{\delta}_1(\boldsymbol{q}_1) \cdots \tilde{\delta}_1(\boldsymbol{q}_n), \tag{A.23}$$

which describes the n-th order solution as n convolution of the linear solution with kernels F_n and G_n. Note that we already used the previously derived initial solution to replace $\tilde{\theta}_1$ by $\tilde{\delta}_1$.

Inserting this Ansatz into the right-hand side of Equations (A.20) and (A.21), and setting them equal to the right-hand side of Equations (A.22) and (A.23) leads to rather lengthy expressions which are however rather straightforward to simplify by integrating out the Dirac delta function in $R_i^{(n)}$. This finally leads to another set of recursion relations for F_n and G_n

$$\begin{aligned} F_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_m)}{(2n+3)(n-1)} \Big[& (2n+1) \alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) F_{n-m}(\boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n) \\ & + 2\, \beta(\boldsymbol{k}_1, \boldsymbol{k}_2) G_{n-m}(\boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n))\Big] \end{aligned} \tag{A.24}$$

$$\begin{aligned} G_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_m)}{(2n+3)(n-1)} \Big[& 3 \alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) F_{n-m}(\boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n) \\ & + 2n \beta(\boldsymbol{k}_1, \boldsymbol{k}_2) G_{\mathrm{n-m}}(\boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n)\Big] \end{aligned} \tag{A.25}$$

with $F_1 \equiv G_1 \equiv 1$, $\boldsymbol{k}_1 = \boldsymbol{q}_1 + \ldots + \boldsymbol{q}_m$ and $\boldsymbol{k}_2 = \boldsymbol{q}_{m+1} + \ldots + \boldsymbol{q}_n$.

Before going further, it is important to consider how $R_1^{(n)}$ and $R_2^{(n)}$ behave under the exchange of $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$, which is determined by α and β. We can see from Equations (A.8) and (A.11) that $R_2^{(n)}$ is symmetric, but $R_1^{(n)}$ is not. This means that if we want to integrate out the Dirac delta function that appears in $R_i^{(n)}$, we have to symmetrise the terms containing α. This then leaves us with expressions for δ_n and θ_n which require $n-1$ 3D integrals. This form is useful because it relates everything back to the linear density field, and higher order power- and cross-spectra can be related to the linear power spectrum that is easy to compute (e. g. Heavens et al., 1998). However, it is not advantageous for the point-by-point computation performed in Chapter 4.

An alternative solution can be found by directly considering Equations (A.20) and (A.21) (Juszkiewicz, 1981; Vishniac, 1983; Makino et al., 1992). Integrating over the Dirac delta function in $R_i^{(n)}$ and symmetrising the terms with $R_1^{(n)}$

immediately leads to

$$\begin{aligned}\tilde{\delta}_n(\boldsymbol{k}) = \frac{1}{(2n+3)(n-1)} \int \frac{\mathrm{d}^3 k_1}{(2\pi)^3} \Bigg\{ \frac{(2n+1)}{2} \Bigg[\sum_{m=1}^{n-1} \tilde{\theta}_{n-m}(\boldsymbol{k}_1)\, \tilde{\delta}_m(\boldsymbol{k}-\boldsymbol{k}_1) \\ \alpha(\boldsymbol{k}_1, \boldsymbol{k}-\boldsymbol{k}_1) + \sum_{m=1}^{n-1} \tilde{\theta}_{n-m}(\boldsymbol{k}-\boldsymbol{k}_1)\, \tilde{\delta}_m(\boldsymbol{k}_1)\, \alpha(\boldsymbol{k}-\boldsymbol{k}_1, \boldsymbol{k}_1) \Bigg] \\ +2\,\beta(\boldsymbol{k}_1, \boldsymbol{k}-\boldsymbol{k}_1) \sum_{m=1}^{n-1} \tilde{\theta}_{n-m}(\boldsymbol{k}_1)\, \tilde{\theta}_m(\boldsymbol{k}-\boldsymbol{k}_1) \Bigg\} \qquad \text{(A.26)}\end{aligned}$$

$$\begin{aligned}\tilde{\theta}_n(\boldsymbol{k}) = \frac{1}{(2n+3)(n-1)} \int \frac{\mathrm{d}^3 k_1}{(2\pi)^3} \Bigg\{ \frac{3}{2} \Bigg[\sum_{m=1}^{n-1} \tilde{\theta}_{n-m}(\boldsymbol{k}_1)\, \tilde{\delta}_m(\boldsymbol{k}-\boldsymbol{k}_1) \\ \alpha(\boldsymbol{k}_1, \boldsymbol{k}-\boldsymbol{k}_1) + \sum_{m=1}^{n-1} \tilde{\theta}_{n-m}(\boldsymbol{k}-\boldsymbol{k}_1)\, \tilde{\delta}_m(\boldsymbol{k}_1)\, \alpha(\boldsymbol{k}-\boldsymbol{k}_1, \boldsymbol{k}_1) \Bigg] \\ +2n\beta(\boldsymbol{k}_1, \boldsymbol{k}-\boldsymbol{k}_1) \sum_{m=1}^{n-1} \tilde{\theta}_{n-m}(\boldsymbol{k}_1)\, \tilde{\theta}_m(\boldsymbol{k}-\boldsymbol{k}_1) \Bigg\} \qquad \text{(A.27)}\end{aligned}$$

which is equivalent to Equations (2.15a) and (2.15b) in Makino et al. (1992) with $H(\boldsymbol{k}-\boldsymbol{p}) = \alpha(\boldsymbol{k}-\boldsymbol{p}, \boldsymbol{p})$ and $F = \beta$. Note that there is a small typo in their Equation (2.15a), it should say $F(\boldsymbol{p}, \boldsymbol{k}-\boldsymbol{p})$ instead of $F(\boldsymbol{p}, \boldsymbol{p}-\boldsymbol{k})$.

This solution is much more suitable for a direct, numerical calculation of the third-order density field as done in Chapter 4, because it only involves one 3D integral, and the mode-coupling kernels are always the same. It will thus take roughly the same time to calculate each order in the expansion, because the solution is based on the previous orders, while this fact is not used in the previous method. Of course this also means that even if we are only interested in δ_n we also need to calculate θ_{n-1}. However this is actually an advantage, as studying the velocity divergence is relevant e. g. in the context of redshift-space distortions (Kaiser, 1987; Hawkins et al., 2003).

A.2 Solution for a flat $\Lambda\mathrm{CDM}$ universe

The derivation above is technically only correct in the case of an EdS universe. However, it can be easily extended even in the case of a non-zero cosmological constant. We start by replacing the scale factor $a(t)$ for the time evolution of the density contrast by the growth factor $D(t)$. The time evolution for the velocity divergence then immediately follows from the linearised Continuity equation, and we obtain

$$\tilde{\delta}(\boldsymbol{k}, \tau) = D(\tau)\tilde{\delta}_1(\boldsymbol{k}), \qquad \text{(A.28)}$$

$$\tilde{\theta}(\boldsymbol{k}, \tau) = -\mathcal{H} f(\Omega_\mathrm{m}, \Omega_\Lambda) D(\tau)\tilde{\delta}_1(\boldsymbol{k}), \qquad \text{(A.29)}$$

where

$$f(\Omega_\mathrm{m}, \Omega_\Lambda) \equiv \frac{\mathrm{d}\ln D}{\mathrm{d}\ln a} = \frac{1}{\mathcal{H}} \frac{\mathrm{d}\ln D}{\mathrm{d}\tau} \qquad \text{(A.30)}$$

encodes the difference of growth between a ΛCDM and an EdS universe. Based on the similarities to Equations (A.12) and (A.13), it is sensible to make the Ansatz

$$\tilde{\delta}(\boldsymbol{k},\tau) = \sum_{n=1}^{\infty} D_n(\tau)\tilde{\delta}_n(\boldsymbol{k}), \tag{A.31}$$

$$\tilde{\theta}(\boldsymbol{k},\tau) = -\mathcal{H} f(\Omega_\mathrm{m},\Omega_\Lambda) \sum_{n=1}^{\infty} E_n(\tau)\tilde{\theta}_n(\boldsymbol{k}) \tag{A.32}$$

where $D_n(\tau)$ and $E_n(\tau)$ have to be determined. By insertion, it can be shown that this Ansatz is a solution for the non-linear equations of motions with $D_n \approx E_n \approx D^n$, as long as $f(\Omega_\mathrm{m},\Omega_\Lambda) \approx \Omega_\mathrm{m}^{0.5}$, and the latter approximation is highly accurate for flat ΛCDM models independent of Ω_m (Scoccimarro et al. 1998, also shown numerically by Bernardeau 1994a).

The mode-coupling kernels are also slightly affected because this new time dependence is not an exact power law (Bernardeau, 1994a), but the differences are only of percent order even at redshift 0 where the deviation from EdS is largest, so this effect is usually neglected.

References

Acquaviva V., Bartolo N., Matarrese S., Riotto A., 2003, Nuclear Physics B, 667, 119, `astro-ph/0209156`

Albrecht A., Steinhardt P. J., 1982, Physical Review Letters, 48, 1220

Allen S. W., Evrard A. E., Mantz A. B., 2011, Annu. Rev. Astron. Astrophys., 49, 409, `arXiv:1103.4829`

Alpher R. A., Bethe H., Gamow G., 1948, Phys. Rev., 73, 803

ATLAS Collaboration et al., 2012, Physics Letters B, 716, 1, `arXiv:1207.7214`

Baldauf T., Seljak U., Senatore L., 2011, JCAP, 4, 6, `arXiv:1011.1513`

Bardeen J. M., 1980, Phys. Rev. D, 22, 1882

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

Barnes J., Hut P., 1986, Nature, 324, 446

Barnes J. E., Hernquist L., 1992, Nature, 360, 715

Bassett B. A., Kunz M., 2004, Phys. Rev. D, 69, 101305, `astro-ph/0312443`

Bassett B. A., Tsujikawa S., Wands D., 2006, Reviews of Modern Physics, 78, 537, `astro-ph/0507632`

Baudis L., 2012, Physics of the Dark Universe, 1, 94, `arXiv:1211.7222`

Baumann D., 2009, ArXiv e-prints, `arXiv:0907.5424`

Becker A., Huterer D., Kadota K., 2012, JCAP, 12, 34, `arXiv:1206.6165`

Belokurov V., 2013, ArXiv e-prints, `arXiv:1307.0041`

Bennett C. L. et al., 2012, ArXiv e-prints, `arXiv:1212.5225`

Benson A. J., Cole S., Frenk C. S., Baugh C. M., Lacey C. G., 2000, MNRAS, 311, 793, `astro-ph/9903343`

Bernabei R. et al., 2008, European Physical Journal C, 56, 333, `arXiv:0804.2741`

Bernardeau F., 1994a, ApJ, 433, 1, `astro-ph/9312026`

Bernardeau F., 1994b, A & A, 291, 697, `astro-ph/9403020`

Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1, `astro-ph/0112551`

Bernardeau F., Kofman L., 1995, ApJ, 443, 479, `astro-ph/9403028`

Bertschinger E., 1985, ApJS, 58, 39

Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122, `arXiv:1005.2239`

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

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

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645, `astro-ph/0511338`

Boylan-Kolchin M., Bullock J. S., Sohn S. T., Besla G., van der Marel R. P., 2013, ApJ, 768, 140, `arXiv:1210.6046`

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150, `arXiv:0903.3041`

Branch D., Tammann G. A., 1992, Annu. Rev. Astron. Astrophys., 30, 359

Buchalter A., Kamionkowski M., 1999, ApJ, 521, 1, `astro-ph/9903462`

Bullock J. S., Johnston K. V., 2005, ApJ, 635, 931, `astro-ph/0506467`

Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559, `astro-ph/9908159`

Bunch T. S., Davies P. C. W., 1978, Royal Society of London Proceedings Series A, 360, 117

Cai Y., Bernstein G., Sheth R. K., 2011, MNRAS, 412, 995, `arXiv:1007.3500`

Carlin J. L., Casetti-Dinescu D. I., Grillmair C. J., Majewski S. R., Girard T. M., 2010, ApJ, 725, 2290, `arXiv:1010.5257`

Carlson J., White M., Padmanabhan N., 2009, Phys. Rev. D, 80, 043531, `arXiv:0905.0479`

Carrasco J. J. M., Foreman S., Green D., Senatore L., 2013, ArXiv e-prints, `arXiv:1310.0464`

Catelan P., Lucchin F., Matarrese S., Porciani C., 1998, MNRAS, 297, 692, `astro-ph/9708067`

Catelan P., Porciani C., Kamionkowski M., 2000, MNRAS, 318, L39, `astro-ph/0005544`

Chandrasekhar S., 1943, ApJ, 97, 255

Chisari N. E., Zaldarriaga M., 2011, Phys. Rev. D, 83, 123505, `arXiv:1101.3555`

Christensen S. M., 1976, Phys. Rev. D, 14, 2490

Clifton T., Ferreira P. G., Padilla A., Skordis C., 2012, Phys. Rep., 513, 1, `arXiv:1106.2476`

Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781, `astro-ph/9402001`

Cole S., Kaiser N., 1989, MNRAS, 237, 1127

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168, `astro-ph/0007281`

Cole S. et al., 2005, MNRAS, 362, 505, `astro-ph/0501174`

Coles P., Jones B., 1991, MNRAS, 248, 1

Conn A. R. et al., 2013, ApJ, 766, 120, `arXiv:1301.7131`

Conn B. C. et al., 2007, MNRAS, 376, 939, `astro-ph/0701664`

Conn B. C., Lane R. R., Lewis G. F., Irwin M. J., Ibata R. A., Martin N. F., Bellazzini M., Tuntsov A. V., 2008, MNRAS, 390, 1388

Conn B. C., Lewis G. F., Irwin M. J., Ibata R. A., Ferguson A. M. N., Tanvir N., Irwin J. M., 2005a, MNRAS, 362, 475, `astro-ph/0506522`

Conn B. C., Martin N. F., Lewis G. F., Ibata R. A., Bellazzini M., Irwin M. J., 2005b, MNRAS, 364, L13, `astro-ph/0508366`

Conn B. C. et al., 2012, ApJ, 754, 101, `arXiv:1205.3177`

Cooper A. P. et al., 2010, MNRAS, 406, 744, `arXiv:0910.3211`

Crane J. D., Majewski S. R., Rocha-Pinto H. J., Frinchaboy P. M., Skrutskie M. F., Law D. R., 2003, ApJL, 594, L119, `astro-ph/0307505`

Creminelli P., Nicolis A., Senatore L., Tegmark M., Zaldarriaga M., 2006, JCAP, 5, 4, `astro-ph/0509029`

Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369, `astro-ph/0606505`

Crocce M., Scoccimarro R., 2006, Phys. Rev. D, 73, 063519, `astro-ph/0509418`

Dalal N., Doré O., Huterer D., Shirokov A., 2008, Phys. Rev. D, 77, 123514, `arXiv:0710.4560`

Dalcanton J. J. et al., 2009, ApJS, 183, 67, `arXiv:0905.3737`

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

de Blok W. J. G., 2010, Advances in Astronomy, 2010, `arXiv:0910.3538`

de Jong J. T. A., Yanny B., Rix H.-W., Dolphin A. E., Martin N. F., Beers T. C., 2010, ApJ, 714, 663, `arXiv:0911.3900`

Dehnen W., 2001, MNRAS, 324, 273, `astro-ph/0011568`

Dekel A., Lahav O., 1999, ApJ, 520, 24, `astro-ph/9806193`

Dekel A., Rees M. J., 1987, Nature, 326, 455

Desjacques V., Seljak U., 2010a, Classical and Quantum Gravity, 27, 124011, `arXiv:1003.5020`

Desjacques V., Seljak U., 2010b, Phys. Rev. D, 81, 023006, `arXiv:0907.2257`

Dicke R. H., Peebles P. J. E., Roll P. G., Wilkinson D. T., 1965, ApJ, 142, 414

Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859, `astro-ph/0703337`

Dinescu D. I., Martínez-Delgado D., Girard T. M., Peñarrubia J., Rix H.-W., Butler D., van Altena W. F., 2005, ApJL, 631, L49, `astro-ph/0508247`

Dodelson S., 2003, Modern cosmology, Academic Press

Efstathiou G., Frenk C. S., White S. D. M., Davis M., 1988, MNRAS, 235, 715

Efstathiou G., Sutherland W. J., Maddox S. J., 1990, Nature, 348, 705

Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605, `astro-ph/9709112`

Eisenstein D. J. et al., 2005, ApJ, 633, 560, `astro-ph/0501171`

Elia A., Kulkarni S., Porciani C., Pietroni M., Matarrese S., 2011, MNRAS, 416, 1703, `arXiv:1012.4833`

Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23, `astro-ph/9304022`

Ferrell R., Bertschinger E., 1994, International Journal of Modern Physics C, 5, 933

Fillmore J. A., Goldreich P., 1984, ApJ, 281, 1

Flender S. F., Schwarz D. J., 2012, Phys. Rev. D, 86, 063527, `arXiv:1207.2035`

Freeman K., Bland-Hawthorn J., 2002, Annu. Rev. Astron. Astrophys., 40, 487, `astro-ph/0208106`

Frenk C. S., White S. D. M., 2012, Annalen der Physik, 524, 507, `arXiv:1210.0544`

Frigo M., Johnson S. G., 2005, Proceedings of the IEEE, 93, 216

Fry J. N., 1984, Astrophys. J., 279, 499

Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447, `astro-ph/9302009`

Gangui A., Lucchin F., Matarrese S., Mollerach S., 1994, ApJ, 430, 447, `astro-ph/9312033`

Gao L., Navarro J. F., Cole S., Frenk C. S., White S. D. M., Springel V., Jenkins A., Neto A. F., 2008, MNRAS, 387, 536, `arXiv:0711.0746`

Gao L., Springel V., White S. D. M., 2005, MNRAS, 363, L66, `astro-ph/0506510`

Gaztanaga E., Frieman J. A., 1994, ApJL, 437, L13, `astro-ph/9407079`

Geen S., Slyz A., Devriendt J., 2013, MNRAS, 429, 633, `arXiv:1204.3327`

Giannantonio T., Porciani C., 2010, Phys. Rev. D, 81, 63530, `arXiv:0911.0017`

Giannantonio T., Porciani C., Carron J., Amara A., Pillepich A., 2012, MNRAS, 422, 2854, `arXiv:1109.0958`

Giannantonio T., Ross A. J., Percival W. J., Crittenden R., Bacher D., Kilbinger M., Nichol R., Weller J., 2014, Phys. Rev. D, 89, 023511, `arXiv:1303.1349`

Gómez F. A., Helmi A., Brown A. G. A., Li Y.-S., 2010, MNRAS, 408, 935, `arXiv:1004.4974`

Gómez F. A., Helmi A., Cooper A. P., Frenk C. S., Navarro J. F., White S. D. M., 2013, ArXiv e-prints, `arXiv:1307.0008`

Goroff M. H., Grinstein B., Rey S., Wise M. B., 1986, ApJ, 311, 6

Graham A. W., Guzmán R., 2003, AJ, 125, 2936, `astro-ph/0303391`

Grebel E. K., 2001, Astrophysics and Space Science Supplement, 277, 231, `astro-ph/0011048`

Green S. R., Wald R. M., 2012, Phys. Rev. D, 85, 063512, `arXiv:1111.2997`

Grillmair C. J., 2006, ApJL, 651, L29, `astro-ph/0609449`

Gunn J. E., Gott, III J. R., 1972, ApJ, 176, 1

Guth A. H., 1981, Phys. Rev. D, 23, 347

Hamilton A. J. S., 1998, in Astrophysics and Space Science Library, Vol. 231, The Evolving Universe, Hamilton D., ed., p. 185

Harrison E. R., 1970, Phys. Rev. D, 1, 2726

Hawking S. W., Moss I. G., Stewart J. M., 1982, Phys. Rev. D, 26, 2681

Hawkins E. et al., 2003, MNRAS, 346, 78, `astro-ph/0212375`

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

Heavens A. F., Matarrese S., Verde L., 1998, MNRAS, 301, 797, `astro-ph/9808016`

Hinshaw G. et al., 2012, ArXiv e-prints, `arXiv:1212.5226`

Hockney R. W., Eastwood J. W., 1988, Computer simulation using particles, Hilger, Bristol

Ibata R., Irwin M., Lewis G., Ferguson A. M. N., Tanvir N., 2001a, Nature, 412, 49, `astro-ph/0107090`

Ibata R., Irwin M., Lewis G. F., Stolte A., 2001b, ApJL, 547, L133, `astro-ph/0004255`

Ibata R. A., Irwin M. J., Lewis G. F., Ferguson A. M. N., Tanvir N., 2003, MNRAS, 340, L21, `astro-ph/0301067`

Ibata R. A. et al., 2013, Nature, 493, 62, `arXiv:1301.0446`

Ivezić Ž. et al., 2008a, ApJ, 684, 287, `arXiv:0804.3850`

Ivezić Ž. et al., 2008b, ArXiv e-prints, `arXiv:0805.2366`

Izotov Y. I., Thuan T. X., Stasińska G., 2007, ApJ, 662, 15, `astro-ph/0702072`

Jain B., Bertschinger E., 1994, ApJ, 431, 495, `astro-ph/9311070`

Jeffreys H., 1961, The Theory of Probability, 3rd edn., Oxford University Press

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372, `astro-ph/0005260`

Jeong D., Komatsu E., 2006, ApJ, 651, 619, `astro-ph/0604075`

Jeong D., Komatsu E., 2007, in Bulletin of the American Astronomical Society, Vol. 39, American Astronomical Society Meeting Abstracts, p. 885

Jeong D., Komatsu E., 2009a, ApJ, 691, 569, `arXiv:0805.2632`

Jeong D., Komatsu E., 2009b, ApJ, 703, 1230, `arXiv:0904.0497`

Jurić M. et al., 2008, ApJ, 673, 864, `astro-ph/0510520`

Juszkiewicz R., 1981, MNRAS, 197, 931

Kaiser N., 1984, ApJL, 284, L9

Kaiser N., 1987, MNRAS, 227, 1

Kass R. E., 1993, The Statistician, 42, 551

Katz N., White S. D. M., 1993, ApJ, 412, 455

Kauffmann G., Nusser A., Steinmetz M., 1997, MNRAS, 286, 795, `astro-ph/9512009`

Kayo I., Taruya A., Suto Y., 2001, ApJ, 561, 22, `astro-ph/0105218`

Kazantzidis S., Bullock J. S., Zentner A. R., Kravtsov A. V., Moustakas L. A., 2008, ApJ, 688, 254, `arXiv:0708.1949`

Kim J., Park C., Rossi G., Lee S. M., Gott, III J. R., 2011, Journal of Korean Astronomical Society, 44, 217, `arXiv:1112.1754`

Knebe A. et al., 2011, MNRAS, 415, 2293, `arXiv:1104.0949`

Knebe A. et al., 2013, ArXiv e-prints, `arXiv:1304.0585`

Knobel C., 2012, ArXiv e-prints, `arXiv:1208.5931`

Kodama H., Sasaki M., 1984, Progress of Theoretical Physics Supplement, 78, 1

Komatsu E. et al., 2009, ApJS, 180, 330, `arXiv:0803.0547`

Komatsu E. et al., 2011, ApJS, 192, 18, `arXiv:1001.4538`

Kopp J., Schwetz T., Zupan J., 2010, JCAP, 1002, 014, `arXiv:0912.4264`

Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, ApJS, 182, 216, `arXiv:0810.1681`

Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35, `astro-ph/0308519`

Kroupa P., Theis C., Boily C. M., 2005, A & A, 431, 517, `astro-ph/0410421`

Kuhlen M., Vogelsberger M., Angulo R., 2012, Physics of the Dark Universe, 1, 50, `arXiv:1209.5745`

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

Laureijs R. et al., 2011, ArXiv e-prints, `arXiv:1110.3193`

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473, `astro-ph/9911177`

Li Y.-S., White S. D. M., 2008, MNRAS, 384, 1459, `arXiv:0710.3740`

Libeskind N. I., Knebe A., Hoffman Y., Gottlöber S., Yepes G., Steinmetz M., 2011, MNRAS, 411, 1525, `arXiv:1010.1531`

Liddle A. R., Lyth D. H., 1993, Phys. Rep., 231, 1, `astro-ph/9303019`

Linde A., Mukhanov V., 1997, Phys. Rev. D, 56, R535

Linde A. D., 1982, Physics Letters B, 108, 389

Lovell M. R., Eke V. R., Frenk C. S., Jenkins A., 2011, MNRAS, 413, 3013, `arXiv:1008.0484`

Ludlow A. D. et al., 2013, MNRAS, 432, 1103, `arXiv:1302.0288`

Ludlow A. D., Navarro J. F., Li M., Angulo R. E., Boylan-Kolchin M., Bett P. E., 2012, MNRAS, 427, 1322, `arXiv:1206.1049`

Ludlow A. D., Navarro J. F., Springel V., Jenkins A., Frenk C. S., Helmi A., 2009, ApJ, 692, 931, `arXiv:0801.1127`

Lynden-Bell D., 1967, MNRAS, 136, 101

Lynden-Bell D., 1976, MNRAS, 174, 695

Lyth D. H., Ungarelli C., Wands D., 2003, Phys. Rev. D, 67, 023503, `astro-ph/0208055`

Ma C.-P., Bertschinger E., 1995, ApJ, 455, 7, `astro-ph/9506072`

Maddox S. J., Efstathiou G., Sutherland W. J., 1996, MNRAS, 283, 1227, `astro-ph/9601103`

Maggiore M., Riotto A., 2010, ApJ, 717, 526, `arXiv:0903.1251`

Makino N., Sasaki M., Suto Y., 1992, Phys. Rev. D, 46, 585

Maldacena J., 2003, Journal of High Energy Physics, 5, 13, `astro-ph/0210603`

Manera M., Gaztañaga E., 2011, MNRAS, 415, 383, `arXiv:0912.0446`

Manera M., Sheth R. K., Scoccimarro R., 2010, MNRAS, 402, 589, `arXiv:0906.1314`

Martin N. F., Ibata R. A., Bellazzini M., Irwin M. J., Lewis G. F., Dehnen W., 2004, MNRAS, 348, 12, `astro-ph/0311010`

Martin N. F., Ibata R. A., Conn B. C., Lewis G. F., Bellazzini M., Irwin M. J., 2005, MNRAS, 362, 906, `astro-ph/0503705`

Martin N. F., Irwin M. J., Ibata R. A., Conn B. C., Lewis G. F., Bellazzini M., Chapman S., Tanvir N., 2006, MNRAS, 367, L69, `astro-ph/0601176`

Martínez-Delgado D., Peñarrubia J., Gabany R. J., Trujillo I., Majewski S. R., Pohlen M., 2008, ApJ, 689, 184, `arXiv:0805.1137`

Matarrese S., Verde L., 2008, ApJL, 677, L77, `arXiv:0801.4826`

Matarrese S., Verde L., Heavens A. F., 1997, MNRAS, 290, 651, `astro-ph/9706059`

Mateo M., Mirabal N., Udalski A., Szymanski M., Kaluzny J., Kubiak M., Krzeminski W., Stanek K. Z., 1996, ApJL, 458, L13

Mathewson D. S., Cleary M. N., Murray J. D., 1974, ApJ, 190, 291

McConnachie A. W., 2012, AJ, 144, 4, `arXiv:1204.1562`

McConnachie A. W. et al., 2009, Nature, 461, 66, `arXiv:0909.0398`

McDonald P., 2006, Phys. Rev. D, 74, 103512, `astro-ph/0609413`

McMillan S. L. W., Aarseth S. J., 1993, ApJ, 414, 200

Meerburg P. D., van der Schaar J. P., Corasaniti P. S., 2009, JCAP, 5, 18, `arXiv:0901.4044`

Metz M., Kroupa P., Libeskind N. I., 2008, ApJ, 680, 287, `arXiv:0802.3899`

Michel-Dansac L., Abadi M. G., Navarro J. F., Steinmetz M., 2011, MNRAS, 414, L1, `arXiv:1102.4463`

Mirabel I. F., Dottori H., Lutz D., 1992, A & A, 256, L19

Misner C. W., Thorne K. S., Wheeler J. A., 1973, Gravitation, W.H. Freeman and Co., San Francisco

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347, `astro-ph/9512127`

Moitinho A., Vázquez R. A., Carraro G., Baume G., Giorgi E. E., Lyra W., 2006, MNRAS, 368, L77, `astro-ph/0602493`

Momany Y., Zaggia S., Gilmore G., Piotto G., Carraro G., Bedin L. R., de Angeli F., 2006, A & A, 451, 515, `astro-ph/0603385`

Mukhanov V. F., Feldman H. A., Brandenberger R. H., 1992, Phys. Rep., 215, 203

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493, `astro-ph/9611107`

Navarro J. F. et al., 2010, MNRAS, 402, 21, `arXiv:0810.1522`

Newberg H. J. et al., 2002, ApJ, 569, 245, `astro-ph/0111095`

Niederste-Ostholt M., Belokurov V., Evans N. W., Peñarrubia J., 2010, ApJ, 712, 516, `arXiv:1002.0266`

Nishimichi T. et al., 2009, PASJ, 61, 321, `arXiv:0810.0813`

Oh S.-H., Brook C., Governato F., Brinks E., Mayer L., de Blok W. J. G., Brooks A., Walter F., 2011, AJ, 142, 24, `arXiv:1011.2777`

Okamoto T., Frenk C. S., 2009, MNRAS, 399, L174, `arXiv:0909.0262`

Pawlowski M. S., Pflamm-Altenburg J., Kroupa P., 2012, MNRAS, 423, 1109, `arXiv:1204.5176`

Peñarrubia J. et al., 2005, ApJ, 626, 128, `astro-ph/0410448`

Peebles P. J. E., 1967, ApJ, 147, 859

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

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Percival W. J. et al., 2001, MNRAS, 327, 1297, `astro-ph/0105252`

Perlmutter S. et al., 1999, ApJ, 517, 565, `astro-ph/9812133`

Perryman M. A. C. et al., 2001, A & A, 369, 339, `astro-ph/0101235`

Pietroni M., 2008, JCAP, 10, 36, `arXiv:0806.0971`

Pillepich A., Porciani C., Hahn O., 2010, MNRAS, 402, 191, `arXiv:0811.4176`

Pillepich A., Porciani C., Reiprich T. H., 2012, MNRAS, 422, 44, `arXiv:1111.6587`

Planck Collaboration et al., 2013a, ArXiv e-prints, `arXiv:1303.5076`

Planck Collaboration et al., 2013b, ArXiv e-prints, `arXiv:1303.5084`

Porciani C., Norberg P., 2006, MNRAS, 371, 1824, `astro-ph/0607348`

Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14, `astro-ph/0201544`

Predehl P. et al., 2010, in SPIE Conference Series, Vol. 7732, SPIE Conference Series

Preskill J. P., 1979, Physical Review Letters, 43, 1365

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

Regan D. M., Schmittfull M. M., Shellard E. P. S., Fergusson J. R., 2012, Phys. Rev. D, 86, 123524, `arXiv:1108.3813`

Riess A. G. et al., 1998, AJ, 116, 1009, `astro-ph/9805201`

Rocha M., Peter A. H. G., Bullock J., 2012, MNRAS, 425, 231, `arXiv:1110.0464`

Rocha-Pinto H. J., Majewski S. R., Skrutskie M. F., Crane J. D., 2003, ApJL, 594, L115

Roth N., Porciani C., 2011, MNRAS, 415, 829, `arXiv:1101.1520`

Roth N., Porciani C., 2012, MNRAS, 425, L81, `arXiv:1205.3165`

Sales L. V., Navarro J. F., Abadi M. G., Steinmetz M., 2007, MNRAS, 379, 1464, `arXiv:0704.1770`

Sawala T., Frenk C. S., Crain R. A., Jenkins A., Schaye J., Theuns T., Zavala J., 2013, MNRAS, 431, 1366, `arXiv:1206.6495`

Scherrer R. J., Weinberg D. H., 1998, ApJ, 504, 607, `astro-ph/9712192`

Schmittfull M. M., Regan D. M., Shellard E. P. S., 2013, Phys. Rev. D, 88, 063512, `arXiv:1207.5678`

Schneider P., 2006, in Saas-Fee Advanced Courses, Vol. 33, Gravitational Lensing: Strong, Weak and Micro, Springer Berlin Heidelberg, pp. 269–451

Scoccimarro R., 2004, Phys. Rev. D, 70, 083007, `astro-ph/0407214`

Scoccimarro R., Colombi S., Fry J. N., Frieman J. A., Hivon E., Melott A., 1998, ApJ, 496, 586, `astro-ph/9704075`

Sefusatti E., 2009, Phys. Rev. D, 80, 123002, `arXiv:0905.0717`

Sefusatti E., Scoccimarro R., 2005, Phys. Rev. D, 71, 063001, `astro-ph/0412626`

Senatore L., Smith K. M., Zaldarriaga M., 2010, JCAP, 1, 28, `arXiv:0905.3746`

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1, `astro-ph/9907024`

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119, `astro-ph/9901122`

Sheth R. K., Tormen G., 2004, MNRAS, 350, 1385, `astro-ph/0402237`

Skrutskie M. F. et al., 2006, AJ, 131, 1163

Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, JCAP, 8, 31, `arXiv:0805.3580`

Smith K. M., Ferraro S., LoVerde M., 2012, JCAP, 3, 32, `arXiv:1106.0503`

Smith K. M., LoVerde M., 2011, JCAP, 11, 9, `arXiv:1010.0055`

Smith R. E. et al., 2003, MNRAS, 341, 1311, `astro-ph/0207664`

Smith R. E., Scoccimarro R., Sheth R. K., 2007, Phys. Rev. D, 75, 063512, `astro-ph/0609547`

Sofue Y., Rubin V., 2001, Annu. Rev. Astron. Astrophys., 39, 137, `astro-ph/0010594`

Sollima A., Valls-Gabaud D., Martinez-Delgado D., Fliri J., Peñarrubia J., Hoekstra H., 2011, ApJL, 730, L6, `arXiv:1102.2137`

Somerville R. S., Lemson G., Sigad Y., Dekel A., Kauffmann G., White S. D. M., 2001, MNRAS, 320, 289, `astro-ph/9912073`

Spergel D. N. et al., 2007, ApJS, 170, 377, `astro-ph/0603449`

Springel V., 2005, MNRAS, 364, 1105, `astro-ph/0505010`

Springel V. et al., 2008, MNRAS, 391, 1685, `arXiv:0809.0898`

Springel V., White S. D. M., 1999, MNRAS, 307, 162, `astro-ph/9807320`

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726, `astro-ph/0012055`

Stark D. P., Bunker A. J., Ellis R. S., Eyles L. P., Lacy M., 2007, ApJ, 659, 84, `astro-ph/0604250`

Stark D. P., Ellis R. S., Chiu K., Ouchi M., Bunker A., 2010, MNRAS, 408, 1628, `arXiv:1003.5244`

Starobinsky A. A., 1982, Physics Letters B, 117, 175

Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074, `astro-ph/0602350`

Suzuki N. et al., 2012, ApJ, 746, 85, `arXiv:1105.3470`

Tammann G. A., 1994, in ESOC, Vol. 49, ESOC, Meylan G., Prugniel P., eds., p. 3

Taruya A., Hiramatsu T., 2008, ApJ, 674, 617, `arXiv:0708.1367`

Taruya A., Koyama K., Matsubara T., 2008, Phys. Rev. D, 78, 123534, `arXiv:0808.4085`

Taruya A., Nishimichi T., Saito S., 2010, Phys. Rev. D, 82, 063522, `arXiv:1006.0699`

Tegmark M. et al., 2006, Phys. Rev. D, 74, 123507, `astro-ph/0608632`

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709, `arXiv:0803.2706`

Trotta R., 2008, Contemporary Physics, 49, 71, `arXiv:0803.4089`

Verde L. et al., 2002, MNRAS, 335, 432, `astro-ph/0112161`

Verde L., Matarrese S., 2009, ApJL, 706, L91, `arXiv:0909.3224`

Vielva P., Sanz J. L., 2010, MNRAS, 404, 895, `arXiv:0910.3196`

Vishniac E. T., 1983, MNRAS, 203, 345

Wagner C., Verde L., 2012, JCAP, 3, 2, `arXiv:1102.3229`

Wagner C., Verde L., Boubekeur L., 2010, JCAP, 10, 22, `arXiv:1006.5793`

Wald R. M., 1983, Phys. Rev. D, 28, 2118

Walker T. P., Steigman G., Kang H.-S., Schramm D. M., Olive K. A., 1991, ApJ, 376, 51

Wang J., Frenk C. S., Navarro J. F., Gao L., Sawala T., 2012, MNRAS, 424, 2715, `arXiv:1203.4097`

Weinberg S., 2008, Cosmology, Oxford University Press

Xia J.-Q., Baccigalupi C., Matarrese S., Verde L., Viel M., 2011, JCAP, 8, 33, `arXiv:1104.5015`

Yanny B. et al., 2003, ApJ, 588, 824, `astro-ph/0301029`

York D. G. et al., 2000, AJ, 120, 1579, `astro-ph/0006396`

Zehavi I. et al., 2011, ApJ, 736, 59, `arXiv:1005.2413`

Zel'dovich Y. B., 1970, AAP, 5, 84

Zel'dovich Y. B., 1972, MNRAS, 160, 1P

Zhang P., 2011, Phys. Rev. D, 83, 063510, `arXiv:1101.5164`

List of Figures

List of Tables

Acronyms

2LPT	Second-order Lagrangian Perturbation Theory	39
BE	Boltzmann Equation	12
BBN	Big Bang Nucleosynthesis	9
CAMB	Code for Anisotropies in the Microwave Background	15
CDM	Cold Dark Matter	4
CMa	Canis Major	86
CMB	Cosmic Microwave Background	4
DE	Dark Energy	2
DM	Dark Matter	2
EdS	Einstein-deSitter	9
EFE	Einstein Field Equations	5
EoS	Equation of State	6
FE	Friedmann Equation	6
FoF	Friends-of-Friends	39
FRW	Friedmann-Robertson-Walker	3
GR	General Relativity	2
GUT	Grand Unified Theory	22
ΛCDM	Lambda Cold Dark Matter	1
LSS	Large-scale structure	10
MO	Monoceros Overdensity	86
MW	Milky Way	83
NFW	Navarro-Frenk-White	40
PDF	Probability Density Function	50
PM	Particle-Mesh	35
PNG	Primordial non-Gaussianity	11
REE	Radial Energy Equation	93
r. m. s.	root mean square	38
SDSS	Sloan Digital Sky Survey	83
SO	Spherical Overdensity	39
SPT	Standard Perturbation Theory	17
SSIM	Self-similar Secondary Infall Model	99
ZA	Zel'dovich Approximation	38